Privacy protection of data with safe watermark extraction using data hiding and sparse sampling

H.N.RAnotkar[#1], M.S.Deshmukh[*2]

[1]*Department of Information Technology, PRMITR, Badnera,Maharshtra,India*
[2] *Assistant professor,Department of Information Technology, PRMITR, Badnera, Maharshtra,India*

Abstract: *Data privacy protection is a very important aspect in this cloud computing era. Also,it becomes very easy for the user to collect the data from the various sources without worrying about the copyright information as a result of the rapid growth of the Internet and social networks. Hence adding certain ownership information into the data(image) becoming necessary for the data owners now-a-days. To do so watermarking technique is used. This restricts the reuse and republishing of the data to authenticated users only. In this paper, we propose a system that enhances privacy protection of image data as well as the safe watermark extraction in a simultaneous manner by using sparse sampling, data hiding and secure computation. We provide enhancement in the privacy of the image data by applying layer of encryption over the data using data hiding technique in which we add identity bits to the image data before it is passed to the sparse sampling transformation. This will make the proposed architecture tolerable against the semi-honest security assumption required for the simultaneous operation in sparse sampling domain and without these bits an attacker could not get the original data. In SS transformation we deal with DWT coefficients of image.*

Keywords : *Watermark Embedding, Watermark extraction, data hiding sparse sampling, secure computation, DWT*

I. INTRODUCTION

The security of information is one of the most important factors of information technology and communication because of the rapid growth of Internet as well the concept of cloud computing. So for providing the security to the privacy of communication cryptography was introduced.The different methods are proposed to perform encryption and decryption over the data to keep the information secret. But it is found that keeping the contents of a message secret, is not enough sometimes .As a result it may also be necessary to keep the existence of the message secret. The technique used for doing so, is called steganography. Steganography is the art and science of invisible communication. This is done by hiding information in other information. Thus, it hides the existence of the communicated information. The word steganography is derived from the Greek words "*stegos*" meaning "cover" and "*grafia*" meaning "writing" defining it as "covered writing". In image steganography the information is hidden exclusively in images. In simple the steganographic process is defined as:

cover_medium + hidden_data + stego_key = stego_medium

where, the *cover_medium* is the file in which we are going to hide the *hidden_data*, which may also be encrypted using the *stego_key*. The result of the operation will be the *stego_medium* .It must be of same type as that of cover medium.

The other technology that is closely related to steganography is watermarking. It is strongly focused on the copyright protection of intellectual property. Hence obviously the requirement of

algorithm is different from those of steganography. In watermarking all of the instances of an object are "marked" in the same way. In watermarking generally a signature to verify the ownership of the data is embedded within the image data that helps to protect copyright issues. It becomes important as user can easily get the multimedia data(image) from the internet very easily from different sources without knowing the copyrights. Watermarking is performed in two ways:

1. Invisible watermarking
2. Visible watermarking

The first type provides embedding of copyright information imperceptibly into host media. To identify the ownership information, the hidden information can be retrieved from the protected host. The watermarked image must be resistant enough to common image operations to ensure that the hidden information is still retrievable after such alterations. The second method, on the other hand, yield visible watermarks which are generally clearly visible after common image operations are performed. In addition, visible watermarks convey ownership information directly on the media and can deter attempts of copyright violations.

The different safe watermark extraction techniques were proposed to convince a verifier whether or not a watermark is embedded without exposing the watermark pattern so that an untrusted verifier cannot remove the watermark from the watermark protected copy [3], [11]. There are two approaches for safe watermark extraction: Asymmetric watermarking [18] and Zero-knowledge watermark detection [4],[7]. However, most of the existing safe watermark extraction works target on the security of the watermark pattern and very less attention is paid over the target data on which the extraction is performed. But as stated above, in some application, it is required to protect the multimedia data's(image) privacy in the watermark extraction process. Implementing such kind of storage and safe watermark extraction simultaneously is possible by using the existing safe watermark extraction technologies such as zero-knowledge proof protocols [4],[7] that transform the multimedia data to a public key encryption domain. But, it also has certain limitations, as complicated algorithms, high computational and communication complexity [11], and large storage consumption in the public key encryption domain, may impede their practical applications.

So to overcome these issues a Sparse sampling technique with secure multiparty computation protocol is used to form a framework that simultaneously perform privacy protection and safe watermark extraction in the storage place[19].Sparse sampling is known to reconstruct a signal or image from few samples of the image. This system is secured under semi-honest assumption. It makes use of DCT coefficients of the image over which the CS transformation is performed. For privacy preserving storage, as the DCT coefficients are not perfectly sparse, the CS reconstruction will introduce distortion to the reconstructed image and as a result the quality of the reconstructed image is an issue.

In our work, we implement lossless visible watermarking over the image data. After that we perform data hiding over the watermarked data to in which we add identity bits to the watermarked data. It will provide enhancement to the privacy of data. As a result, a new enhanced encrypted watermarked data is formed. In SS transformation, a DWT coefficients [1],[17] of this image is obtained. After this a SS matrix is generated for this image from which we will reconstruct the signal. This SS transformation makes the attacks over the data probable only. After SS transformation we can decrypt the data by providing identity bits. Once he knows to be authorised user he can extract the

watermark from the image data. We then observed performance evolution parameters of the proposed system using MATLAB as MSE, PSNR and correlation factor.

II. Proposed work:

There are three steps involved in the process of proposed system.

1. Watermark Embedding and Data hiding
2. Sparse Sampling Transformation
3. Watermark Extraction and Data Recovery

In Watermark embedding and data hiding process, firstly, a multimedia data (image) and a watermark pattern is accepted. Then, a watermark pattern is embedded within an image that ensures the proof of ownership of the user over the data[20].It discourages the misuse or unauthorised use of the data over the internet and gives the guarantee that signal is unmodified. After embedding a watermark a data hiding is applied over the watermarked data to form a new encrypted data in which identity bits called authenticated bits are substituted at the second bit to Least significant bit of the original data so that only the authenticated user can get the original data. It will enhance the privacy of data and the change in data is not even visible to human eye. If the user does not possess right authenticated bits he could not reveal the original data.

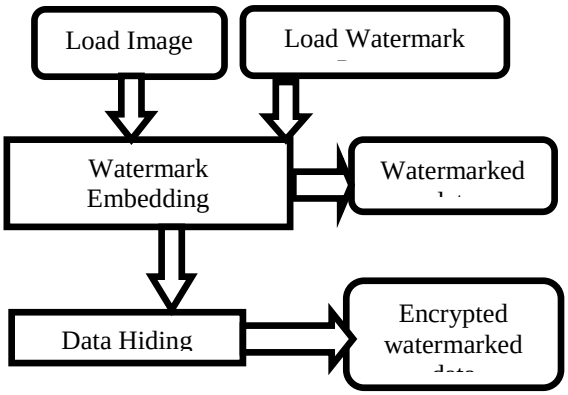

Fig.2.1 Watermark embedding and data hiding Process

In Sparse Sampling Transformation process [19], an encrypted watermarked image formed is passed as input. After getting an image, it will calculate DWT coefficients of an encrypted image. Once the DWT coefficient calculation is done a new Sparse Sampling matrix is generated for it. From this matrix the original data is reconstructed by the user. It makes use of the sparsity in an image where from few samples it reconstructs the signal. If the SS matrix is wrong, a user could not reconstruct the original data. So there is double protection to the data. If the attacker guesses random SS matrix of the image he also need to know authentication bits

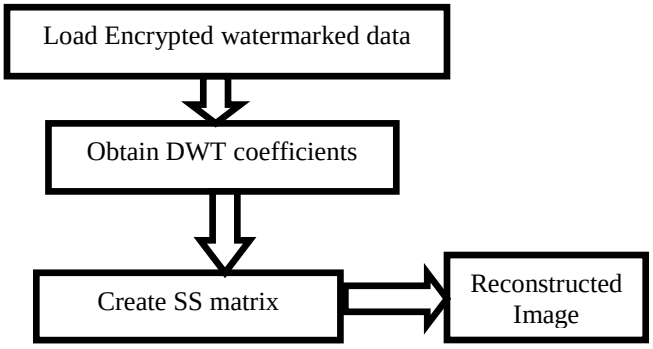

Fig. 2.2 Sparse sampling Tramsformation

In Watermark Extraction and Data Recovery process, firstly to recover the data it checks whether the identity bits provided by user are equal to the embedded one or not. If they are identical then the decrypted watermarked data is passed to the user over which he then applied watermark extraction.

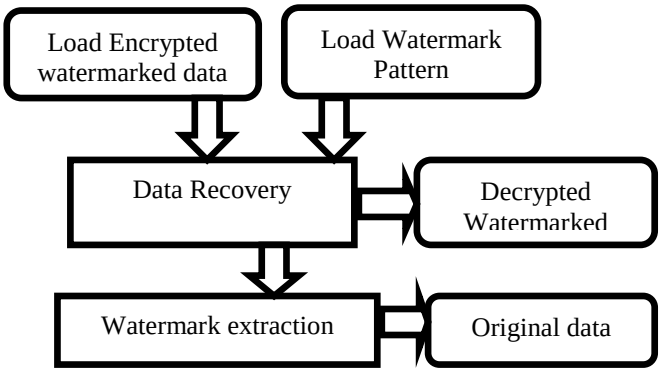

Fig. 2.3 Watermark extraction and data recovery process

III. Result and discussion:

Here we perform operations over greyscale images. We implemented the visible watermarking over the image data and applied data hiding policy by inserting eight identity bits to the original data at the second bit to the LSB. Thus forming the encrypted data.

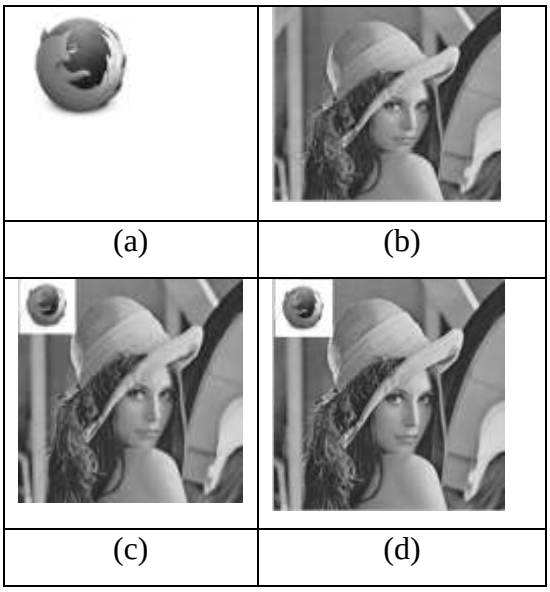

Fig. 3.1 a) watermark pattern b) image data c) watermarked data d) encrypted watermarked data(after data hiding)

Over this encrypted data we perform SS transformation and calculates its DWT coefficients. From this we reconstruct the signal and calculate its performance parameters as MSE,PSNR and correlation factor.

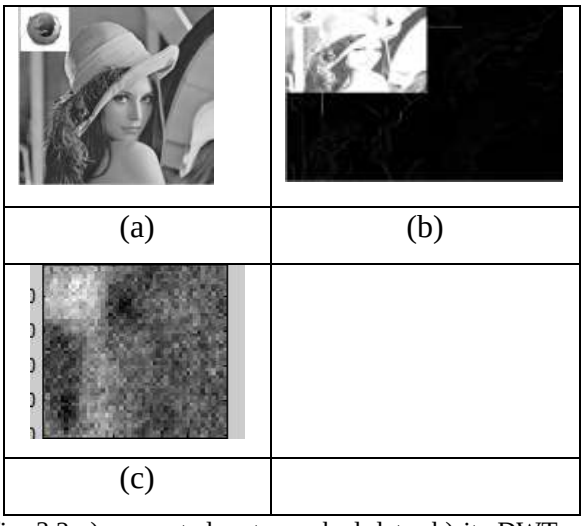

fig. 3.2 a) encrypted watermarked data b) its DWT coefficients c) part of the reconstructed image

Though the SS reconstruction is done, to get the original data and to perform watermark extraction we performed decryption of the hidden data(identity bits) and recovered original image after safe watermark extraction.

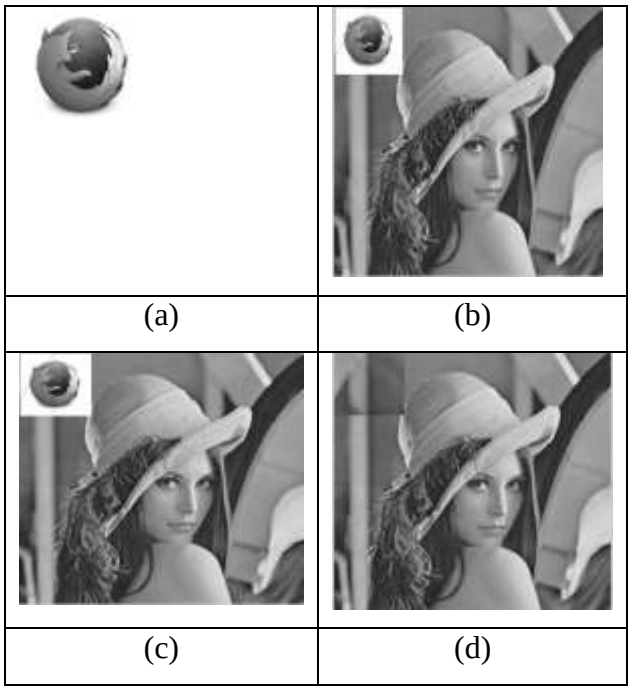

Fig.3.3a)watermarkpattern b)encrypted watermarked data c) decrypted watermarked data d) Recoverd image

Table 3.1 MSE of recovered image

Image name	MSE
Tulips	2.12E-02
Lenna	1.37E-04
Baboon	0
Penguins	3.51E-04
Peppers	1.22E-04
Desert	1.22E-04
Hydragens	2.14E-04
Koala	1.53E-04
Jellyfish	3.01E-02
Lighthouse	3.40E-03

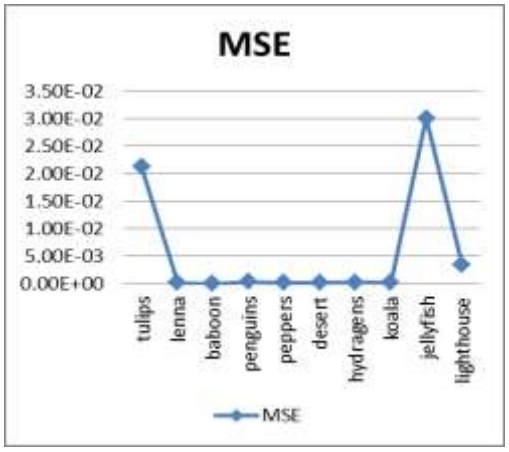

Fig. 3.4 Line graph of MSE of different images after watermark extraction and data recovery.

Table 3.2 PSNR of recovered image

Image name	PSNR
Tulips	6.49E+01
Lenna	8.68E+01
Penguins	8.27E+01
Peppers	8.73E+01
Desert	8.73E+01
Hydragens	8.48E+01
Koala	8.63E+01
Jellyfish	6.33E+01
Lighthouse	7.28E+01
Baboon	inf

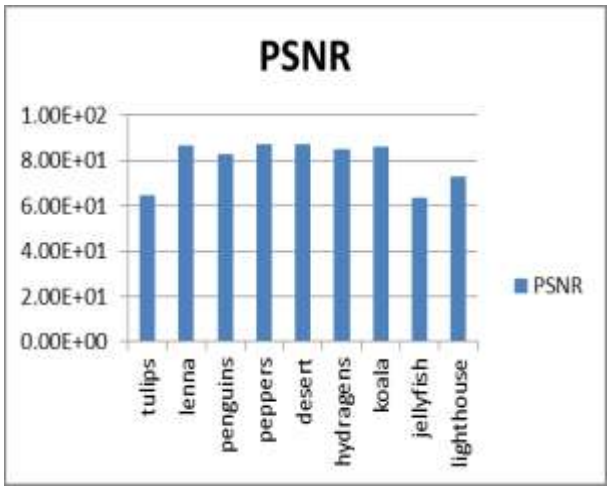

Fig. 3.5 Bar graph of PSNR of different images after watermark extraction and data recovery.

Table 3.3 corelation factor of recovered image

Image name	correlation
tulips	2.12E-02
lenna	1.37E-04
baboon	0
penguins	3.51E-04
peppers	1.22E-04
desert	1.22E-04
hydragens	2.14E-04
koala	1.53E-04
jellyfish	3.01E-02
lighthouse	3.40E-03

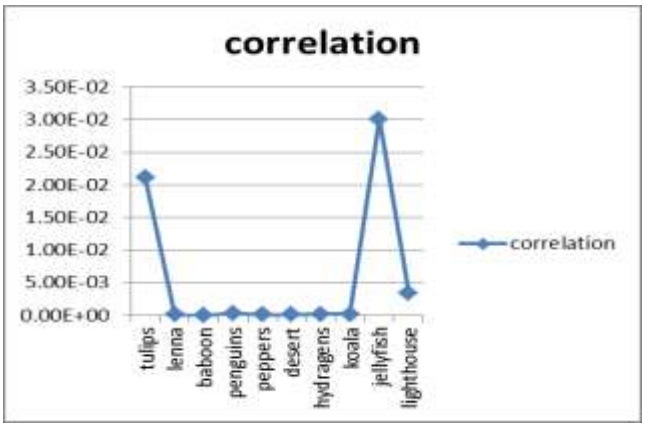

Fig. 3.6 Line graph of corelation of different images after watermark extraction and data recovery.

Table 3.4 MSE of recovered image using SS transformation

Image name	MSE
Tulips	5.80E+02
Lenna	1.44E+03
Baboon	1.50E+03
Penguins	1.93E+03
Peppers	1.34E+03
Desert	8.88E+01
Hydragens	1.99E+03
Koala	2.25E+03
Jellyfish	2.53E+03
Lighthouse	2.46E+03

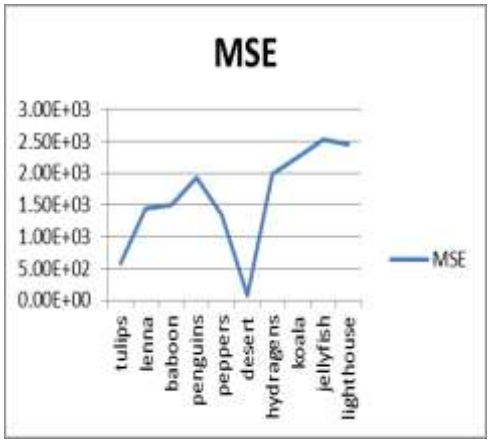

Fig. 3.7 Line graph of MSE of different images after recovery using SS matrix.

Table 3.5 PSNR of recovered image using SS ransformation

Image name	PSNR
Tulips	2.05E+01
Lenna	1.65E+01
Penguins	1.63E+01
Peppers	1.53E+01
Desert	1.69E+01
Hydragens	2.86E+01
Koala	1.52E+01
Jellyfish	1.46E+01
Lighthouse	1.41E+01
Baboon	14.2222

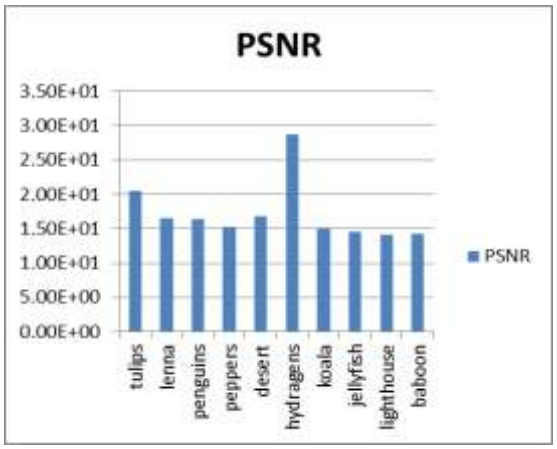

Fig. 3.8 Line graph of PSNR of different images after recovery using SS matrix.

Table 3.4 corelation of recovered image using SS transformation

Image name	correlation
tulips	8.00E-01
lenna	5.67E-01
baboon	0.662
penguins	5.47E-01
peppers	6.14E-01
desert	7.94E-01
hydragens	7.29E-01
koala	6.42E-01
jellyfish	6.65E-01
lighthouse	8.08E-01

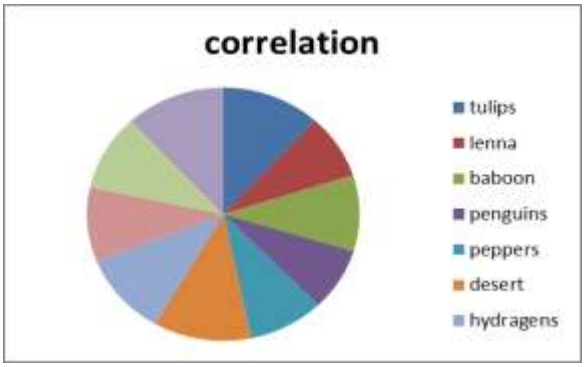

Fig. 3.7 pi-chart of correlation of different images after recovery using SS matrix

Conclusion:

The proposed system provides the protection to the image data in two steps with safe watermark extraction in simultaneous manner. SS matrix makes the attacks to the data probable and identity bits added to the image by data hiding provide additional layer of protection, although, a SS matrix calculated rightly by attacker. It also makes the system tolerable against the semi-honest assumption required in case of simultaneous access. The watermark embedding and extraction is carried out in lossless manner and decryption of the hidden data also results in a good quality recovered images as the original one.

Future Scope:

In future the proposed system may be extended to other data types as audio/video of different types. Also, developing a secure computation protocol for safe sparse sampling reconstruction is a future work

References

[1] Rafael C. Gonzalez, Richard E. Woods.(1992), Digital Image Processing(2nd edition), NJ:Prentice Hall

[2] P. Paillier, "Public-key cryptosystems based on composite degree residuosity classes," in Proc. Adv. Cryptology-Eurocrypt, 1999, pp. 223–238.

[3] J. Eggers, J. Su, and B. Girod, "Public key watermarking by eigenvectors of linear transforms," in Proc. Euro. Signal Process. Conf., 2000.

[4] A. Adelsbach and A. Sadeghi, "Zero-knowledge watermark detection and proof of ownership," in Proc. 4th Int. Workshop Inf. Hiding,vol. 2137. 2001, pp. 273–288.

[5] B. Goethals, S. Laur, H. Lipmaa, and T. Mielikainen, "On private scalar product computation for privacy-preserving data-mining," in Proc. 7th Int. Conf. Inf. Security Cryptology, 2004, pp. 104–120.

[6] O. Goldreich, The Foundations of Cryptography. Cambridge, U.K.:Cambridge Univ. Press, 2004.

[7] M. Malkin and T. Kalker, "A cryptographic method for secure watermark detection," in Proc. 8th Int. Workshop Inf. Hiding, 2006, pp. 26–41.

[8] K. Liu, H. Kargupta, and J. Ryan, "Random projection-based multiplicative data perturbation for privacy preserving distributed data mining,"IEEE Trans. Knowl. Data Eng., vol. 18, no. 1, pp. 92–106, Jan. 2006.

[9] D. Donoho, "Compressed sensing," IEEE Trans. Inf. Theory, vol. 52,no. 4, pp. 1289–1306, Apr.2006.

[10] M. Rudelson and R. Vershynin, "Sparse reconstructions by convex relaxation: Fourier and Gaussian measurements," in Proc. Conf. Inf. Sci. Syst., Mar. 2006, pp. 207–212.

[11] Z. Erkin, A. Piva, S. Katzenbeisser, R. Lagendijk, J. Shokrollhi,G. Neven, et al., "Protection and retrieval of encrypted multimedia content: When cryptography meets signal processing," EURASIP J. Inf. Security, vol. 7, no. 2, pp. 1–20, 2007

[12] J. Tropp and A. Gilbert, "Signal recovery from random measurements via orthogonal matching pursuit," IEEE Trans. Inf. Theory, vol. 5, no. 12, pp.4655-4666. Dec. 2007

[13] A. Orsdemir, H. O. Altun, G. Sharma, and M. F. Bocko, "On the security and robustness of encryption via compressed sensing," in Proc. IEEE Military Commun. Conf., Nov. 2008, pp. 1040–1046

[14] D. Hsu, S. M. Kakade, J. Langford, and T. Zhang, "Multi-label prediction via compressed sensing," in Proc. NIPS, 2009, pp. 772–780

[15] W. Lu, A. L. Varna, and M.Wu, "Security analysis for privacy preserving search for multimedia," in Proc. IEEE 17th Int. Conf. Image Process.,Sep. 2010, pp. 2093–2096.

[16] M. Davenport, P. Boufounos, M. Wakin, and R. Baraniuk, "Signal processing with compressive measurements," IEEE J. Sel. Topics SignalProcess., vol. 4, no. 2, pp. 445–460, Apr. 2010.

[17] Anilkumar Katharotiya ,Swati Patel, Mahesh Goyani, "Comparative Analysis between DCT & DWT Techniques of Image compression" in Journal of Information Engineering and Applications Vol 1, No.2, 2011

[18] T. Bianchi and A. Piva, "Secure watermarking for multimedia content protection: A review of its benefits and open issues," IEEE Signal Process. Mag., vol. 30, no. 2, pp. 87–96, Mar. 2013.

[19] Qia Wang, Wenjun Zeng, Fellow, IEEE, and Jun Tian,Member "A compressive sensing based secure watermark detection and privacy preserving storage framework" in, IEEE issues in IEEE Transactions on image processing, vol. 23, no. 3,march 2014

[20] Tsung-Yuan Liu and Wen-Hsiang Tsai "Generic Lossless Visible Watermarking-A New Approach" in IEEE transactions on image processing, vol. 19, no. 5, may 2010