Steganalysis and Image Quality Measures

Neha Singh
Assoc. Prof., Department of Electronics and Communication Engineering
Institute of Engineering and Technology
Alwar, India

***Abstract*—Steganography is the art/ science of covert communication and steganalysis is the counter to it. Though the first goal of steganalysis is detection of hidden message, there can be additional goals such as disabling, extraction and /or manipulating the original hidden message. Detection of the secret message is enough to defeat the very purpose of *steganography* even if it is not extracted because detecting the existence of hidden data is enough if it needs to be destroyed. The performance of steganalysis approach depends on identifying the appropriate Image Quality Measures. A good objective quality measure should well reflect the distortion on the image due to embedding, blurring, noise and compression. A survey of Image Quality Measures used by the various reported Steganalysis techniques is presented in this paper.**

***Keywords*- Image Quality Measures, Steganography, Steganalysis, HVS based IQM, Pixel based IQM.**

I. INTRODUCTION

Hiding information in digital content has a wide class of application and the techniques involved in such applications are collectively referred to as *information hiding techniques* [1]. Steganography refers to the art/science of embedding information in the digital images, aiming to convey messages secretly by concealing its very existence. Compared with the Cryptography, modern steganography not only encrypts messages but also masks the very presence of the communication. A wide range of steganography techniques have been introduced which exploit spatial or transform domain characteristics of images to hide information. A survey of these techniques can be found in [2-7].

The message embedding into an image results in distortions which can be visible under human observation with an experienced observer but are generally imperceptible to human eyes. But there exists some detectable artifacts in the images depending on the steganographic algorithm, which the steganalyst uses for the detection of data hiding. Steganalysis is the practice of attacking the steganographic technique and aims at detecting the mere presence or estimating potentially hidden information from observed data with little or no knowledge about the steganography algorithm and/or its parameters. The success of an attack is application dependent [7]; for a secret communication application the mere detection of hidden data is a success but for a steganalyst attempting to defeat a copyright mark, a successful attack requires that he not only detects the mark but also destroys or modifies the mark without significant degradation of the perceptual quality of the stego-image. Detection of presence of hidden information can be performed by close examination of the stego-image for distortions or exploration of the strong inter-pixel dependencies that are characteristic of natural images. But the underlying steganographic algorithm(s) affect the steganalysis method.

A study of various steganalysis techniques is done with respect to different Image Quality Measures used, in the rest of the paper. The terms 'method', 'algorithm' and 'technique' are used interchangeably throughout the paper. The term 'cover' refers to the original image devoid of any hidden information and 'stego-image' refers to the image with hidden information.

II. CLASSIFICATION OF STEGANALYSIS TECHNIQUES

Primarily, steganalysis techniques are labeled as: active or passive. If mere detection of presence or absence of any hidden information is the purpose, the technique is referred to be passive else if the hidden information is to be retrieved or some properties of the hidden message are to be extracted, the technique is referred to be active [8-9]. Nissar and Mir [10] classify steganalysis into two classes: signature steganalysis and statistical steganalysis. The classification is based on whether the signature of the steganography technique or the statistics of image is used to detect the presence of hidden messages in images. Under each class, the methods are further sub-divided into specific approaches and universal or generic approaches. Specific steganalysis techniques are those which depend upon the underlying steganographic approach and are expected to have high success rate. The generic steganalysis techniques are not technique specific but are universal, that is, these are expected to produce effective results with all steganographic techniques of a class.

On the basis of knowledge of cover, stego-image and message available for steganalysis, the techniques are also classified [10] as: *Stego-only* : where only the stego-image is available, *Known- cover* : where both the original cover and a corresponding stego-image is available, *Known-message* : when the steganalyst knows the secret message embedded in a stego-image, *Chosen-stego* : when access to the message extraction tool is available so that the attacker does not have to deduce the decoding algorithm, *Chosen-message* : where the steganalyst has access to the steganography encoding tool itself and can embed and analyze messages of his own choosing.

The research in the field of steganalysis has focused on extracting features from images. Based on the approach [11] used for steganalysis, the techniques are classified as: *classifier based* or *statistical parameter based* or *hybrid*. The classifier based techniques essentially construct a classifier which extracts image features from the training set in the first phase to update iteratively the classification rule. In the second phase, the classifier classifies the test inputs as cover or stego-image based on the best classification rule obtained in first phase. The statistical parameter based steganalysis approach tests parametric statistical models based on the cover, message and stego-images to detect if the test image has some hidden data or not. The hybrid techniques are those which use combinations of other techniques.

III. Image Quality Measures

Image quality measures are figures of merit used for the evaluation of coding/processing techniques. Image quality measurement continues to be the subject of intensive research and experimentation. The selected features should be sensitive to the hidden message while insensitive to other. These measures are generally pixel difference-based, correlation-based, histogram-based, spectral-based, context-based and HVS-based (Human Visual System-based). In spite of their complicated algorithms, the HVS-based objective measures do not appear to be superior to the simple pixel-based measures like the Mean Square Error (MSE), Peak Signal-to-Noise Ratio (PSNR) or Root Mean Squared Error (RMSE). The statistical evidence left by steganography is captured by a combination of selective Image Quality Measures which is then exploited for steganalysis.

Ismail et al. used IQMs as a steganalysis tool rather than as an indicator of image quality or algorithmic performance. They used MSE, multiresolution distance measure, structural content, cross correlation, weighted spectral distance, median block weighted spectral distance, HVS Based Normalized absolute error, HVS Based L2, and gradient measure as IQM for their proposed universal steganalysis method. In [12] it is indicated that the image quality metrics form a multidimensional feature space whose points cluster well enough for the classifier to successfully classify watermarked and non-watermarked images even when the tested images come from an embedding technique unknown to it. Rather than an arbitrary set of features, [13] uses the ratio of any two Fourier coefficients of the DCT coefficients, by modeling their distribution as a Laplacian and was reported effective against pixel domain hiding.

The image histogram is greatly explored for IQM as it is essentially the probability mass function (pmf) of the image and probability density function (pdf) can be thought as the normalized version of a histogram. The histogram characteristic function center of mass (HCF-COM), which gives general information about the energy distribution in HCF, is exploited to capture the low pass filter effect of the additive noise. The HCF-COM can successful detect the steganographic techniques of additive noise type. It is effective to observe the statistical changes in wavelet domain. In [14] Yun Q. Shi et al. use the statistical moments of the Characteristic Functions (Fourier transform of the pdf, with a reversal in the sign of exponential) of both a test image and its wavelet subbands as features for steganalysis. The moments of wavelet CF's can reflect the differentiation property of the associated histograms, hence, reflecting sensitively the changes caused by data hiding. Another variant for the use of statistical moments of CF obtained for all subbands after 2 level wavelet decomposition of the image is reported by Yun Q. Shi et al. in [15]. The first and higher order wavelet statistics are employed with a non-linear support vector machines (SVM) to detect steganographic messages and exploit color statistics using the classifier in [16].

Binary similarity measures are exploited greatly to determine similarity between two images. The early measures are based upon bit-by-bit matching of the corresponding pixels. Many other binary similarity measures are listed in [17]. The steganalysis technique proposed in [17] is based on the regression of seven binary similarity measures between successive bit planes of an image to determine the presence of a message hidden by LSB technique. This technique is however not for active warden case.

Combining the features of histogram and binary similarity measures, [17] proposes an improvement over Difference Image Histogram [17] method of steganalysis for LSB embedding, to reduce the mean square error by 50% for embedding ratios greater than 40%. Difference Image Histogram method uses the measure of weak correlation between successive bit planes to construct a classifier for discrimination between stego-images and cover images. The proposed algorithm in [18] reduces the initial-bias, and estimates the LSB embedding message ratios by constructing equations with the statistics of difference image histogram.

The steganalysis techniques proposed in [19] employs only a 5-dimensional feature vector of four Huffman length statistics (H) and the ratio of file size to resolution (FR Index) which are unique, accurate and monotonic over a wide range of settings for YASS and several supervised classifiers with the accuracy of prediction superior to most blind steganalyzers.

Artificial Neural Network (ANN) for pattern classification of feature data extracted from the cover image and stego image data has been used by [20] with three selected IQMs. They are the median block spectral phase, the median weighted block spectral distortion and the normalized mean square HVS error.

In [21] Bin Li et al. extract steganalytic features by local linear transform, a signal processing based approach for texture classification to work on gray scale as well as color images for universal steganalysis.

The choice of Quality measure is highly dependent on the steganography technique to be defeated. Most of the above techniques could detect the LSB steganography but an attempt to defeat message hiding by halftoning and coordinate projection was made by Kim and Park in [22] by extracting the features from the frequency domain of the histogram. They selected three features: the ratio of the sum of higher frequencies and the sum of lower frequencies, the position of the second peak in the frequency domain of the histogram and the difference of between the first moment of test image and modified test image using data hiding via halftoning and coordinate projection algorithm in the frequency domain.

IV. CONCLUSION

It is very difficult to construct a universal steganalysis tool. The field of steganalysis is still growing. The steganalysis methods proposed in literature are inexahustive and out of the scope of this paper to include all. An attempt has been made to bring to light the IQM measures used generally for some of the steganographic techniques. It is hardly possible to achieve high correct classification rate with a single feature. So, multi-dimensional feature vector should be used. Some of the commercially available steganalysis tools are listed in [23]. These are expected to defeat most of the techniques for LSB embedding. While it is possible to design a reasonably good steganalysis technique for a known steganography algorithm, the long term goal must be to develop a steganalysis framework that can work effectively at least for a class of steganography methods, if not for all.

V. REFERENCES

[1] Peticolas Fabien A., Anderson Ross J., Kuhn Markus G., “Information Hiding- A Survey”, *Proceedings of IEEE, Special issue on protection of multimedia content*, 87(7): 1062-1078, July 1999.

[2] Rey Christian, Dugelay Jean-Luc, “A Survey of Watermarking Algorithms for Image Authentication”. *EURASIP Journal on Applied Signal Processing*, 2002:6, 613-621.

[3] Liu Lin, ”A Survey of Digital Watermarking Techniques” ,*unknown*.

[4] Agarwal Nitika, Singh Neha, “Transform Domain Digital Image Watermarking: A Survey”, *Proc. of International Conference on Wavelet Transforms and Its Applications*, March 2011.

[5] Natanj Sara, Taghizadeh Seyed Reza, “Current Steganography Approaches: A survey”, *International Journal of Advanced Research in Compouter Science and Software Engineering*, Vol 1, Issue 1, Dec. 2011.

[6] Agarwal Nitika, Singh Neha, “Techniques Used for Digital Image Watermarking:A Survey”, *Proc. of International Conf. on Global Trends in Technology: Impact on Industry and Society*, Oct. 2011.

[7] Lin Eugene T., Delp Edward J.,“A Review of Data Hiding in Digital Images”, Prudune University.

[8] Chandramouli R., Memon N.D., “Steganographic Capacity:A Steganalysis Perspective” *Proc. SPIE Security and Watermarking of Multimedia Contents*, vol. 5020, pp. 173–177, Jan. 2003.

[9] Chandramouli R., “On Information Hiding With Incomplete Information About Steganalysis” ,*IEEE Proc. of International Conference on Image Processing,* Vol.2,pp.1161-1164, Oct. 2004.

[10] Arooj Nissar, A.H. Mir, “Classification of steganalysis techniques: A study”, *Academic Presss, Journal of Digital Signal Processing*, Vol 20, Issue 6, Dec. 2010

[11] Chandramouli R., Subbalakshmi K.P., “Current Trends In Steganalysis: A Critical Survey” *Proc. of Control, Automation, Robotics and Vision Conference*, 2004.

[12] Ismail Avcibay, Nasir Memon, Bulent Sankur, “Steganalysis Based On Image Quality Metrics” *IEEE Trans.on Image Processing,* Vol 12, No. 2, Feb, 2003.

[13] Tu-Thach Quach, Fernando Pérez-González and Gregory L. Heileman, "Model-based steganalysis using invariant features", *Proc. SPIE 7254*, 72540B (2009); doi:10.1117/12.810507.

[14] Yun Q. Shi, Guorong Xuan, Dekun Zou, Jianjiong Gao, Chengyun Yang, Zhenping Zhang, Peiqi Chai, Wen Chen, Chunhua Chen, "Steganalysis Based on Moments of Characteristic Functions Using Wavelet Decomposition, Prediction-Error Image, and Neural Network" *IEEE ICME, pp 269-272, 2005.*

[15] Yun Q. Shi, Guorong Xuan, Chengyun Yang, Jianjiong Gao, Zhenping Zhang, Peiqi Chai, Dekun Zou, Chunhua Chen, Wen Chen, " Effective Steganalysis Based on Statistical Moments of Wavelet Characteristic Function", *IEEE Proceedings of the International Conference on Information Technology: Coding and Computing*, Vol.1, 2005.

[16] Siwei Lyu, Hany Farid, " Steganalysis Using Color Wavelet Statistics and One-Class Support Vector Machines", *Proc. Security, Steganography, and Watermarking of Multimedia Contents*, 2004, pp.35-45.

[17] Bhanu Prakash Battula , R. Satya Prasad, "Essentials Of Image Steganalysis Measures", *Journal of Theoretical and Applied Information Technology,* Vol 11, No.1, January 2009.

[18] Sanjay Kumar Jena, G.V.V. Krishna, "Blind Steganalysis: Estimation of Hidden Message Length", *International Journal of Computers, Communications & Control,* Vol. II (2007), No. 2, pp. 149-158.

[19] Bhat Veena H., Krishna S., Shenoy P. Deepa, Venugopal K. R., Patnaik L. M., "Steganalysis of YASS Using Huffman Length Statistics", *International Journal of Hybrid Information Technology*, Vol. 4 No. 3, July, 2011.

[20] Jennifer Davidson, Clifford Bergman, Eric Bartlett, "An artificial neural network for wavelet steganalysis", *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 5916, Mathematical Methods in Pattern and Image Analysis, 2005, pp. 1-10.

[21] Bin Li, Jiwu Huang, Yun Q. Shi, "Textural Features Based Universal Steganalysis" *Proc. of SPIE: Forensics, Steganography, and Watermarking of Multimedia Contents X*. Volume 6819, pp. 681912-681912-12, 2008..

[22] Kim Woong Hee, Park Ilhwan, "Steganalysis of Data Hiding via Halftoning and Coordinate Projection", *World Academy of Science, Engineering and Technology*, 2005.

[23] Hayati Pedram, Potdar Vidyasagar, Chang Elizabeth, "A Survey of Steganographic and Steganalytic Tools for the Digital Forensic Investigator" Workshop of Information Hiding and Digital Watermarking, July 2007.

About The Author

Ms. Neha Singh received her B.Tech. (Hons.) with 81.5% in ECE from University of Rajasthan in the year 2004 and M.Tech. in VLSI Design from Malaviya National Institute of Technology, Jaipur with CGPA of 9.22 in the year 2009. She has been teaching since 2004. She has guided many M.Tech Dissertations and has published/presented more than 20 papers in National/International Conference/Journals. She has authored books on Signals and Systems, Digital Signal Processing, Digital Logic Design and Basic Electrical and Electronics Engineering and edited many engineering books too. Her areas of interest include Analog Electronics, VLSI Design and Technology, IC Technology, Signal Processing and Image Processing. She is associated with World Bank/MSME projects of high repute.