A Novel Approach to Enhance Robustness in Steganography Using Multiple Watermark Embedding Algorithm

S. Praveen Kumar, K. Anusha, R.Venkata Ramana

***Abstract*- This paper deals with an enhanced technique that improves robustness in steganography. A multiple watermark embedding algorithm has been proposed to embed multiple text messages as watermarks simultaneously in the same watermark space while minimizing the watermark (distortion) energy. The watermark is embedded in the DCT coefficients of the green channel of the color image. The algorithm takes into account the perceptual capacity of each coefficient inside the DCT blocks before embedding the watermark information. Therefore, the first 16 low frequency coefficients (excluding the DC value) in the 8×8 DCT block was screened and the eight coefficients with the maximum magnitudes were selected for embedding. The algorithm used is blind and does not require the original image for extracting the watermark. The watermarking method is robust against JPEG compression, additive noise, cropping, scaling, low-pass and median filtering.**

***Index Terms*- Steganography, DCT, MWE, pseudo random bit sequence, pseudo random positive real numbers, zero –mean Gaussian with variance.**

I. INTRODUCTION

In this highly digitized world, the internet services as an important role for data transmission and sharing. However, since it is a world-wide and publicized medium, some confidential data might be stolen, copied, modified or destroyed by an unintended observer. Therefore, security problems become an essential issue. Encryption is a well-known procedure for secure data transmission. The frequently used encryption methods include RSA and DES algorithms. Although these two methods do achieve certain security effects, they make the secret messages unreadable and unnatural. These unnatural messages usually attract some unintended observers attention. This is the reason a new security approach called "STEGANOGRAPHY" arises.

Manuscript received March 10, 2011

S.Praveen Kumar, Department of Information Technology, GITAM University, Visakhapatnam, Andhra Pradesh, India, 9989590690., praveen@gitam.edu.

K.Anusha, Department of Information Technology, GITAM University, Visakhapatnam, Andhra Pradesh, India, 9951971641., anusha.keri@gmail.com.

R.Venkata Ramana, Department of Information Technology, GITAM University, Visakhapatnam, Andhra Pradesh, India, 9703450094., ragolu.venkataramana13@gmail.com

Steganography is derived from the Greek for covered writing and essentially means "to hide in plain sight". As defined by Cachin [1] steganography is the art and science of communicating in such a way that the presence of a message cannot be detected. Simple steganographic techniques have been in use for hundreds of years, but with the increasing use of files in an electronic format new techniques for information hiding have become possible. Steganography can be used to hide a message intended for later retrieval by a specific individual or group or for copyright marking, where the message to be inserted is used to assert copyright over a document which can be further divided into watermarking and fingerprinting.

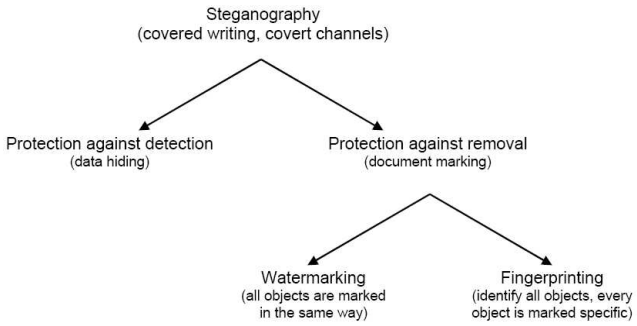

Fig 1: Types of Steganography

In this paper we propose a method to hide the text image and also protect it from modifications using MWE(Multiple Watermark Embedding) algorithm.

Almost all digital file formats can be used for steganography, but the formats that are more suitable are those with a high degree of redundancy. Redundancy can be defined as the bits of an object that provide accuracy far greater than necessary for the object's use and display. The redundant bits of an object are those bits that can be altered without the alteration being detected easily. Image and audio files especially comply with this requirement, while research has also uncovered other file formats that can be used for information hiding. Given the proliferation of digital images, especially on the Internet, and given the large amount of redundant bits present in the digital representation of an image, images are the most popular cover objects for

steganography. In the domain of digital images many different image file format exit, most of them for specific applications. For these different image file formats, different steganographic algorithms exist. Among all these file formats, the JPEG file format [3] is the most popular image file format on the Internet, because of the small size of the images.

Our algorithm for colour digital invisible image watermarking is carried out in the frequency domain. A DCT block based approach has been applied to the green channel of the RGB image [3].The text messages are converted to images using their ascii values and embedded as watermarks simultaneously in the host JPEG image in the selected pixels using MWE algorithm.

II. THE PROPOSED METHOD

The proposed algorithm takes into account the perceptual capacity of each coefficient inside the DCT blocks before embedding the watermark information. Therefore, the first 16 low frequency coefficients (excluding the DC value) in the 8×8 DCT block was screened and the eight coefficients with the maximum magnitudes were selected for embedding. This range of frequencies is chosen because the high frequency components may be discarded in some image processing operation such as JPEG compression. Each DCT block consists of 8×8 coefficients. The 16 lower frequencies are screened to find eight coefficients with the highest magnitude and register its locations. This process is repeated for all DCT blocks. The locations which are repeated more are selected. These locations will vary from one image to another according to the spatial frequency contents of the image. Eight binary bits of the watermark will be embedded in these locations. A flow graph of the coefficients selection process is shown in Fig.2. In order to test the security of the coefficients selection process the images were screened again after embedding to verify that the method is secure and an attacker would not be able to use the selection process again to detect the originally selected locations. Screening the DCT blocks again after embedding will result in totally different locations from the previously registered locations in the original unwatermarked images.

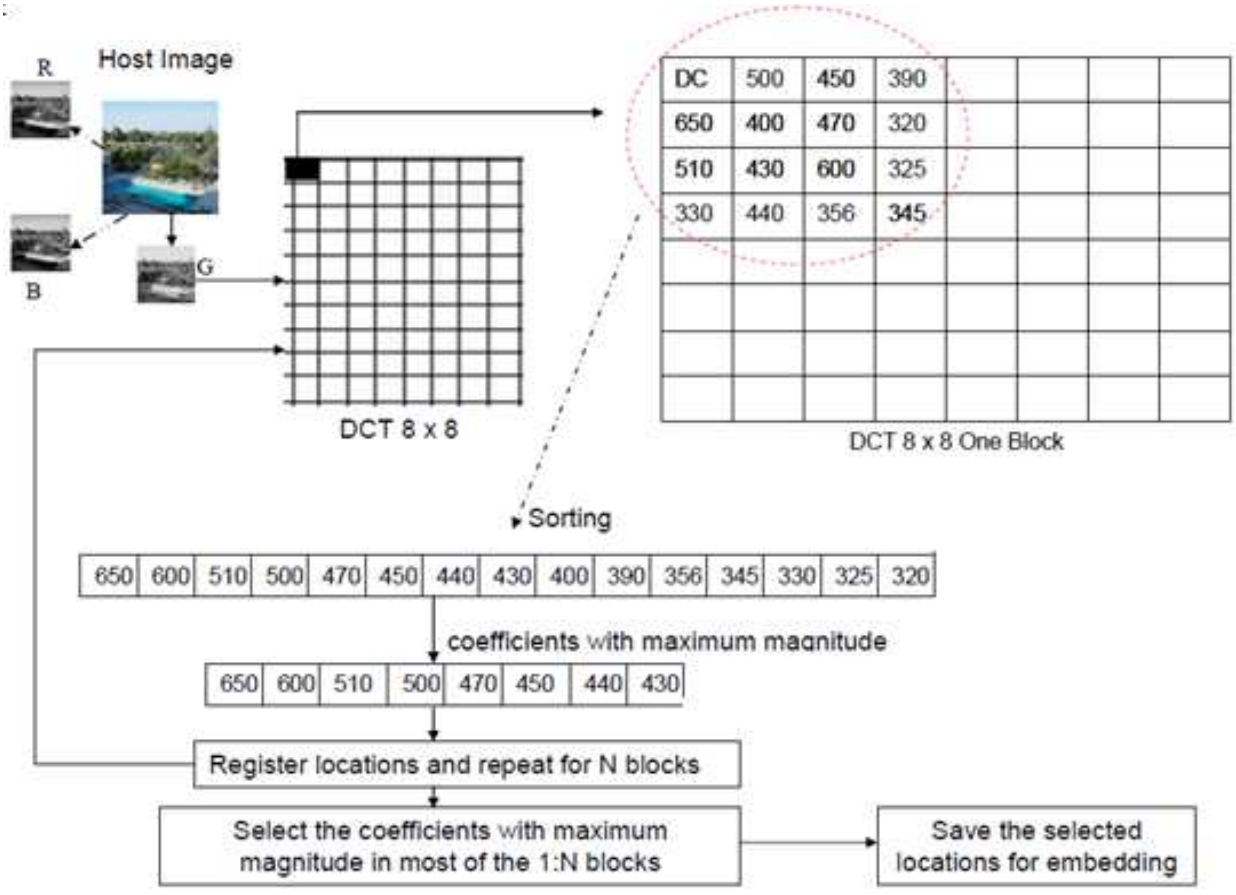

Fig 2: A flow graph of the coefficients selection process

III. MULTIPLE WATERMARK EMBEDDING ALGORITHM

A multiple watermark embedding algorithm[4] has been proposed to embed multiple text messages as watermarks simultaneously in the same watermark space while minimizing the watermark (distortion) energy These

pixels that have been selected in DCT are stored as host vector. Let the host vector be $\mathbf{H=[h_1,h_2,.....h_L]}$ with length L. The text document to be embedded is converted into set of images whose size N is less than L. Let the no.of images be Q and are denoted as $\mathbf{I_n}$ where $\mathbf{0<n<Q}$.The watermarks are

modulated with a pseudorandom bit sequence $\mathbf{s=[s_1,s_2,.....s_N]}$ with $\mathbf{s_i \in \{0,1\}}$ as

$$w_{i,n} = s_i \oplus I_{i,n} \quad (1)$$

where $\mathbf{0 < i \le N}$, $\mathbf{0<n<Q}$

Divide the host vector into N sub vectors of equal length. Then we embed each bit in all watermarks in each subvector using two keys. The first key is Q sets of pseudorandom positive real numbers denoted as $\mathbf{R_1,R_2,.....R_Q}$ where $\mathbf{R=[r_1,r_2,.....,r_N/r_i \in R^+ \le i \le N]}$ each of length N. The second key denoted as $\mathbf{K_1,K_2.....K_Q}$ of length L with each $K_{j,l}$ being zero –mean Guassian with variance .

The i^{th} bit of j^{th} watermark sequence is denoted as $w_{j,i}$.The watermarked i^{th} subvector, denoted as H_i' is

$$H_i' = H_i + \alpha_{1,i}K_{1,i} + \alpha_{2,i}K_{2,i} + \ldots + \alpha_{Q,i}K_{Q,i}, \quad (2)$$

The scaling factors form a row vector $V_i = [\alpha_{1i}, \alpha_{2iI}, \ldots\ldots\ldots \alpha_{Q,i},]$ with $\alpha_{j,i}, \in R$

We derive the scaling factors based on two conditions.The first condition is that the projection of H_i' onto the direction of Kj,i corresponds to the correct watermark bit as

$$\left[\text{Round}\left(\frac{\langle H_i', K_{j,i}\rangle}{r_{j,i}}\right)\right]\%2 = w_{j,i}, \quad 1 \le j \le Q. \quad (3)$$

The second condition is that the distortion or the Euclidean distance E_w between Y_i and Y_i' is minimized. The Euclidean distance is also the energy of the watermark and is equal to

$$E_w = \| H_i' - H_i \|_2^2 = V_i C_i V_i^T \quad (4)$$

Where

$$C_i = \begin{pmatrix} \|K_{1,i}\| & \langle K_{1,i}, K_{2,i}\rangle & \ldots\ldots & \langle K_{1,i}, K_{Q,i}\rangle \\ \langle K_{2,i}, K_{1,i}\rangle & \|K_{2,i}\| & \ldots\ldots & \langle K_{2,i}, K_{Q,i}\rangle \\ \vdots & \ddots & \ddots\ddots & \vdots \\ \vdots & \ddots & \ddots\ddots & \vdots \\ \langle K_{Q,i}, K_{1,i}\rangle & \langle K_{Q,i}, K_{2,i}\rangle & \ldots\ldots & \|K_{Q,i}\| \end{pmatrix} \quad (5)$$

Substituting (2) into (3), Q simultaneous equations can be obtained.They are, in matrix form

$$\mathbf{V_iC_i=B_i=[b_{1,i},b_{2,i},.....b_{Q,i}]} \quad (6)$$

With

$$b_{j,i} = r_{j,i} \cdot (2 \cdot x_{j,i} + w_{j,i}) \; - \langle H_i, K_{j,i}\rangle, \quad 1 \le j \le Q \quad (7)$$

where $x_{j,i}$ is an integer for any i,j . If all the $x_{j,i}$ are determined, the row vector B_i can be computed using (7) and the scaling vector V_i can then be obtained as from (6). It is important to choose the integers $x_{j,i}$ such that the E_w is as small as possible. By substituting (6) into (4), the E_w can be rewritten in terms of C_i and B_i as

$$E_w = B_i C_i^{-1} B_i^T \quad (8)$$

An approach called iterative approach(IA)[4] is used to choose $x_{j,i}$, which is more suitable for the general case of nonorthogonal randon key vectors.

A. A. Decoding Process:

To decode the watermark bits we extract the watermark vector Y' from the test image and segment it into N subvectors. The i^{th} modulated watermark bit w_i' is decoded from the subvector Y_i' as

$$w_i' = \text{round}(\langle H_i', K_i\rangle / r_i)\%2 \quad (9)$$

using the two keys and the i^{th} demodulated watermark bit is obtained as

$$I_i' = w_i' \oplus s_i \quad (10)$$

Watermark detection is performed in a similar way by decoding the watermark bit sequence and computing the average detection scores s_1 or s_2 [4]

Then the watermark obtained is converted to text. Each 8 bits in the image are converted from binary to decimal and based on the ascci values the decimal values are converted to text. Thus we obtain the text message.

IV. EXPERIMENTAL RESULTS

We tested the proposed algorithms on many testing images as shown in Fig. 3. All the images are 512×512 pixels (e.g., see Fig. 4) and only the luminance components are used. Six 10×8 binary images, as shown in Fig. 5, are used as perceptual meaningful watermarks in the experiments. The binary images are raster-scanned to form 1-dimensional bit sequences and modulated with a pseudorandom binary sequence. We simulate MWE in the DCT domain. The whole original image is transformed to the (512×512) DCT domain and scanned in a zigzag order. We use the first 8 high magnitude low frequency DCT AC coefficients to form the host vector to embed the watermark. The length of the host vector is H= 8×64=512(no of pixels in selected 8×8 blocks in host image). Based on the host vector length we can embed 64 characters as each character requires eight bits. The length of each watermark we used is 8×10= 80. The length of the subvector is floor(512/80)=6. We choose low-frequency components of DCT to form the host vector as these components tend to have large energies such that the embedded watermarks tend to be robust against different kinds of attack.

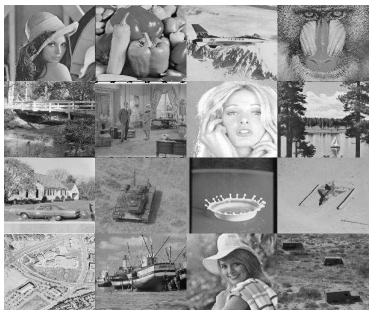

Fig. 3. Testing images used in experiments

Fig. 4. Original 512×512 "Lena'".

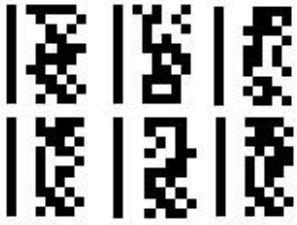

Fig. 5. Watermarks

The six watermarks in Fig. 5 are embedded with MWE using IA-R. A typical MWE-watermarked image (IA-R with Q=5) is shown in Fig. 6. With PSNR of 44.94 dB, the IA-R image has very good visual quality. The sample distributions of $S_1$of MWE (IA-R,Q=5) watermark detection are shown in Figs.7,9,11 for JPEG attack, LPF, and noise attack, respectively. The corresponding total detection errors (sum of type 1 and 2 error probability) against different detection thresholds are shown in Figs. 8, 10, 12.

In the JPEG attack on MWE in Fig. 7, the distribution of "no watermark" intersects with the distribution corresponding to SF=3. As a result, no threshold can give zero detection error probability at SF=3 in Fig.8. A typical example of MWE under JPEG attack (SF=2 , zero error) is shown in Fig. 13. Similar observations can be made for the LPF attack on MWE in Figs. 9,10, and 14. With five watermarks embedded, the sample distributions of "watermarked" and "no watermark" are intersecting slightly. Error can occur in some cases. An example of MWE with no error is shown in Fig. 14. In the noise attack on MWE in Fig. 11, the distribution of "no watermark" intersects with many distributions and thus no zero detection error is possible in the corresponding situations in Fig. 11. Fig. 15 is an example of MWE when no error occurs. Fig.16 is an example of the MWE under print-and-scan attack. By observing these results, more embedded watermarks tend to result in lower robustness.

Fig.6. MWE-watermarked image with five watermarks embedded (Q = 5, IA-R). PSNR = 44:94 dB.

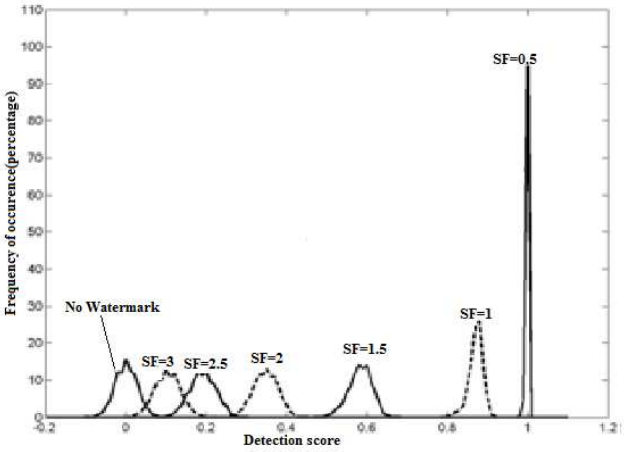

Fig. 7. Distribution of S of MWE (IA-R,Q = 5) under JPEG attack.

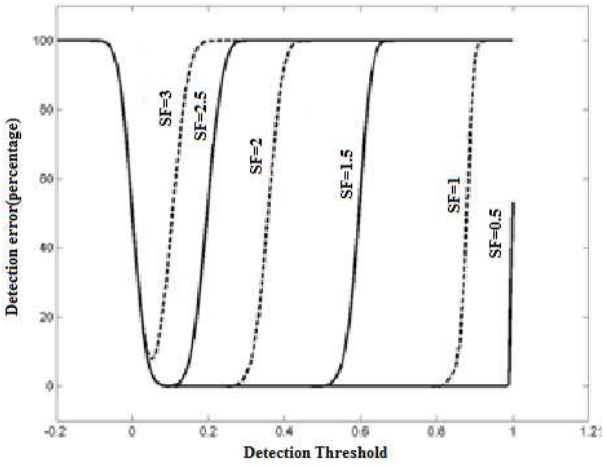

Fig. 8. Detection error of MWE under JPEG attack.

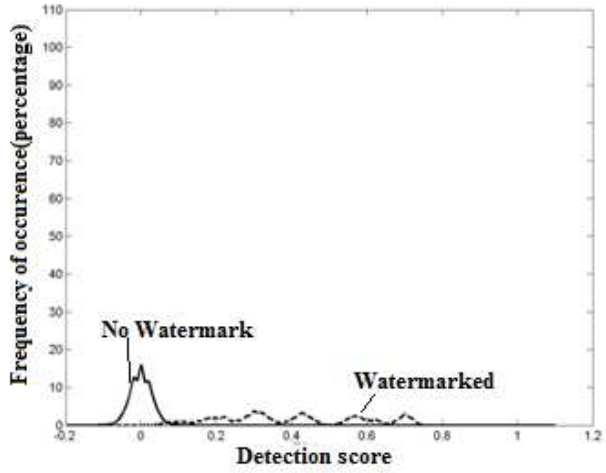

Fig. 9. Distribution of S of MWE under LPF attack.

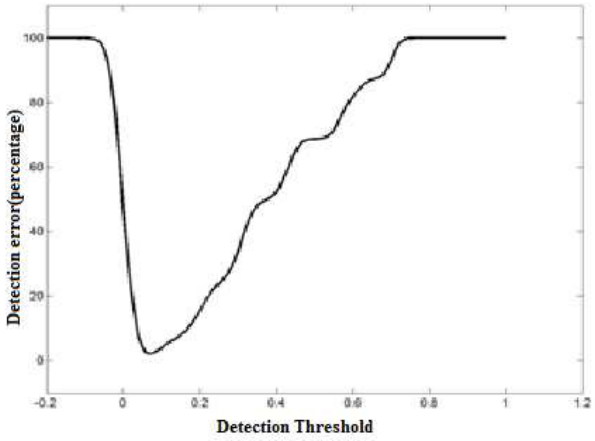

Fig. 10. Detection error of MWE under LPF attack.

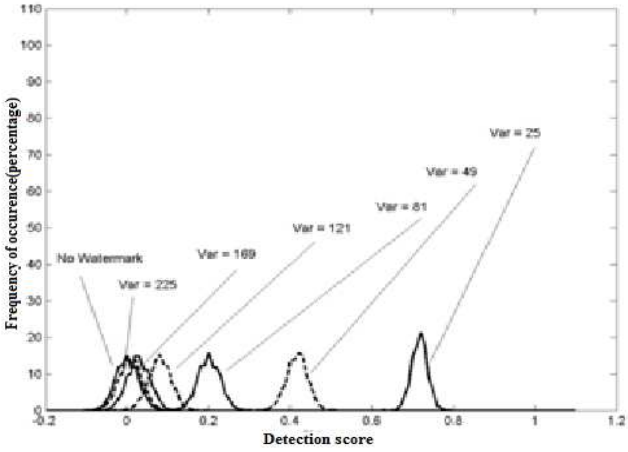

Fig. 11. Distribution of S of MWE under noise attack.

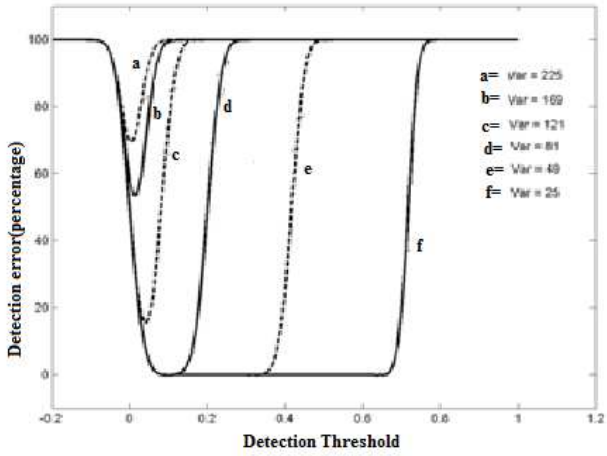

Fig.12. Detection error of MWE under noise attack.

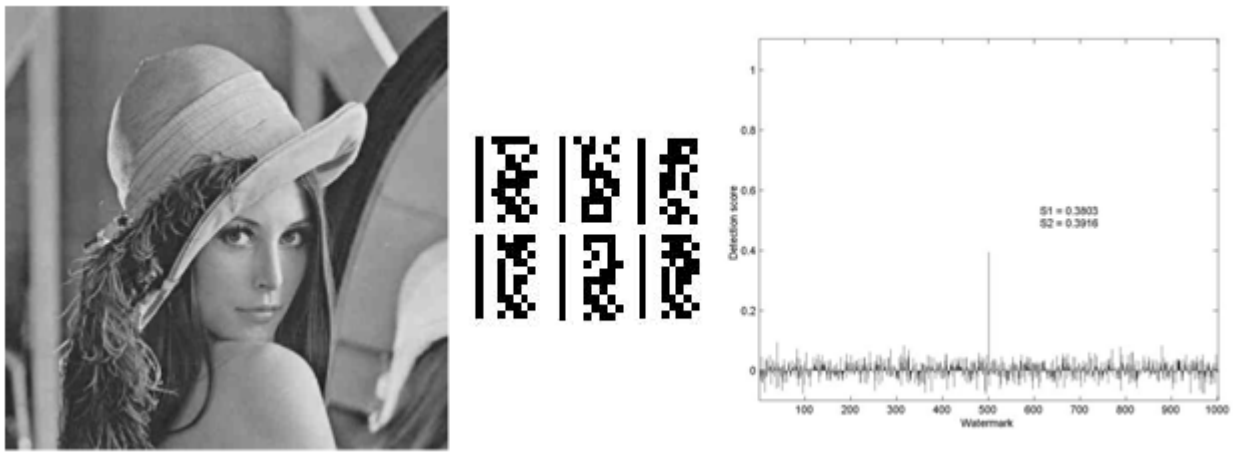

Fig. 13. (Left) MWE-watermarked image under JPEG attack, SF = 2, PSNR = 33:46 dB, bpp = 0:412. (Middle) Decoded watermark. (Right) Random watermark detection results.

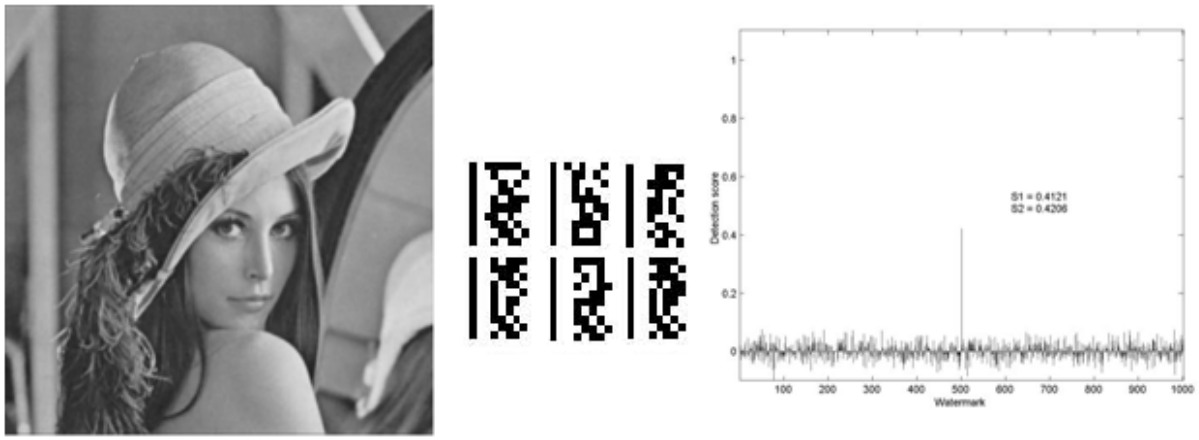

Fig.14. (Left) MWE-watermarked image under LPF attack. PSNR = 31:82 dB. (Middle) Decoded watermark. (Right) Random watermark detection results.

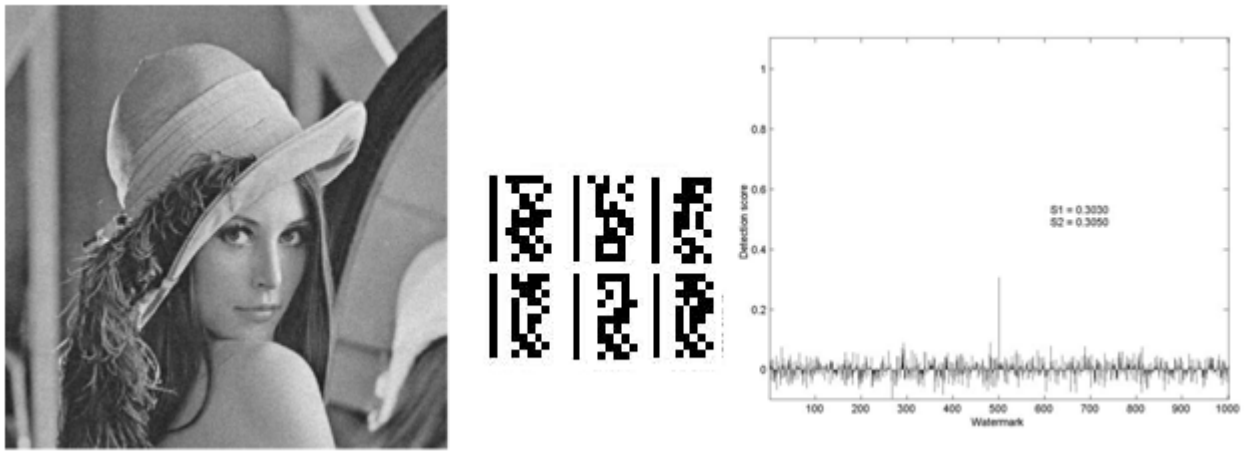

Fig. 15. (Left) MWE-watermarked image under noise attack. PSNR = 29:94 dB, variance = 64. (Middle) Decoded watermark. (Right) Random watermark detection results.

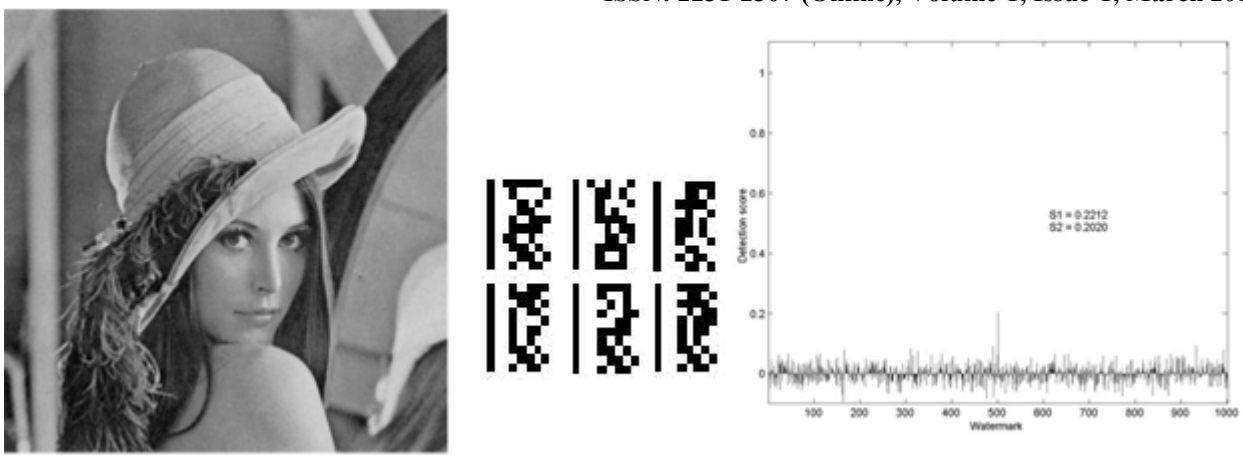

Fig. 16. (Left) MWE-watermarked image under print-and-scan attack. PSNR = 27:41 dB. (Middle) Decoded watermark. (Right) Random watermark detection results.

V. CONCLUSION

The proposed method includes novel invisible blind watermarking scheme to embed watermarks into digital images. The watermarks are designed to be decoded or detected without the original images. The results obtained through this method are similar to that of MWE with added robustness due to the selection process in DCT. Thus the method is robust in varying degrees to unintentional attacks such as JPEG compression, transcoding, LPF, additive noise, and print-and-scan.

REFERENCES

[1] C. Cachin, "*An Information-Theoretic Model for Steganography*", Proceedings of second Workshop on Information Hiding, MIT Laboratory for Computer Science, May 1998

[2] R. Popa, "*An Analysis of Steganographic Techniques*", The "Politehnica" University of Timisoara, Faculty of Automatics and Computers, Department of Computer Science and Software Engineering, http://ad.informatik.unifreiburg.de/mitarbeiter/will/dlib_bookmarks/digital-watermarking/popa/popa.pdf, 1998

[3] A. Al-Gindy, H. Al-Ahmad, R. Qahwaj, and A. Tawfik, "*A novel blind Image watermarking technique for colour RGB images in the DCT domain using green channel* " in Mosharaka International Conference on Communications, Computers and Applications (MICCCA 2008). , Amman, Jordan, 2008.

[4] P.H.W. Wong, Oscar C. Au and Gene Y.M. Yeung, "*A Novel* Blind *Multiple Watermarking Technique for Images*," IEEE Trans on Circuits and Syst. For Video Technol., vol. 13, no. 8, pp. 813-830, Aug. 2003.

[5] Pseudo random bit sequence: http://en.wikipedia.org/wiki/Pseudorandom_binary_sequence

[6] Zero mean Gaussian with variance: http://en.wikipedia.org/wiki/Normal_distribution