

LSB Matching Revisited

Jarno Mielikainen, *Member, IEEE*

Abstract—This letter proposes a modification to the least-significant-bit (LSB) matching, a steganographic method for embedding message bits into a still image. In the LSB matching, the choice of whether to add or subtract one from the cover image pixel is random. The new method uses the choice to set a binary function of two cover pixels to the desired value. The embedding is performed using a pair of pixels as a unit, where the LSB of the first pixel carries one bit of information, and a function of the two pixel values carries another bit of information. Therefore, the modified method allows embedding the same payload as LSB matching but with fewer changes to the cover image. The experimental results of the proposed method show better performance than traditional LSB matching in terms of distortion and resistance against existing steganalysis.

Index Terms—Information hiding, steganography.

I. INTRODUCTION

STEGANOGRAPHY is an application of information hiding. Generally, a good steganographic technique should have good visual/statistical imperceptibility and a sufficient payload [1]. The proposed method is a modified version of the least-significant-bit (LSB) method, resulting in improved imperceptibility.

LSB replacement embeds a message into the cover image by replacing the LSBs of the cover image with message bits to arrive at the stego image. The method increases even pixel values either by one or leaves them unmodified, while odd values are left unchanged or decreased by one. As a result, there exists an imbalance in the embedding distortion in the stego image. Due to this imbalance, LSB replacement can be detected by the current detection methods, e.g., [2]. The terminology LSB replacement/LSB matching was introduced in [2].

LSB matching also modifies the LSBs of the cover image for message embedding. LSB matching does not simply replace the LSBs of the cover image as LSB replacement does. Instead, if the message bit does not match the LSB of the cover image, then one is randomly either added or subtracted from the value of the cover pixel. To ensure the invertibility of the process, pixel values are never modified to fall outside of the allowable range [3].

Only a few detection methods for LSB matching have been proposed. The best-known detector for LSB matching is based on the center of mass (COM) of the histogram characteristic function (HCF) introduced by Harmsen *et al.* [4]. Ker [5] improved the detection rates for gray-scale images by using a

down-sampled image for calibration. He also proposed using the adjacency histogram instead of the usual histogram.

The proposed method allows an embedding of the same amount of information as LSB matching but with fewer changes to the cover image. This makes the detection of the proposed method harder than the conventional LSB matching method.

II. PROPOSED METHOD

The proposed method uses gray-scale cover images. The message embedding is performed for the two cover image pixels at a time. The gray-level values of those two pixels are x_i and x_{i+1} . After the message embedding, the value of the i th message bit m_i is equal to the LSB of stego image's i th pixel y_i . The value of the $i+1$ th message bit m_{i+1} is a function of y_i and y_{i+1} . The proposed method allows a selection of addition/subtraction of y_i to carry information, because the selection can set a binary function $f(y_i, y_{i+1})$ to the desired value. If a binary function has the following property

$$f(l-1, n) \neq f(l+1, n), \quad \forall l, n \in \mathbb{Z} \quad (1)$$

then a controlled change of the value of l by one allows the setting of $f(l, n)$ to the desired value. If a binary function $f(l, n)$ is of the form

$$f(l, n) \neq f(l, n+1), \quad \forall l, n \in \mathbb{Z} \quad (2)$$

then both an increase and a decrease of n by one will change the value of the function $f(l, n)$. The function $f(y_i, y_{i+1}) = \text{LSB}(\lfloor y_i/2 \rfloor + y_{i+1})$ has the properties of (1) and (2).

The embedding algorithm for a pair of pixels is presented in Fig. 1. The embedding cannot be performed for saturated pixels, i.e., pixels that have either a minimal or maximal allowable value. Table I shows an example of embedding message bits into the cover image. Four of the resulting stego pixels y_{i+1} could be set to either one of the two alternative values. A step in the algorithm $y_{i+1} = x_{i+1} \pm 1$ either adds or subtracts one from the cover image pixel x_{i+1} . The choice of whether to add or subtract one is performed separately, both for the even- and odd-valued regions. The even- and odd-valued regions refer to the areas of the cover image where the pixel values are even and odd, respectively. Inside both regions, the increment/decrement selection is made to minimize the absolute value of the differences between the cover image and the stego image. The selection thus avoids the LSB replacement style imbalance. The message embedding order is determined by the same pseudo-random sequence generator as in LSB matching [3].

The expected number of modifications per pixel can be computed using

$$\frac{P(x_i \neq m_i) + P(x_i = m_i)P(f(x_i, x_{i+1}) \neq m_{i+1})}{2} \quad (3)$$

Manuscript received October 6, 2005; revised December 6, 2005. This work was supported by the Academy of Finland. The associate editor coordinating the review of this manuscript and approving it for publication was Dr. Min Wu.

The author is with the Department of Information Technology, Lappeenranta University of Technology, 53851 Lappeenranta, Finland (e-mail: mielikai@lut.fi).

Digital Object Identifier 10.1109/LSP.2006.870357

```

input: a pair of cover image pixels  $x_i, x_{i+1}$ 
      two message bits  $m_i, m_{i+1}$ 
output: a pair of stego image pixels  $y_i, y_{i+1}$ 

if  $m_i = \text{LSB}(x_i)$ 
  if  $m_{i+1} \neq f(x_i, x_{i+1})$ 
     $y_{i+1} = x_{i+1} \pm 1$ 
  else
     $y_{i+1} = x_{i+1}$ 
  end
   $y_i = x_i$ 
else
  if  $m_{i+1} = f(x_i - 1, x_{i+1})$ 
     $y_i = x_i - 1$ 
  else
     $y_i = x_i + 1$ 
  end
   $y_{i+1} = x_{i+1}$ 
end

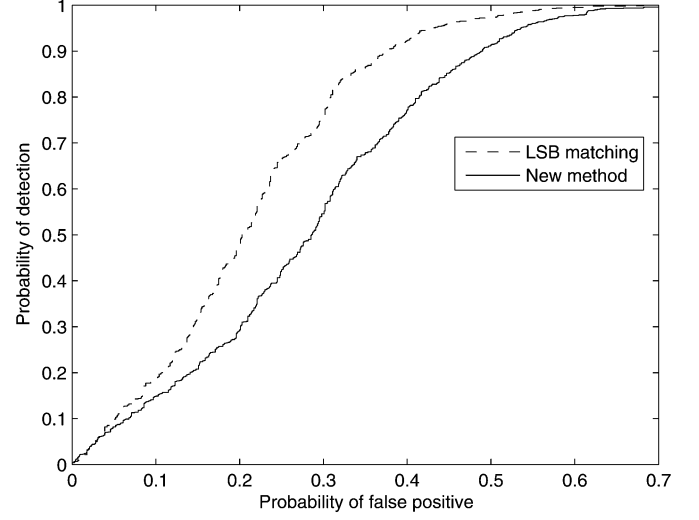
```

Fig. 1. Embedding algorithm for a pair of pixels.

TABLE I
EXAMPLE OF EMBEDDING MESSAGE BIT PAIRS (m_i, m_{i+1}) INTO
COVER IMAGE PIXEL PAIRS (x_i, x_{i+1}), RESULTING IN STEGO
IMAGE PIXEL PAIRS (y_i, y_{i+1})

x_i	x_{i+1}	m_i	m_{i+1}	y_i	y_{i+1}
1	1	0	0	2	1
1	1	0	1	0	1
1	1	1	0	1	0 or 2
1	1	1	1	1	1
1	2	0	0	0	2
1	2	0	1	2	2
1	2	1	0	1	2
1	2	1	1	1	1 or 3
2	1	0	0	2	1
2	1	0	1	2	0 or 2
2	1	1	0	3	1
2	1	1	1	1	1
2	2	0	0	2	1 or 3
2	2	0	1	2	2
2	2	1	0	1	2
2	2	1	1	3	2

where $P()$ is the probability function. If ones and zeros are equally likely both in the message and in the LSBs of the

Fig. 2. ROC curves for the calibrated HCF COM. The x -axis has been scaled to focus on a region of interest.

cover image and they have no correlation between them, then $P(x_i \neq m_i) = P(x_i = m_i) = P(f(x_i, x_{i+1}) = m_{i+1}) = 0.5$. Therefore, the expected number of modifications per pixel is 0.375.

III. EXPERIMENTAL RESULTS

One thousand JPEG images [6], all sized 384×256 , were used in our experiments. JPEG images were used as the cover images, because the HCF COM detectors work best on them [5]. The images were converted to gray-scale before use. The cover images were embedded with maximal-length random messages, such as messages obtained by encryption.

On average, LSB matching and the new method required 0.500 and 0.375 modifications per pixel, respectively. The experimental results agree with the theoretical results.

Also, the absolute values of the sum of differences per pixel between the cover and stego image for the even and odd regions were computed. These experiments were used to determine if the new method exhibits bias. For both regions, the average results were 0.500 and 0.000 for LSB replacement and the new method, respectively. The results confirm that the new method does not have LSB replacement style imbalance.

Fig. 2 depicts receiver operating characteristic (ROC) curves using the calibrated HCF COM for message detection. The curves show how the probability of detection and probability of false positive vary as the detection threshold is adjusted. Similarly, in Fig. 3, ROC curves for the calibrated adjacency HCF COM are given. From the ROC curves, it can be seen that the novel embedding algorithm decreases the probability of detection for the HCF COM detectors compared to LSB matching.

IV. CONCLUSION

The proposed steganographic method allows an embedding of the same amount of information into the stego image as LSB matching. At the same time, the number of changed pixel values is smaller. The proposed method does not have the asymmetric

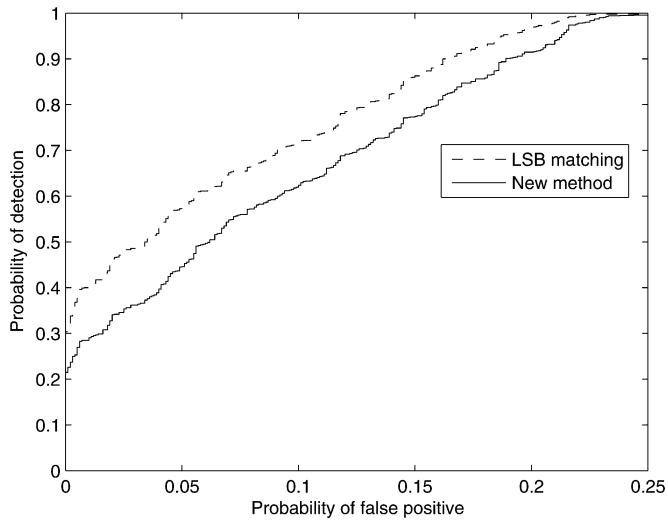


Fig. 3. ROC curves for the calibrated adjacency HCF COM. The x -axis has been scaled to focus on a region of interest.

property of LSB replacement method. Therefore, it is immune against steganographic attacks that utilize the asymmetric property. The new embedding procedure could be used for any discrete-valued cover medium, not just images. Finally, the detection of the existence of the hidden messages using the HCF COM-based detectors is less efficient against the method compared to LSB matching. Future work includes extending the

method for color images. Also, the message embedding order will be researched in order for the method to work better with saturated pixels and to spread the distortion more evenly over the image.

ACKNOWLEDGMENT

The author would like to thank the anonymous reviewers for suggestions, which improved the quality of this letter. The author would also like to thank Prof. S. K. Mitra, who took the time to carefully review the letter.

REFERENCES

- [1] X. Zhang and S. Wang, "Steganography using multiple-base notational system and human vision sensitivity," *IEEE Signal Process. Lett.*, vol. 12, no. 1, pp. 67–70, Jan. 2005.
- [2] A. Ker, "Improved detection of LSB steganography in grayscale images," in *Proc. Information Hiding Workshop*, vol. 3200, Springer LNCS, 2004, pp. 97–115.
- [3] T. Sharp, "An implementation of key-based digital signal steganography," in *Proc. Information Hiding Workshop*, vol. 2137, Springer LNCS, 2001, pp. 13–26.
- [4] J. Harmsen and W. Pearlman, "Steganalysis of additive-noise modelable information hiding," in *Proc. SPIE Security Watermarking Multimedia Contents*, vol. 5020, 2003, pp. 131–142.
- [5] A. Ker, "Steganalysis of LSB matching in greyscale images," *IEEE Signal Process. Lett.*, vol. 12, no. 6, pp. 441–444, Jun. 2005.
- [6] J. Wang, J. Li, and G. Wiederhold, "SIMPLicity: Semantics-sensitive integrated matching for picture Libraries," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 9, pp. 947–963, Sep. 2001.