Least Significant Bit Matching Steganalysis Based on Feature Analysis

Lalit Kumar Vashishtha Tanima Dutta Arijit Sur
Department of Computer Science and Engineering
Indian Institute of Technology Guwahati, India
Email: {v.lalit, d.tanima, arijit}@iitg.ernet.in

***Abstract*—Steganography is a science of hiding messages into multimedia documents. In steganography, there is a technique in which the least significant bit is modified to hide the secret message, known as the least significant bit (LSB) steganography. Several steganalyzers are developed to detect least significant bit (LSB) matching steganography. Least significant bit matching images are still not well detected, especially, at low embedding rate. In this paper, we have improved the least significant bit steganalyzers by analyzing and manipulating the features of some existing least significant bit matching steganalysis techniques. A comprehensive set of experiments is carried out to justify proposed method's applicability and evaluate its performance against the existing least significant bit matching steganalysis techniques.**

***Index Terms*—Least significant bit (LSB) matching, median filter, wiener filter, steganalysis, feature set, steganography, histogram**

I. Introduction

The science of hiding messages into multimedia documents is called steganography. Although, the embedded message may change the characteristics and statistical nature of the document, it is required that these changes are difficult to identified, except by the sender and the intended receiver.

Steganalysis develops theories, methods and techniques that can be used to detect hidden messages in multimedia documents. The documents without any hidden messages are called cover documents and the documents with hidden messages are denoted by stego documents. In steganography, there is a technique in which the least significant bit is modified to hide the secret message, this technique is known as the least significant bit (LSB) steganography or LSB embedding. It can be categorized as follows:

- LSB Replacement Steganography
- LSB Matching Steganography

In LSB replacement steganography, the LSB of cover pixel is replaced by a bit of the secret message (before embedding, the secret message is converted into the stream of bits) in a pseudo random order generated by a shared secret key. LSB replacement steganography can be easily detected even if the embedding rate (secret message bit embedded per pixel) is very low [1]. In LSB replacement steganalysis, not only the presence of the secret message can be detected, but also, length of the message can be estimated.

978-1-4673-5952-8/13/$31.00 © 2013 IEEE

In LSB matching, which is an improved version of LSB replacement, 1 is randomly either added to or subtracted from the cover pixel value if the secret message bit does not match the LSB of cover pixel. LSB matching is harder to detect than the LSB replacement. Most of the LSB matching steganalyzers first select a set of statistical features and then trains a supervised learning classifier based on these statistical features to classify stego image from cover image.

Assuming steganographic embedding adds some noise to the original cover images, a set of filters is used to predict cover images from stego images. The difference between stego features and predicted cover features are used to train the supervised learning classifier. In this paper, we extract the features from techniques proposed in [2], [3].

The paper is organized as follows. In Sec. II, a brief description of several LSB Matching Steganalyzers are given. The proposed scheme is mentioned in Sec. III. Sec. IV shows a comprehensive sets of results. Finally, the paper is concluded in Sec. V.

II. LSB Matching Steganalysis

Authors of [4] have modeled the hiding of data as an additive noise process. They proposed a steganalysis method for RGB images based on Histogram Characteristic Function Center of Mass (HCF COM) for the detection of additive noise based steganography. According to them, histogram of stego message is the convolution of noise probability mass function (PMF) and the original histogram. The feature set for their experiment is quite small.

Ker [5] has pointed out that the method proposed by Harmsen and Pearlman in [4] have performed well for detecting LSB matching in RGB color images, but it is not reliable for gray scale images. Ker [5] has showed that for gray scale images, although HCF COM decreases after LSB matching embedding, its range for cover and stego images are heavily overlapped. Hence, the HCF COM cannot distinguish well between the cover and the stego for gray scale images. In order to overcome this shortcoming, Ker [5] proposed an approach based on calibration (down sample) technique, which is much reliable in this respect. The main idea of Ker [5] is use a procedure of down sample that can reduce embedding noise.

Li *et al.* [1] have improved two detectors for detecting LSB matching: (a) Calibrated HCF COM, and (b) Calibrated Adjacency HCF COM. Instead of using center of mass (COM) of HCF, they have considered the ratio of the histogram's DFT coefficients of the image to the corresponding coefficients of the down-sampled image. They have proposed down-sampling only for non oscillating pixels.

Qin *et al.* [6] have proposed a method in which difference between the neighboring pixels (DNPs) and difference between local extrema (DLENs) and their neighbors in gray scale histogram are used as distinguishing features and SVM is adopted to construct classifier.

Abolghasemi *et al.* [7] have shown for gray scale images that the changes due to LSB matching embedding are limited up to five least significant bits (approximately 95%). In [7], Abolghasemi *et al.* have removed 3 MSBs from 8 bit gray scale images. They have created co-occurrence matrix in $0°$, $45°$, $90°$ and $135°$ and have done average on these matrices. They have extracted the features along the diagonal ($135°$). The gray-levels of the neighbor pixels in natural images are highly correlated and the gray-level co-occurrence matrix of the natural image tends to be diagonally distributed. After the data embedding, the high concentration along the main diagonal of gray-level co-occurrence matrix spreads because the high correlation between the pixels in the original image have been reduced.

Tao *et al.* [3] have proposed a steganalysis method which is designed to detect the presence of LSB matching LSB matching steganography in gray scale images. They have 'used four types of distinguishing features which are derived from the pixel histogram, the 1D, 2D and 3D joint pixel difference histogram respectively.

Liu *et al.* [8], [2] have proposed a scheme based on feature mining and pattern classification to detect LSB matching steganography in gray scale images.

Xia *et al.* [9] have proposed a method, which is based on neighborhood node degree histogram characteristic function (NDHCF), in which they take neighborhood node degree in 3×3 and 5×5 neighborhood. They have calculated the center of mass of neighborhood node degree histogram in $1D$ and $2D$. They have observed that after embedding cover and stego images, the value of the neighborhood node degree of cover images is more than that of the stego images, have considered the relative difference of neighborhood node degree of cover images and that of the stego image as 1 feature in 3×3 and 5×5 neighborhood.

We have used the features of [3], as feature set 1. and features of paper [2] as feature set 2 in our work. The feature set 1 [3] consists of following features:

1) Features derived from pixel histogram
2) Features derived from the 1D pixel difference histogram
3) Features derived from 2D joint pixel difference histogram
4) Features derived from 3D joint pixel difference histogram

While, the feature set 2 [2] consists of following features:

1) Entropy and high order statistics of the histogram of the nearest neighbors
2) Correlation features
    a) The correlation between the least significant bit plane (LSBP) and second least significant bit plane (LSBP2) and correlation in the LSBP
    b) The autocorrelation in the image histogram
    c) The correlation in the difference between the image and the de-noised version

III. Proposed Scheme

It is observed from the literature that the domain from which the steganalytic features are taken, plays an important role in the accuracy of steganalytic detection. In some recent steganalytic approaches [10], [11], cover images are predicted from the stego images and then the steganalytic features are taken from the difference between stego and predicted cover images. Intuitively this difference domain represents the steganalytic noise component and features taken from these noise domains are more sensitive to the steganalytic embedding. In this paper, a new way of feature generation is proposed. Intuitively, if no secret bits are embedded in the images under suspicion, the feature values from stego images should be identical feature values generated from predicted cover images. In this case, difference between features from stego and features from predicted cover will be around zero. On the other hand, if secret bits are embedded in the images under suspicion, this difference between features should have values higher than zero. Thus this feature difference itself can be used as features for a steganalytic classifier.

A. Cover Image Prediction using Image De-noising

Since steganpgraphic noise is treated as additive noise, different filters are used to predict the cover images by de-noising the stego images. In this paper, two filters are used as de-noising filters, namely median filter and wiener filter.

1) Median Filtering: The median filter is a nonlinear digital filtering technique, often used to remove noise. Such noise reduction is a typical pre-processing step to improve the results of later processing (for example, edge detection on an image). In median filtering, the neighboring pixels are ranked according to intensity values and the median value is considered as the new value for central pixel [12].

2) Wiener Filtering: In recent past, Fridrich *et al.* proposed a spatial domain steganalyzer [13] where the stego image is assumed to be an additive mixture of a non-stationary Gaussian signal (the cover image) and a stationary Gaussian signal with a known variance (the noise). Authors have used a wiener filter [14] for de-noising such kind of noise to predict the cover images. The Wiener filter, first proposed by Norbert Wiener [15] during the 1940s and published in 1949, is used to reduce the amount of noise present in a signal by comparison with an estimation of the desired noiseless signal.

B. Steganalytic Features

In the proposed scheme, two sets of features are generated. Firstly, the difference between features from stego images and features from predicted cover images are calculated. A

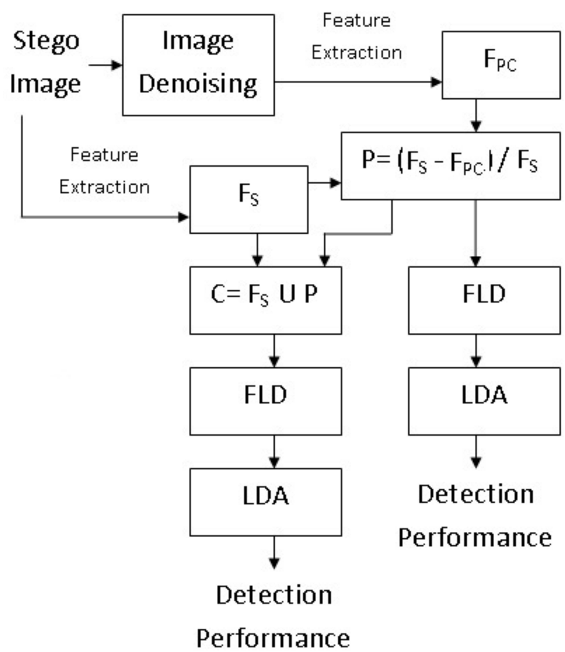

Fig. 1. Block diagram of the proposed scheme

normalized version of these difference, called relative feature (P), is calculated as follows:

$$P = \frac{F_S - F_{PC}}{F_S} \tag{1}$$

where F_S is the features taken from stego images whereas F_{PC} is taken from predicted cover images. Secondly, two feature set P and F_S are merged to get another feature set (say C). C feature set is double dimension of feature set P.

$$C = P \cup F_S \tag{2}$$

C. Steganalytic Classifier

First, the Fisher Linear Discriminant (FLD) is used to reduce the dimension of the feature space to single dimension. Then the Linear Discriminant Analysis (LDA) classifier [10] is trained using this single dimension feature.

A block diagram of the proposed scheme is depicted in Fig. 1.

IV. Experimental Results

In this section, experimental results are presented to demonstrate the performance of the proposed method by comparing it with feature set 1 (Tao *et al.*'s scheme [3]) and feature set 2 (Liu et al's scheme [2]). A set of 200 uncompressed images [16], all the images are uncompressed digital TIFF files with size 512×384 or 384×512, are used as image dataset. Random sequence is taken as embedding message and it is embedded in images. Table 1 shows the experimental setup. From the experimental results, it is observed that the proposed scheme outperforms other two methods ([2], [3]) for the uncompressed images.

TABLE I
Table shows the experimental setup

Data set	200 8-bit gray scale images
Embedding rate	0.08/0.20
Classifier used	LDA
Compared by	ROC
LSB matching scheme	• Median Filter Vs [3] schemes • Median Filter Vs [2] schemes • Median Filter Vs Concatenation of [3] and [2] schemes • Wiener Filter Vs [2] schemes • Wiener Filter Vs [3] schemes • Wiener Filter Vs Concatenation of [3] and [2] schemes

Several experiments are performed for different parameters (like Noise Suppression, Relative Difference, and Concatenated Feature Set) using Median and Wiener filtering for feature set 1 and feature set 2. Results of these experiments are described below:

1) Noise Suppression using Median Filtering: In the first experiment, it is shown that the detection performance of an existing steganalyzer (here, Tao *et al.*'s Scheme [3] is used as steganalyzer) is decreased if it applies on median filtered version of stego images. In Fig. 2(a), it is observed that the detection performance of Tao *et al.*'s attack on stego images (red colored graph '–') is higher than detection performance of same attack on median filtered version of same stego images (black colored graph '–'). It implies that median filtering removes some of stegangraphic noise such that it is less detectable. In Fig. 2(a) experiment, embedding rate is 0.08 bpp (bit per pixel). Results of the same experiment but with embedding rate 0.20 is presented in Fig. 2(b).
   Same experiment as above (alike, Tao *et al.*'s Scheme [3]) is done here with another existing steganalyzer (Liu *et al.* [2]). In Fig. 2(c) (@ embedding rate 0.08bpp) and Fig. 2(d) (@ embedding rate 0.20 bpp), it is shown that detection performance of steganalyzer is substantially decreased when it applied on median filtered version of stego images.
2) Relative Difference using Median Filtering: In second experiment, relative difference feature set (P) using Median Filtering (refer to Sec. III-A1) is used as steganalytic features. In Fig. 2(e) and Fig. 2(f), it is observed that proposed steganalyzer outperforms Tao *et al.*'s method [3] for embedding rate 0.08 bpp and 0.20 bpp, respectively.
   Again, same experiment as above (alike, Tao *et al.*'s Scheme [3]) is done for Liu *et al.*'s Scheme [2]. Fig. 2(g) and Fig. 2(h), shows the proposed scheme gives comparative results with existing scheme for embedding rate 0.08 bpp and 0.20 bpp, respectively.
3) Concatenated Feature Set using Median Filtering: In third experiment, concatenated feature set (C) using Median Filtering (refer to Sec. III-A1) is used as steganalytic features. In Fig. 2(i) and Fig. 2(j), it is observed

that proposed steganalyzer outperforms Tao et. al.'s method [3] for embedding rate 0.08 bpp and 0.20 bpp respectively.

Similarly, the above experiment is performed for Liu *et al.*'s Scheme [2]. In Fig. 2(k) and Fig. 2(l), the proposed scheme gives comparative results with existing scheme for embedding rate 0.08 bpp and 0.20 bpp, respectively.

4) Noise Suppression using Wiener Filtering: In this experiment, the detection performance of an existing steganalyzer is decreased if it applies on wiener filtered version of stego images. It implies that wiener filtering (refer to sec. III-A2) removes some of stegangraphic noise such that it is less detectable. In Fig. 2(n) experiment, embedding rate is 0.20 bpp (bit per pixel) which shows that the proposed scheme (black colored graph '–') outperforms Tao *et al.*'s method [3] (red colored graph '–'). Results of the same experiment but with embedding rate 0.08 is presented in Fig. 2(m).

   Again, the same experiment, as mentioned above, is performed for Liu *et al.*'s Scheme [2]. Fig. 2(p), for embedding rate 0.20 bpp, shows that the proposed scheme gives better results than existing scheme whereas Fig. 2(o) (embedding rate 0.08 bpp) gives comparative results.

5) Relative Difference using Wiener Filtering: In this experiment, relative difference feature set (P) using Wiener Filtering (refer to Sec. III-A2) is used as steganalytic features. In Fig. 2(q) and Fig. 2(r), it is observed that proposed steganalyzer outperforms Tao *et al.*'s method [3] for embedding rate 0.08 bpp and 0.20 bpp, respectively.

   Similarly, the above experiment is performed for Liu *et al.*'s Scheme [2]. In Fig. 2(s) and Fig. 2(t), the proposed scheme gives comparative results with existing scheme for embedding rate 0.08 bpp and 0.20 bpp, respectively.

6) Concatenated Feature Set using Wiener Filtering: In last experiment, concatenated feature set (C) using Wiener Filtering (refer to Sec. III-A2) is used as steganalytic features. In Fig. 2(u) and Fig. 2(v), it is observed that proposed steganalyzer outperforms Tao et. al.'s method [3] for embedding rate 0.08 bpp and 0.20 bpp respectively.

   Same experiment as mentioned above is performed for Liu *et al.*'s Scheme [2]. In Fig. 2(w) and Fig. 2(x), the proposed scheme gives comparative results with existing scheme for embedding rate 0.08 bpp and 0.20 bpp, respectively.

V. Conclusions

In this paper, a new way of feature derivation is proposed. It is experimentally shown that proposed method outperforms existing method most of the time at different embedding rates. It is also experimentally shown that detection performance is increased when newly derived features are concatenated with existing features for different steganalyzers.

References

[1] X. Li, T. Zeng, and B. Yang, "A Further Study on Steganalysis of LSB Matching by Callibration," in *Proc. Intl. Conference on Image Processing*, 2009, pp. 2072–2075.

[2] Q. Liu, A. Sung, Z. Chen, and J. Xu, "Feature Mining and Pattern Classification For Steganalysis of LSB Matching Steganography in Grayscale Images," *Pattern Recognition*, vol. 41, pp. 56–66, 2008.

[3] Z. Tao, Z. Yan, L. Wenxiang, and P. Xijian, "Steganalysis of LSB Matching Exploiting High-Dimensional Correlations Between Pixel Differences," in *Proc. 5th Intl. Conference On Natural Computation*, 2009, pp. 307–310.

[4] J. Harmsen and W. Pearlman, "Steganalysis of Additive Noise Modelable Information Hiding," *SPIE Information Technology*, 2003.

[5] A. Ker, "Steganalysis of LSB Matching in Grayscale Images," *IEEE Signal Processing Letter*, vol. 12, pp. 441–444, 2005.

[6] J. Qin, X. Su, X. Xiang, and Z. Xia, "Steganalysis Based on Difference Statistics For LSB Matching Steganography," *Information Technology*, vol. 8, no. 8, pp. 1281–1286, 2009.

[7] M. Abolghasemi, H. Aghainia, K. Faez, and M. Mehrabi, "Steganalysis of LSB matching Based on Co-Occurrence Matrix And Removing Most Significant Bit Planes," in *Proc. Intl. Conference on Intelligent Information Hiding and Multimedia Signal Processing*, 2008, pp. 1527–1530.

[8] Q. Liu, A. Sung, Z. Chen, and J. Xu, "Image Complexity and Feature Extraction For Steganalysis of LSB Matching," in *Proc. 18th Intl. Conference on Pattern Recognition*, vol. 2, 2006, pp. 267–270.

[9] B. Xia, X. Sun, and J. Qin, "Steganalysis Based on Neighborhood Node Degree Histogram for LSB Matching Steganography," in *Proc. Intl. Conference on Multimedia Information Networking and Security*, 2009, pp. 79–82.

[10] J. Fridrich, "Feature-Based Steganalysis for JPEG Images and its Implications for Future Design of Steganographic Schemes," in *Proc. 6th Intl. Workshop on Information Hiding*, May 2004, pp. 67–81.

[11] T. Pevny and J. Fridrich, "Merging Markov and DCT features for multiclass JPEG steganalysis," in *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. 6505, January 2007, pp. 03–04.

[12] G.R.Arce, *Nonlinear Signal Processing: A Statistical Approach*. Wiley: New Jersey, USA, 2005.

[13] J. F. M. Goljan and T. Holotyak, "New blind steganalysis and its implications," in *Proc. SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents VIII*, vol. 6072, January 2006, pp. 1–13.

[14] A. Khireddine, K. Benmahammed, and W. Puech, "Digital image restoration by Wiener filter in 2D case," *Advances in Engineering Software*, vol. 38, no. 7, pp. 513–516, 2007.

[15] N. Wiener, *Extrapolation, Interpolation, and Smoothing of Stationary Time Series*. New York: Wiley, 1949.

[16] J. L. J. Wang and G. Wiederhold, "SIMPLicity : Semantics-sensitive integrated matching for picture LIbraries," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 9, pp. 947–963, September 2001.

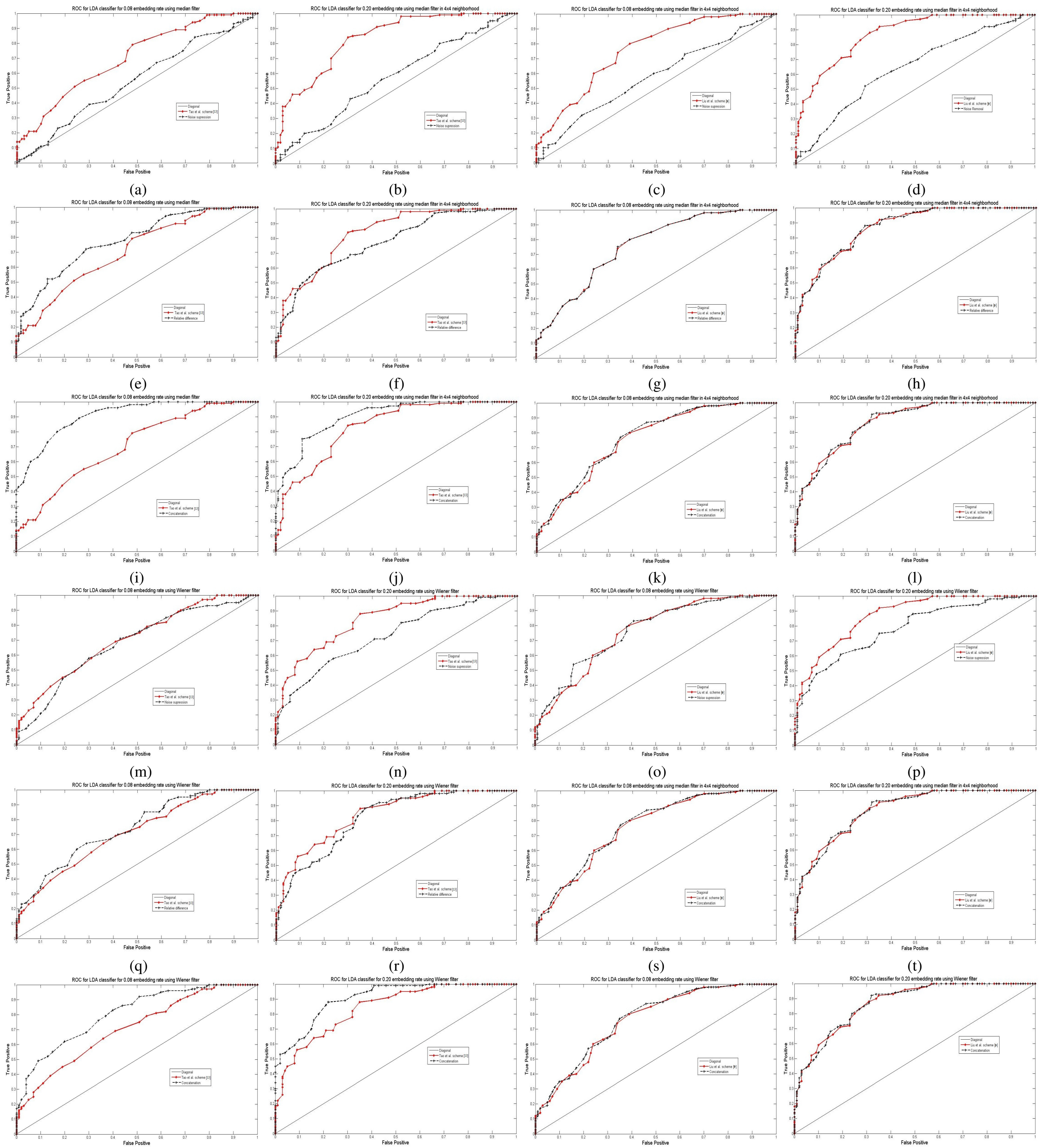

Fig. 2 ROC for 8% and 20% embedding for median filter scheme for noise suppression Vs Tao *et al.* [3] scheme (a, b) and Liu *et al.* [2] scheme (c, d), 8% and 20% embedding for relative difference by median filter scheme Vs Tao *et al.* [3] scheme (e, f) and Liu *et al.* [2] scheme (g, h), 8% and 20% embedding for concatenation by median filter scheme Vs Tao *et al.* [3] scheme (i, j) and Liu *et al.* [2] scheme (k, l), 8% and 20% embedding for wiener filter scheme for noise suppression Vs Tao *et al.* [3] scheme (m, n) and Liu *et al.* [2] scheme (o, p), 8% and 20% embedding for relative difference by wiener filter scheme Vs Tao *et al.* [3] scheme (q, r) and Liu *et al.* [2] scheme (s, t), and 8% and 20% embedding for concatenation by wiener filter scheme Vs Tao *et al.* [3] scheme (u, v) and Liu *et al.* [2] scheme (w, x).