Spread Spectrum Watermark Design under Noisy Compressive Sampling

Anirban Bose and Santi P. Maity

Department of Information Technology, Bengal Engineering and Science University, Shibpur
P.O. Botanic Garden, Dist-Howrah, Pin-711 103

ABSTRACT
This paper proposes an algorithm for spread spectrum watermark design under compressive sampling (CS) attack using hybridization of genetic algorithm (GA) and neural network. In watermarking application, CS may be viewed as a typical fading-like attack operation on the watermarked image. GA is used to determine the watermark strength taking into consideration of both robustness and imperceptibility in the paradigm of CS with additive white Gaussian noise (AWGN) attack channel. Then NN assisted improved detector is developed to classify two image classes i.e. watermarked and non-watermarked one. Simulation results demonstrate the effectiveness of the proposed method.

Key Words
Spread spectrum watermark, compressive sampling, Genetic algorithms, Neural Network.

1. INTRODUCTION
With the widespread use of internet and the development in computer industry, the digital media i.e. images, audios and videos become an integral part in our daily life. But digital multimedia content often suffers from infringement in the form of copyright violation, collusion, deterioration of content etc. To countermeasure these problems digital watermarking has emerged as one form of potential solutions over the last one and half decade. Digital watermarking is one way to hide a secret data (watermark) so that visual quality of the host image is not degraded much (imperceptibility requirement) and at the same time hidden data would not be removed after various common and malicious degradations (robustness requirement). To meet simultaneously both these conflicting requirements, spread spectrum (SS) watermarking becomes popular where watermark is spread throughout the spectrum of the host image [1]. However, SS watermark system suffers from poor detection performance for the widely used correlator in presence of non-stationary fading-like gain attack [2]. Performance is further affected when watermarked signal is reconstructed from the lower measurement spaces i.e. from under sampled version.

In recent years compressed sampling (CS) has attracted much attention due to reconstruction of broadband signals in real time application. The principle of operation in CS technique lies with the fact that reconstruction of signal to be done from its sparse representation, on some basis, by taking only a small number of measurements which is incoherent relative to the sparsity basis. In watermarking application, CS may be viewed as a typical fading-like attack operation on the watermarked image. Design of robust system in such fading attack channel becomes crucial. One design criteria is the proper selection of the watermark strength (α). On communication theoretic perspective, calculation of watermark strength may be treated as optimal power allocation problem in fading channel under power constraint scenario i.e. imperceptibility requirement in watermarking application.

In CS, reconstruction of signal is done from low measurement space using sensing matrix which takes values using independent and identically distribution (i.i.d). Faithful recovery of watermark signal from these reconstructed image through the calculation of proper watermark strength using analytic approach has no real and own solution. Soft computing tools like genetic algorithm (GA), fuzzy logic, artificial neural network (ANN) look appealing to determine low cost, tractable, robust and optimal (or sub-optimal) solution to a constrained optimization problem. To this aim, GA is used lot to determine an optimum watermark strength so that a balance between the two conflicting requirements i.e. robustness and imperceptibility, could be achieved. At the same time artificial neural network (ANN) is used in detection phase to classify the unknown set of images into watermarked and non-watermarked images. ANN represents a highly parallelized dynamic system with a directed graph topology which can generate output signal by means of reaction of its states on the input signal. Architecturally, an ANN is an ensemble of interconnected artificial neurons that are generally organized into layers, namely input layer, output layer, hidden layer.

This paper proposes an improved technique for embedding as well as detecting a SS watermark under CS attack using a hybrid GA-Neuro (GA-NN) model. GA is used to determine the watermark strength taking into consideration both robustness and imperceptibility in the paradigm of compressed sensing and additive white Gaussian noise (AWGN) attack channel. For classification purpose, a two layer feed-forward ANN structure with back propagation is used to classify two image classes i.e. watermarked and non-watermarked.

Rest of the paper is organized as follows. Section II expresses the basic idea of compressive sampling. A very brief discussion on spread spectrum watermarking is made in Section III. Section IV presents watermark strength calculation using GA. Section V presents NN assisted watermark detection. Section VI illustrates the simulation results. Finally this paper is concluded in Section VII along with the scope of the future works.

2. Brief introduction to Compressive Sampling
This section briefly introduces the CS theory. Let us assume that an unknown signal $x_0 \in R^N$, could be sparsely represented in a certain domain by the transform matrix Ψ. In other words, if there are only K non-zero coefficients in the Ψ domain, we can say that $x_0$is K-sparse. The purpose of CS is to recover the sparse signal $x_0$by taking random measurements less than N [3]. In order to take CS measurements, we first let Φ denote an M by N matrix with M<<N. The measurement matrix Φ should be uncorrelated with the transform matrix Ψ. M measurements are obtained by a linear system

$$y = \Phi x_0 \quad (1)$$

where the symbol y is the number of measurements. Different from the traditional sampling, each measurement from CS measures the whole information of the signal. Therefore, each CS measurement contains a little information from all positions of the original signal. The CS theory states that the signal could be recovered exactly if the number of measurements M obeys the condition $M > Const.K.\log N$. The constant 'Const' is an over-measuring factor which is more than 1. Since y is lower dimension vector compared to x_0, it is impossible to get exact x_0 directly by the inverse transform of (1). The reconstruction has to be done as solving the following optimization problem

$$(L_1) \quad min\|\Psi^T x\|_1 \ subject\ to\ y = \Phi x \quad (2)$$

The reconstructed signal x is among all signals generating the same measured data, which has transform coefficients with the minimal l1-norm (L1) [4]. The reconstruction procedure can be solved as a linear programming problem.

3. Spread Spectrum Watermarking

Watermark embedding using SS is followed as was done in Cox et. al. method [1]. The structure of the watermark is chosen to be identical independent distributed (i.i.d) Gaussian patterm. The watermark is embedded in low frequency AC coefficients of DCT domain of the 2-D image signal, following the equation,-

$$\hat{X}_l^{DCT} = X_l^{DCT} + \alpha.\mathcal{W}_l\,;\ 1 \le l \le k\,; \quad (3)$$

X_l^{DCT} is the zig-zag scanned DCT coefficients of the original image ($I_{N\times N}$), excluding the DC coefficient.

$\hat{X}_l^{DCT}$ is the modified AC coefficients.

α is the watermark strength to be determined by GA and

k is the number of largest AC coefficients. Watermark detection is done based on correlation detection, although one may use other detector structure as suggested in [5].

4. Watermark strength calculation using GA

GA, a concept based on natural genetics, has long been deployed in numerous application where an optimal (or sub-optimal) solution to a given optimization problem is required [6]. GA works on the encoding of the parameters not on the parameters itself and the elements in the encoded stream, or the "genes", are adjusted to minimize or maximize the fitness value. In the present application, GA is used to calculate watermark embedding strength (α) values selected from the initial population of α. So we get a class of watermarked images i.e.

$$\Omega = \{I_{\alpha_1}^W, I_{\alpha_2}^W, \dots, I_{\alpha_r}^W\};\ 1 \le \alpha_i \le r \quad (4)$$
$$r\ is\ the\ total\ number\ of\ \alpha\ in\ a\ population$$

Watermark extraction is performed after each watermarked image has undergone CS attack. To design the fitness function for GA routine, both the imperceptibility criteria and the robustness are considered. Structural similarity index (SSIM) is used as imperceptibility measure, while the value of normalized correlation coefficient (NCC) is used for robustness measure. The fitness function of GA incorporates both to select the best individual i.e. α value. The fitness function of GA is taken to be as,-

$$F_j = \max_j(\lambda_1.SSIM_j + \lambda_2.NCC_j); 1 \le j \le r \quad (5)$$
$$\lambda_1\ and\ \lambda_2\ are\ weighing\ factor$$

The GA routine is initialized with the fitness function shown above and is run for 40 generation. Fig. 1 shows the schematic diagram of the GA routine to determine optimum α value. The value of α for which the maximum fitness value is obtained is used as watermark embedding strength. The input image, $I_{N\times N}$, is watermarked with this value of α and then the watermarked image is prepared for detection phase by ANN.

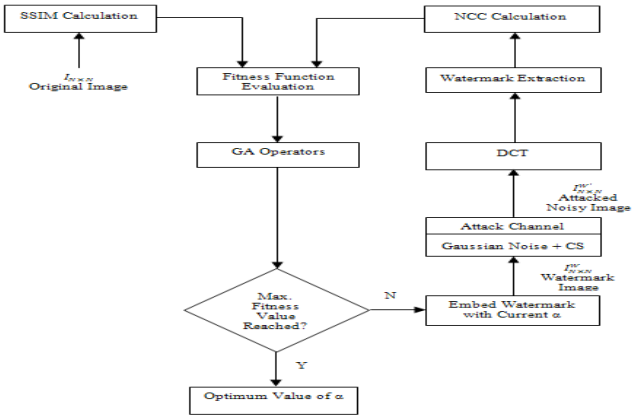

Fig. 1 Block diagram of GA routine for embedding i.i.d Gaussian watermark

V. Neural Network Assisted Watermark Detection

A feed forward back propagation ANN is used as a classifier system to detect the presence or absence of watermark within an image [7]. In order to derive the data set for training the network, feature vectors for both the original (host) image and the attacked watermarked image are determined from their gray level co-occurrence matrix (GLCM) [8]. The feature set selected is {correlation, contrast, energy, and entropy}. The feature vectors thus formed are used as inputs to the network as a congruent feature vector to train the network. This single feature vector works as the dataset for the network. This dataset is partitioned into training, validation and test data spaces at random. The network is trained with the training data and to increase the accuracy of the network validation data is used. Test data is used to check the output of the network and the extent to which the errors are produced by it. One of the major concerns in designing a NN system is to determine the optimal number of hidden neurons. Our empirical study is out by keeping the sample space constant and observing the effect of change in the number of hidden layer neurons to the output of the network. It decides that the optimum number of neurons in the hidden layer would be 15. Fig. 2 below portrays the result of our empirical study.

The performance measure taken for the network is mean square error (MSE) and to ensure faster learning as well as the minimum computational time, gradient descent algorithm with moments is selected. Tan-sigmoid function acts as the activation function for each neuron in the hidden as well as output layer. The network is initialized with random weights and zero biases.

The overall pictorial representation of the proposed neural network assisted watermark detection is shown in the Fig. 3.

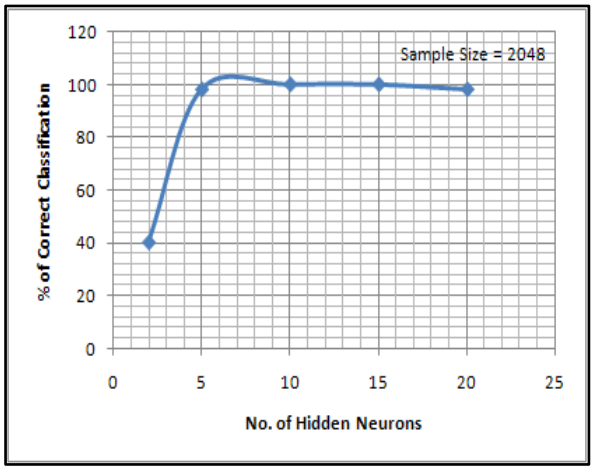

Fig 2. Detection probability for different hidden layer size

5. Performance Evaluation and Discussion

We study the performance of our proposed method by carrying out simulations over large number of images. One such test image Lena with size 256×256 is shown in Fig. 4. An i.i.d Gaussian watermark, as shown in Fig. 5, is embedded in the test image. The watermark is embedded in the low frequency coefficients of the DCT domain excluding the DC coefficient. After embedding the watermark, inverse DCT is performed to obtain the watermarked image.

A. Attack Channel: CS plus AWGN

As mentioned earlier, in this work, CS is considered as a typical watermark attack followed by addition of noise signal. In other words, the watermark image undergoes noisy CS attack and faithful recovery/detection of watermark is important. To simulate this, the watermarked image is then undergone through a CS operation and AWGN attack channel having noise variance (NV) of 0.01.

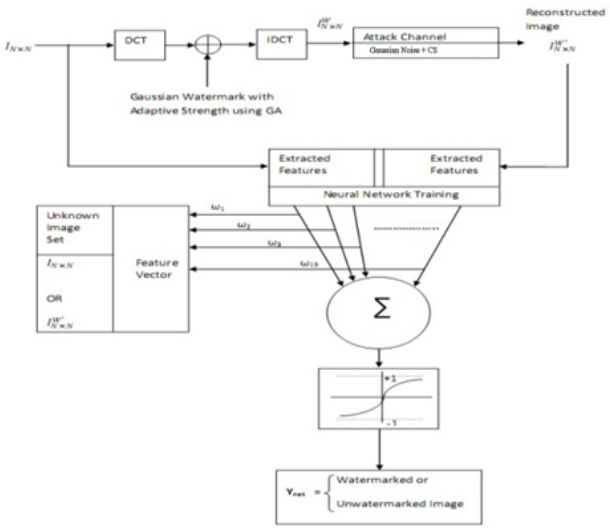

Fig. 3.Schematic representation of the proposed neural model

VI. Results and Discussion

Fig. 4 Test image *Lena*

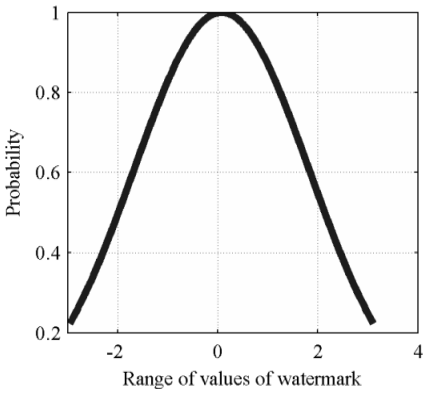

Fig. 5.i.i.d Gaussian Watermark

Wavelet transform is used for sparse representation of the image signal and the measurement matrix is generated by randomly sampling the normal distribution having zero mean and unity standard deviation $\mathbb{N}(0,1)$.

The image $I_{N\times N}^{CS}$is decomposed into various frequency bands applying the transform.

$$X_{N\times N} = \Psi_{N\times N} I_{N\times N}^{CS} \quad (6)$$

Hence, X would have approximation, horizontal, vertical, diagonal set of frequencies as seen in normal DWT decomposition and is shown in the Fig 5.

APPROXIMATION(A)	**HORIZONTAL(H)**
VERTICAL(V)	**DIAGONAL(D)**

Fig 6.Decomposition of Image into various frequency bands

Leaving the approximation coefficients, CS measurements are carried out from all the other subbands coefficients. Only 44.85% of the total coefficients are kept for reconstruction and rest are discarded.

$$Y^{p}_{k\times 1} = \Phi_{k\times N} X^{p}_{N\times 1} \; ; where\ p = \{H, V, D\} \quad and\ 1 \leq k \leq M; M \ll N \tag{7}$$

Each measurement vector $Y^{p}_{k\times 1}$ is formed by taking measurement from each column of the horizontal, vertical and diagonal matrix, one at a time. After that each lower dimension vector is reconstructed from the low measurement space, by minimizing the l1-norm (L1) and stored in a temporary buffer matrix. This process is repeated for all the three detail i.e. horizontal, vertical and diagonal coefficients to get complete watermarked reconstructed image ($I^{W'}_{N\times N}$). Perturbation introduced to the image because of this CS attack is determined by calculating peak signal to noise ratio (PSNR) of the image.

Fig. 7 below displays the noise corrupted Lena image and Fig. 8 shows the effect of CS attack on the noise corrupted image. Next, SSIM and NCC values of the resulting watermarked image are evaluated to form the fitness function of the GA routine. In the GA training process, ten individuals are used for each generation run, with crossover probability 0.8, mutation probability 0.02. The performance of the GA routine is depicted in Fig. 9.

At the end of GA run we get the optimum α and the watermarked image generated with this α is shown in Fig. 10. The selection of λ_1 and λ_2 would influence both the watermark image quality and robustness measure. In this work both have been given the corresponding weightage i.e. λ_1= 3 and λ_2= 2.5. GA based watermark embedding along with CS attack for the test image Lena is given in Table 1.

Fig. 7.Watermarked Image with AWGN with noise variance = 0.01 (PSNR=35.1757dB)

Fig. 8.CS attacked noisy Watermarked Image (PSNR=30.0501dB)

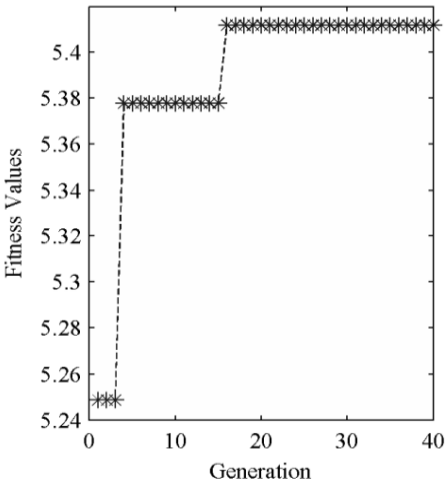

Fig. 9.Performance curve of GA routine

Fig. 10.Watermarked Image with α=5.9235 and PSNR=50.4403

Table 1: Perceptual quality of the watermarked image

Gene-ration Run	Opti-mum Value of α	PSNR of Water-marked image (dB)	SSIM of the water-marked image	PSNR of the noisy water-marked image (dB)	PSNR of the noisy water-marked and CS attacked image (dB)
40	5.9235	50.4403	0.9989	25.7198	25.1481

B. Simulation results

In the classification process, the four textural features mentioned previously is calculated to build up the sample data for the NN. These features are calculated from the gray level co-occurrence matrix (GLCM) of the attacked watermarked image ($I_{N\times N}^{W'}$). Basically the GLCM is measured based on two parameters, the distance between pixel pairs (d) measured in pixel number and their orientation (φ) which is quantized into four directions (0o,45o,90o,135o), hence for each d resulting values for the four φ's are averaged out [9]. But for an image having gray tone level [0,L] computation of GLCM would result in a matrix of dimension L×L×p , where p is array representing the orientations. In view of the above the gray levels of both the watermarked and unwatermarked images are scaled down to 128 levels in lieu of the whole 256 gray levels and for each d, 64 measurements are taken. Therefore, for an image having gray tone level [0,255] a GLCM of 128×128×256 is obtained in place of 256×256×256. Now each GLCM produces 1×256 array of a single feature, hence, for each watermarked and unwatermarked image a feature matrix of 4×256 is calculated. Thus a feature matrix of dimension of 4×512 is obtained and this works as the sample space for the NN [10]. The whole sample space is partitioned into three data sets i) training data, ii) validation data and iii) test data in a random proportion. After the training phase the net is presented with the validation data set. The primary function for validation set is to minimize the over fitting. Adjustment of weights of the network is not done with this data set. It only verifies that any increase in accuracy over the training data set actually yields an increase in accuracy over a data set that has not been shown to the network before. Finally the network is presented with test data set in order confirm the actual predictive power of the classifier. Fig. 11 shows the performance of the network against the aforementioned three sets of data.

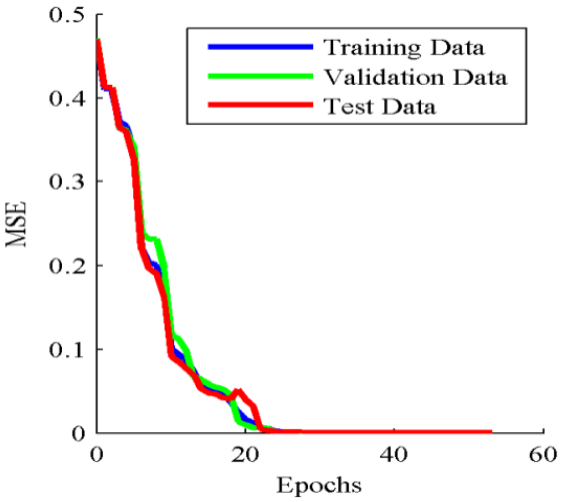

Fig. 11.Performance of the Neural Network for three sets of data

The outcome or the actual predictive power of the network, once again for test image Lena, is shown in Table 2.

Table 2: Watermark detection performance using NN

Image Name	Dimension	Total Testing Samples	% of Correct Classification	% of Incorrect Classification
Lena	256×256	77	100.00	0.00

The receiver operating characteristic (ROC) of the detector/classifier describing the quality of the detector/classifier is shown in the next figure. To obtain the ROC of the neural classifier first threshold values are applied across the interval [0, 1] to the output nodes. For each value of threshold, any feature vector which produces an output greater than or equal to the threshold is classified as a target (watermarked image), otherwise it is classified as a non-target (un-watermarked image). Now the percentage of watermarked images that are correctly classified as watermarked images gives out the true positive rate (TPR). On the other hand, the false positive rate (FPR) is the percentage of watermarked images incorrectly classified as unwatermarked images. The plotting of TPR and FPR against threshold values represents the ROC curve for the network. Fig. 12 shows the ROC of the neural network. Although it is a well-known fact that a perfect test would show points in the upper left corner, with 100% TPR and 0% FPR but still it is clearly evident, from the figure that the network performs with greater accurately. An empirical study is made by changing noise variance and observing the prediction capability of the proposed model to see the effect of noise on the performance of the system. It is found that at lower noise variance the detector performance is dropped, although not by any significant amount, but to a certain extent.

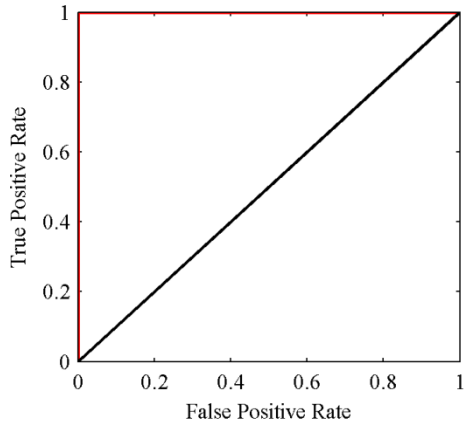

Fig. 12.ROC of the Neural Network

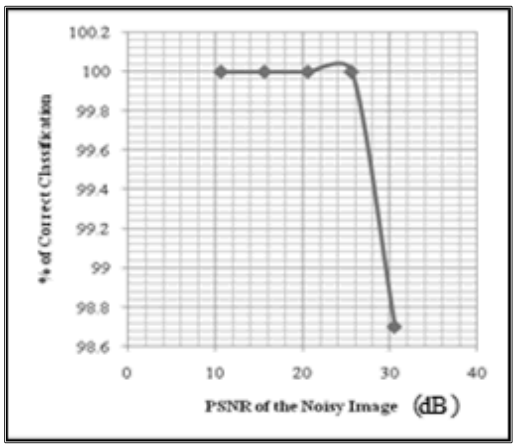

Fig 13.Classifier performance against varying PSNR of Image

The proposed hybrid GA-NN model is tested for images, taken from standard image processing database i.e. SIPI Image

Database, proves the excellence of the model. Due to the space constraint few of the results are shown in Table 3 (contribution of GA operation) and 4 (contribution of NN operation), respectively.

Table 3: Watermark embedding performance using GA

Test image	Opti-mum value of α	PSNR of water-marked image (dB)	SSIM of the water-marked image	PSNR of the noisy water-marked image (dB)	PSNR of noisy CS water-marked and image (dB)
Camera-man	3.3224	54.9923	0.9992	25.6156	24.4753
Mandrill	8.4859	47.6699	0.9987	25.5262	23.7825
Rice	8.8353	47.3196	0.9986	26.6119	26.3483
Moon	8.8402	47.7166	0.9970	25.8219	25.3327

Table 4: Watermark detector performance using NN

Image Name	Dimension	Total Testing Samples	% of Correct Classification	% of Incorrect Classification
Camera	256×256	77	100.00	0.00
man				
Mandrill	256×256	77	100.00	0.00
Rice	256×256	77	100.00	0.00
Moon	256×256	77	100.00	0.00

6. Conclusions and Scope of Future Works

In this paper a new hybrid GA-NN model is developed and tested for efficient SS watermark design under noisy CS attack. Use of GA in determining the optimum value of the embedding strength has shown significant improvement in visual quality of the watermarked images. On the other hand NN assisted watermark design is capable of tackling the adversary attack channel effect. Despite the substantial loss of information because of the reconstruction of the watermarked image from lower measurement space, NN based detector shows almost flawless detection of the watermark. Future work may extend the proposed algorithm for tamper detection and recovery of the image signal. Future work should also include the characteristics of human visual system (HVS) in designing improved visual quality of the watermarked image and at the same time for other receiver, for example, MMSEC type for improved watermark decoder performance.

7. REFERENCES

[1] I.J. Cox, J. Kilian, F.T. Leighton and T. Shamoon, "Secure Spread Spectrum Watermarking for Multimedia," IEEE Trans. on Image Proc., vol. 6, pp.1673-1687, 1997.

[2] S. P. Maity and S. Maity, "Multistage Spread Spectrum Watermark Detection Technique using Fuzzy Logic", IEEE Signal Proc. Letters, vol. 16, no. 4, pp.245-248, 2009.

[3] D. Donoho, "Compressed Sensing," IEEE Trans. Inform. Theory, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[4] E. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," IEEE Trans. Inform. Theory, vol. 52, no. 2, pp. 489–509, Feb. 2006

[5] J.R. Hernandez, M. Amado, F. Perez-Gonzalez, "DCT-domain watermarking techniques for still images: detector performance analysis and a new structure", IEEE Trans. Image Proc., vol. 1, pp.55-68, 2000.

[6] D.E. Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley, Reading, MA, 1992.

[7] Laurene V. Fausett, "Fundamentals of Neural Networks: Architectures, Algorithms, and Applications", 9 December, 1993.

[8] Brian D. Ripley, "Pattern Recognition and Neural Networks", ISBN-10: 0521460867, 18 January, 1996.

[9] Baraldi, F. Parmiggiani, "An Investigation of the Textural characteristics associated with GLCM Matrix Statistical Parameters",IEEE Trans. on Geos. and Rem. Sens., vol. 33(2), pp. 293-304, 1995.

[10] R. Haralick, K. Shanmugam, I. Dinstein, "Texture Features For Image Classification", IEEE Transaction on SMC-3(6), pp. 610-621, 1973.