ISSN: 2278 – 8875

International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering
Vol. 2, Issue 7, July 2013

Efficient Data Embedding Technique Comparison Based On LSB Replacement and Pixel Pair Matching Method

S.Saravana

Assistant Professor, Department of ETE, Bharath University, Chennai-600073, India

ABSTRACT—This paper deals with the detection of hidden bits in the Least Significant Bit (LSB) plane of a natural image. The mean level and the covariance matrix of the image, considered as a quan tized Gaussian random matrix, are unknown. An adaptive statis tical test is designed such that its probability distribution is always independent of the unknown image parameters, while ensuring a high probability of hidden bits detection. This test is based on the likelihood ratio test except that the unknown parameters are re placed by estimates based on a local linear regression model. It is shown that this test maximizes the probability of detection as the image size becomes arbitrarily large and the quantization step van ishes. This provides an asymptotic upperbound for the detection of hidden bits based on the LSB replacement mechanism. Numer ical results on real natural images show the relevance of the method and the sharpness of the asymptotic expression for the probability of detection.

INDEX TERMS—Adaptive detection, information hiding, natural image, nuisance parameters, statistical hypotheses testing.

I. INTRODUCTION

SECRET bits embedding concerns the reliable transmission of information embedded into host signals such as image, video and audio. It has an increasingly wide array of applica tions, from digital watermarking, document authentication to steganography [1]–[4]. Many tools are already available in the public domain and others are easy to create [5]. Unfortunately, all these applications can also be misused and, naturally, there is an interest in knowing if such hiding can be reliably detected. In addition, the huge amount of image available on the Internet shows that there is a real need to use detection algorithms with analytic statistical properties. It is especially crucial to warrant a small prescribed false alarm probability and to know in advance the probability to detect the hidden communication.

II. INFORMATION HIDING IN NATURAL IMAGES

The secret message is imbedded into a harmless natural image which is called the cover, or host, image. The resulting image is called the stegoimage. Ideally the stegoimage is indistin guishable from the original cover image, giving no indication that other information has been encoded. The stegoimage is then transmitted to the receiver via an unsecured channel. The stegoimage can be decoded by a receiver that knows the hiding scheme and the specific parameters (the secret key), used by the encoder, to retrieve the secret message

.

An adversary can detect the hidden communication by ob serving the unsecured channel. The detection of hidden com munication is a difficult problem because, generally, it is an illposed problem: the host image, the hiding rate (if data is hidden), and the secret key are unknown. Despite the intrinsic difficulty of the data hiding detection problem, its importance has led to a number of attempts at developing useful tools; see [2]–[4], [6], and [7] for a survey of the methods available in the open literature.

The most studied algorithm is undoubtedly the simple yet popular technique of hiding in the Least Significant Bit (LSB) of the cover image [8], either in the pixel or transform domain, or its variants. There can be no doubt that replacement of LSBs in digital images is a poor choice for steganography [9] but it remains popular in free steganography software. Moreover, this is the mechanism which inspires the majority of existing hiding methods. Broadly, the literature contains three main classes of detectors for LSB replacement. The first, termed structural detectors in [10], includes [10]–[15] among others; they ana lyze explicitly the combinatorial structure of LSB replacement in pixel groups. The second, known as Weighted Stegoimage (WS) detectors, is found in [16] and [17], and involves filtering the stegoimage to estimate the cover. Last, the third class contains statistical detectors [18], [19] which are derived by applying statistical techniques to the inspected image.

III. THEORETICAL LIMITS OF STANDARD APPROACHES

The vast majority of methods proposed in the literature have generally unknown theoretical statistical performances, which are evaluated by using numerical experiments from large image databases. This involves three main drawbacks.

First, the performances of detection algorithms are generally evaluated with Receiver Operating Characteristic (ROC) curves. It is crucial to bear in mind that ROC curves are especially rele vant for testing two simple hypotheses [20]. In presence of com posite hypotheses (due to unknown cover images and unknown hiding rate), a ROC curve is not sufficient to sum up the perfor mances of a detector. Strictly speaking, a ROC curve has to be calculated for each possible cover image content, except when it is theoretically established that the studied test is independent from the image content. This is not the case for the approaches existing in the literature.

Second, when analyzing a large number of images, it is ad visable to warrant a prescribed probability of false alarm (declare an alarm when inspecting a cover image). Designing a detector whose probability of false alarm is theoretically con trolled for all possible cover images is not addressed by standard approaches. This involves designing a test based on a decision function independent of the cover image parameters. This is the wellknown concept of Constant False Alarm Rate (CFAR) de tection [21]. For a CFAR test, the detection threshold can be set to warrant a prespecified false alarm probability. Such a detector is also referred to as an adaptive detector [22], [23].

Finally, no optimal bound on detection performances has been yet established in the literature, except the square root law [24] which shows that the datahiding capacity of covers grows only with the square root of the available cover size. Hence there are no theoretical means of evaluating a new challenger test. The most common approach consists in comparing the challenger test to the best tests known in the literature. This is time consuming and not satisfactory from a theoretical point of view. Moreover, although it is intuitively clear that such a bound must depend on the hiding rate and the cover image size (among other possible parameters), this dependence is not clearly established, even in an idealized setting.

Main Contributions of the Paper

The goal of this paper is to design a statistical test which ad dresses the above mentioned drawbacks of standard approaches. For this purpose, this paper considers an original approach to de tecting hidden bits in the LSB plane which consists in using a parametric model of natural images together with the theory of statistical hypotheses testing [25], [26]. The main contributions of this paper are the following:

- The proposed approach is based on a parametric model of natural images. Hence, it exploits the physical dependence which naturally exists between the image pixels. These pixels are not supposed to be identically distributed but they admit a joint Gaussian distribution.
- An adaptive Asymptotically Uniformly Most Powerful (AUMP) test is designed (under mild assumptions) to decide if a natural image contains hidden bits. This test maximizes the detection probability of hidden bits, inde pendently of the image parameters, whatever the hiding rate. This test can meet a prescribed false alarm

probability whatever the image parameters.

- The detection threshold, the probabilities of false alarm and detection of the adaptive AUMP test are analytically calcu lated as the size of the inspected image grows to infinity. This provides an asymptotic upperbound for the detection of hidden bits based on LSB replacement.

To allow a simple mathematical formulation, only the LSB replacement mechanism is studied in this paper. However, since the proposed approach is based on general statistical concepts, it can be extended to more general LSB embedding methods provided that a probabilistic description of the data hiding scheme is available. Compared to the previous published works [27]–[31], this paper presents three main differences: i) it is based on the fact that the quantization step vanishes as the size of the inspected image grows to infinity; ii) it studies the AUMP criterion of optimality; and iii) it explicitly considers that the pixel variance is unknown. Finally, as a corollary of this paper, it is shown that the WS detector [17] coincides with the adaptive AUMP test provided that the tuning parameters of the WS detector are conveniently chosen. Hence, the proposed approach theoretically justifies *a posteriori* the performance of the WS detector.

D. Organization of the Paper

The paper is organized as follows. Section II starts with the problem statement. The problem of hidden bits detection is de scribed in the framework of statistical hypothesis testing theory based on parametric models. Next, Section III proposes a sta tistical model for natural cover images. The statistical detec tion approach developed in the paper is based on this model, especially when it is necessary to estimate the unknown cover image parameters. Section IV presents the AUMP criterion of optimality. It proposes a nonadaptive AUMP test which de tects hidden bits in natural images when the mean and variance of the pixels are known, i.e., when the parameters of the nat ural cover model are known. This test provides an asymptotic upperbound for the detection probability of hidden bits based on the LSB replacement mechanism. In Section V, it is proved that the adaptive version of this test is also AUMP when the mean and variance of the pixels are estimated, provided that the quantization step vanishes and the image size becomes ar bitrarily large. In practice, the conditions of asymptotic conver gence are not totally satisfied and a theoretical formula, which estimates the error in the asymptotic approximation, is also pro posed. Section VI studies the numerical performances of the proposed detection algorithm on artificial and real natural im ages. Some comparisons with other detectors are also presented. Finally, Section VII concludes this paper. The appendices give the proofs of the two theorems presented in the paper.

IV. Conclusion

This paper proposes an original approach to detect hidden bits in grayscale natural images. The proposed approach is based on a simplified parametric model of natural images and it ex ploits the image structure. When the parameters of the model are known, Theorem 1 proposes a nonadaptive AUMP test which maximizes the probability of detection, whatever the hiding rate, when the quantization step vanishes as the number of pixels be comes arbitrarily large. Hence, it provides an asymptotic upper bound for the detection of hidden bits based on the LSB replace ment mechanism. In real situations, the image parameters are es timated and Theorem 2 proposes an adaptive AUMP test which maximizes the probability of detection whatever the hiding rate and the true image parameters. Numerical experiments on real images and comparisons with existing detection algorithms con firm the statistical performances of the test.

References

[1] H. Sencar, M. Ramkumar, and A. Akansu, *Data Hiding Fundamentals and Applications: Content Security in Digital Multimedia*. Elsevier: Academic, 2004.

[2] I. Cox, M. Miller, J. Bloom, J. Fridrich, and T. Kalker, *Digital Water marking and Steganography*. San Francisco, CA: Morgan Kaufmann, 2007.

[3] J. Fridrich, *Steganography in Digital Media—Principles, Algorithms, and Applications*. New York: Cambridge Univ. Press, 2009.

[4] R. Böhme, *Advanced Statistical Steganalysis*. New York: Springer, 2010.

[5] N. Provos and P. Honeyman, “Hide and seek: An introduction to steganography,” *IEEE Secur. Priv. J.*, vol. 1, no. 3, pp. 32–44, 2003.

[6] X.Y. Luo, D.S. Wang, P. Wang, and F.L. Liu, “A review on blind detection for image steganography,” *Signal Process.*, vol. 88, no. 9, pp. 2138–2157, Sep. 2008.

[7] A. Nissar and A. Mir, “Classification of steganalysis techniques: A study,” *Digit. Signal Process.*, vol. 20, no. 6, pp. 1758–1770, 2010.

[8] X. X. Qin and M. Wang, “A review on detection of LSB matching steganography,” *Inf. Technol. J.*, vol. 9, pp. 1725–1738, 2010.

[9] A. D. Ker, “Locating steganographic payload via WS residuals,” in *ACM Proc. 10th Multimed. Secur. Workshop*, 2008, pp. 27–31.

[10] A. D. Ker, “A general framework for the structural steganalysis of LSB replacement,” in *Proc. 7th Inf. Hiding Workshop, ser. Springer LNCS*, 2005, vol. 3727, pp. 296–311.

[11] J. Fridrich, M. Goljan, and R. Du, “Detecting LSB steganography in color and grayscale images,” *IEEE Multimedia*, vol. 8, no. 4, pp. 22–28, 2001.

[12] S. Dumitrescu, X. Wu, and Z. Wang, “Detection of LSB steganography via sample pair analysis,” *IEEE Trans. Signal Process.*, vol. 51, pp.
1995–2007, Jun. 2003.

[13] A. Ker, “Steganalysis of LSB matching in grayscale images,” *Signal Process. Lett.*, vol. 12, no. 6, pp. 441–444, Jun. 2005.

[14] A. D. Ker, “A fusion of maximum likelihood and structural steganal ysis,” in *Proc. 9th Inf. Hiding Workshop, ser. Springer LNCS*, 2007, vol. 4567, pp. 204–219.

[15] K. Lee, A. Westfeld, and S. Lee, “Generalised category attack—Im proving histogrambased attack on JPEG LSB embedding,” *Inf. Hiding’07*, pp. 378–391, 2007.

[16] J. Fridrich and M. Goljan, “On estimation of secret message length in LSB steganography in spatial domain,” in *Secur., Steganogr. Water marking of Multimed. Contents VI, ser. Proc. SPIE,* 2004, vol. 5306, pp. 23–34.