

# MINIMIZING ADDITIVE DISTORTION FUNCTIONS WITH NON-BINARY EMBEDDING OPERATION IN STEGANOGRAPHY

Tomáš Filler and Jessica Fridrich

Department of ECE, SUNY Binghamton, NY, USA  
{tomas.filler, fridrich}@binghamton.edu

## ABSTRACT

Most practical steganographic algorithms for empirical covers embed messages by minimizing a sum of per-pixel distortions. Current near-optimal codes for this minimization problem [7] are limited to a binary embedding operation. In this paper, we extend this work to embedding operations of larger cardinality. The need for embedding changes of larger amplitude and the merit of this construction are confirmed experimentally by implementing an adaptive embedding algorithm for digital images and comparing its security to other schemes.

## 1. INTRODUCTION

In steganography, a secret message is embedded in a cover object  $\mathbf{x} = (x_1, \dots, x_n) \in \mathcal{X} = \{\mathcal{I}\}^n$  by slightly modifying its individual elements to produce the stego object  $\mathbf{y} = (y_1, \dots, y_n) \in \mathcal{Y} = \mathcal{I}_1 \times \mathcal{I}_2 \times \dots \times \mathcal{I}_n$ ,  $\mathcal{I}_i \subset \mathcal{I}$ , where  $\mathcal{I}_i$  is the range of the embedding operation at element  $i$  and  $x_i \in \mathcal{I}_i$ . For example, for the Least Significant Bit (LSB) replacement method,  $\mathcal{I}_i = \{x_i, \bar{x}_i\}$ , where  $\bar{x}_i$  is  $x_i$  after flipping its LSB. The embedding operation is *binary* if  $|\mathcal{I}_i| = 2$  or *ternary* if  $|\mathcal{I}_i| = 3$  for all  $i$ . For concreteness, we will call  $\mathbf{x}$  image and  $x_i$  its  $i$ th pixel but other interpretations are certainly possible. For example,  $x_i$  may represent an RGB triple in a color image, a DCT coefficient, etc.

Steganographic schemes for complex cover sources, such as digital images, are usually constructed to minimize some distortion measure  $D$  between  $\mathbf{x}$  and  $\mathbf{y}$  [9] that is assumed to be related to statistical detectability of embedding changes. In this paper, we will consider the following distortion function

$$D(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \rho_i(\mathbf{x}, y_i), \quad (1)$$

---

The work on this paper was supported by Air Force Office of Scientific Research under the research grant number FA9550-08-1-0084. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation there on. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied of AFOSR or the U.S. Government.

where  $\rho_i : \mathcal{X} \times \mathcal{I} \rightarrow \mathbb{R} \cup \{\infty\}$  are cost functions satisfying  $\rho_i(\mathbf{x}, y_i) = \infty$  whenever  $y_i \notin \mathcal{I}_i$  and  $\rho_i(\mathbf{x}, x_i) < \infty$ . The cost functions can reflect higher-order dependencies of cover image pixels (because the dependence on  $\mathbf{x}$  is not constrained) but the additivity of (1) cannot capture dependencies among embedding changes. The most common choice of  $\rho_i$  for binary embedding operations is  $\rho_i(\mathbf{x}, y_i) = \varrho_i \cdot [x_i \neq y_i]$  with scalar costs  $\varrho_i$ , where  $[S]$  is the Iverson bracket defined as 1 when the logical statement  $S$  is true and 0 otherwise. Note that when  $\varrho_i = 1$ ,  $D$  is the number of embedding changes. In the MMx algorithm [14],  $\varrho_i$  is dependent on the quantization error of the  $i$ th DCT coefficient.

Since Crandal [3] pointed out the connection between minimization of  $D$  and syndrome coding, many practical algorithms for the binary embedding operation have been proposed, such as those based on Hamming codes [19, 14], BCH codes [17, 20], random codes [13], and their combination [21]. Special codes, called wet paper codes, were proposed for embedding with wet pixels [4, 11] (pixels prohibited to be modified for which  $\mathcal{I}_i = \{x_i\}$ ). The recently proposed Syndrome-Trellis Codes (STCs) [7, 8] unify the approach, achieve near-optimal performance for various distortion costs  $\varrho_i$ , and perform well even with a large number of wet pixels.

Although it is straightforward to extend STCs to non-binary alphabets and thus apply them to  $q$ -ary embedding operations, their complexity rapidly increases (the number of states in the trellis increases from  $2^h$  to  $q^h$  for constraint height  $h$ ), limiting thus their performance in practice. The main contribution of this paper is extending the STCs to arbitrary  $q$ -ary embedding operations using a simple layered construction without any significant increase in complexity. By moving away from binary embedding operations by increasing the size of  $\mathcal{I}_i$ , the embedding modifications may become larger and a larger payload can be embedded. We demonstrate experimentally that by restricting the larger-amplitude embedding changes adaptively to the content, the multi-layered construction embeds larger payloads with lower statistical detectability, countering thus the established belief that the increase in payload does not outweigh the increase in statistical detectability [9].

Even though the relationship between distortion and

steganographic security is far from clear, it makes sense to embed messages by minimizing a heuristically chosen distortion function. At least, this is how today's least detectable image steganographic schemes work [14, 20, 16]. This paper provides a coding scheme how to embed with minimal distortion once the steganographer agreed on the distortion function and the embedding operation. The important question of how to choose both of these elements to minimize detectability is not addressed and is left as a future direction.

Section 2 restates some known relative payload–relative distortion bounds. Section 3 reviews the binary embedding operation and syndrome-trellis codes [7] for implementation. The main part, the multi-layered construction, is described and analyzed in Section 4. Application to spatial domain steganography is described in Section 5. The paper is concluded in Section 6.

All vectors are typed in bold. Random variables and their realizations are denoted using capital and lower case letters, respectively. Furthermore,  $\log(x) = \log_2(x)$  and  $\ln(x)$  denotes the natural logarithm. We use  $h(x) = -x \log x - (1-x) \log(1-x)$  for the binary entropy function.

## 2. PRELIMINARIES

### 2.1. Problem formulation

We assume the cover image  $\mathbf{x}$  to be *fixed* and known only to the sender. The results of this paper will be general and independent of any particular choice of  $\mathbf{x}$ . For this reason, we simply write  $D(\mathbf{y}) \triangleq D(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^n \rho_i(y_i)$ , while all quantities derived from  $\mathcal{Y}$  or  $D(\mathbf{y})$  should be seen as being conditioned by  $\mathbf{x}$ . The distortion function and its parameters are known only to the sender and not to the receiver. We assume that the embedding algorithm replaces the original cover  $\mathbf{x}$  with  $\mathbf{y} \in \mathcal{Y}$  obtained as a realization of a random variable  $\mathbf{Y}$  defined over  $\mathcal{Y}$  and distributed according to  $\pi$ ,  $\pi(\mathbf{y}) \triangleq P(\mathbf{Y} = \mathbf{y})$ . If the receiver knew  $\mathbf{x}$ , the sender could send up to  $H(\pi)$  bits on average while introducing the average distortion  $E_\pi[D]$ , where

$$H(\pi) = - \sum_{\mathbf{y} \in \mathcal{Y}} \pi(\mathbf{y}) \log \pi(\mathbf{y}), \quad E_\pi[D] = \sum_{\mathbf{y} \in \mathcal{Y}} \pi(\mathbf{y}) D(\mathbf{y}).$$

By the construction of the method, the knowledge of  $\mathbf{x}$  to the receiver has no effect on the bounds between the above quantities as long as  $\mathbf{x}$  is known to the sender.

Having defined the distortion function, the sender is interested in the following optimization problems:

- **Payload-limited sender (PLS):** embed a *fixed average payload* of  $m$  bits while minimizing the average distortion,

$$\underset{\pi}{\text{minimize}} E_\pi[D] \quad \text{subject to } H(\pi) = m. \quad (2)$$

- **Distortion-limited sender (DLS):** maximizes the average payload while introducing a *fixed average distortion*  $D_\epsilon$ ,

$$\underset{\pi}{\text{maximize}} H(\pi) \quad \text{subject to } E_\pi[D] = D_\epsilon. \quad (3)$$

Although the PLS is the most common, the DLS is more suitable for steganography as long as  $D$  is related to statistical detectability.

### 2.2. Performance bounds and comparison metrics

The problems described above bear relationship to the problem of source coding with a fidelity criterion as described by Shannon [18]. Problems (2) and (3) are dual to each other, meaning that the optimal distribution for the first problem is, for some value of  $D_\epsilon$ , also optimal for the second one. Following the maximum entropy principle [2, Th. 12.1.1], the optimal solution has the form of a Gibbs distribution [10]

$$\pi(\mathbf{y}) = \frac{\exp(-\lambda D(\mathbf{y}))}{Z(\lambda)} \stackrel{(a)}{=} \prod_{i=1}^n \frac{\exp(-\lambda \rho_i(y_i))}{Z_i(\lambda)} \triangleq \prod_{i=1}^n \pi_i(y_i), \quad (4)$$

where the parameter  $\lambda \in [0, \infty)$  has to be obtained from the corresponding constraints (2) and (3) by solving an algebraic equation,<sup>1</sup> and  $Z(\lambda) = \sum_{\mathbf{y} \in \mathcal{Y}} \exp(-\lambda D(\mathbf{y}))$ ,  $Z_i(\lambda) = \sum_{y_i \in \mathcal{I}_i} \exp(-\lambda \rho_i(y_i))$  are the corresponding partition functions. Step (a) follows from the additivity of  $D$ , which also leads to mutual independence of individual stego pixels  $y_i$  given  $\mathbf{x}$ . Finally, the impact of embedding can be simulated by changing each pixel  $i$  with probability  $\pi_i$ .

An established way of evaluating practical coding algorithms in steganography is to compare the *embedding efficiency*  $e(\alpha) = \alpha n / E_\pi[D]$  for a fixed expected relative payload  $\alpha = m/n$  with the upper bound derived from (4). When the number of changes is minimized,  $e$  is the average number of bits hidden per one change. For general  $\rho_i$ , the interpretation of this metric becomes less clear. A different and more easily interpretable metric is to compare the payload,  $m$ , of an embedding algorithm w.r.t. the payload,  $m_{\text{MAX}}$ , of the optimal DLS for a fixed  $D_\epsilon$ ,

$$l(D_\epsilon) = \frac{m_{\text{MAX}} - m}{m_{\text{MAX}}}, \quad (5)$$

which we call the *coding loss*.

## 3. BINARY EMBEDDING OPERATION

We start by describing the special case of a binary embedding operation and review a practical coding construction for this problem. In Section 4, we generalize this approach to operations with a larger cardinality. Since the operation is binary,

<sup>1</sup>A simple binary search will do the job because both  $H(\pi)$  and  $E_\pi[D]$  are monotone w.r.t.  $\lambda$ .

we assume that  $\mathcal{I}_i = \{x_i, y_i\}$ . In what follows, the only value the receiver needs to know is the number of message bits  $m$  he wants to receive. This information can be communicated in the same stego image using a different embedding scheme.

According to (4), the coding algorithm for (2) or (3) is optimal if and only if it outputs pixel  $y_i$  with probability

$$\begin{aligned} \pi_i(y_i) &= \frac{\exp(-\lambda\rho_i(y_i))}{\exp(-\lambda\rho_i(x_i)) + \exp(-\lambda\rho_i(y_i))} \\ &= \frac{\exp(-\lambda\varrho_i)}{1 + \exp(-\lambda\varrho_i)}, \end{aligned} \quad (6)$$

where  $\varrho_i = \rho_i(y_i) - \rho_i(x_i)$ .<sup>2</sup> For a fixed value of  $\lambda$ , the values  $\varrho_i$ ,  $i = 1, \dots, n$ , form sufficient statistic for  $\pi$ . Because  $D(\mathbf{y}) = \sum_{i=1}^n \rho_i(x_i) + D'(\mathbf{y})$ , (2) is equivalent to the PLS for the same payload  $m$  and distortion

$$D'(\mathbf{y}) = \sum_{i=1}^n \varrho_i \cdot [x_i \neq y_i]. \quad (7)$$

A solution to the PLS with binary embedding operation can be used to derive the following ‘‘flipping lemma’’ that we will heavily use in Section 4.

**The flipping lemma.** *Given a set of probabilities  $\{p_i\}_{i=1}^n$ , the sender wants to communicate  $m = \sum_{i=1}^n h(p_i)$  bits by sending bit strings  $\mathbf{y} = \{y_i\}_{i=1}^n$  such that  $P(y_i = 0) = p_i$ . This can be achieved by a PLS with a binary embedding operation on  $\mathcal{I} = \mathcal{I}_i = \{0, 1\}$  for all  $i$  by embedding the payload in cover  $x_i = [p_i < 1/2]$  with non-negative per-pixel costs  $\varrho_i = \ln(\tilde{p}_i/(1 - \tilde{p}_i))$ ,  $\tilde{p}_i = \max\{p_i, 1 - p_i\}$ .*

*Proof:* Without loss of generality, let  $\lambda = 1$ . Since the inverse of  $f(z) = \ln(z/(1 - z))$  on  $[0, 1]$  is  $f^{-1}(z) = \exp(z)/(1 + \exp(z))$ , by (6) the cost  $\varrho_i$  causes  $x_i$  to change to  $y_i = 1 - x_i$  with probability  $P(y_i \neq x_i | x_i) = f^{-1}(-\varrho_i) = 1 - \tilde{p}_i$ . Thus,  $P(y_i = 0 | x_i = 1) = f^{-1}(-\varrho_i) = p_i$  and  $P(y_i = 0 | x_i = 0) = 1 - f^{-1}(-\varrho_i) = p_i$  as required.

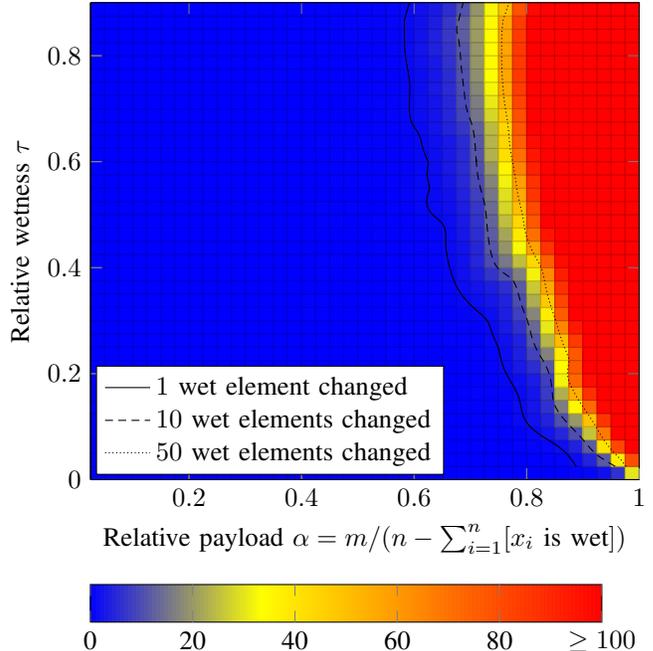
### 3.1. Practical coding algorithms

For a binary embedding operation, both types of senders can be realized in practice using syndrome coding when the message  $\mathbf{m}$  is communicated as a syndrome of a linear code  $\mathbf{m} = \mathbb{H}\mathcal{P}(\mathbf{y})$  with parity-check matrix  $\mathbb{H} \in \{0, 1\}^{m \times n}$ , where  $\mathcal{P} : \mathcal{X} \rightarrow \{0, 1\}$  is a parity function shared between the sender and the receiver, e.g.,  $\mathcal{P}(x) = x \bmod 2$ . The PLS problem (2) then becomes

$$\mathbf{y} = \arg \min_{\mathbb{H}\mathcal{P}(\mathbf{y})=\mathbf{m}} D(\mathbf{y}). \quad (8)$$

With  $m/n = \text{const.}$ , this construction is asymptotically (w.r.t.  $n$ ) optimal with high probability when the elements of  $\mathbb{H}$  are chosen randomly. Such codes are however highly impractical due to the exponential complexity of solving (8).

<sup>2</sup>A common practice in steganography is to define  $\rho_i(x_i) = 0$  which ensures  $D(\mathbf{x}) = 0$ . We do not require this here and thus  $\varrho_i$  can be arbitrary.



**Fig. 1.** Average number of wet elements out of  $n = 10^6$  that need to be changed to find a solution to (8) using STCs with  $h = 11$  versus relative payload  $\alpha$ .

A practical algorithm for solving (2) based on syndrome-trellis codes (STCs) was proposed in [8]. It uses a pseudo-randomly constructed, banded matrix  $\mathbb{H}$  with a band of height  $h$ . For such matrices, the Viterbi algorithm finds the optimal solution to (8) with complexity exponential w.r.t. the constraint height  $h$ . For small values of  $h \in \{7, \dots, 13\}$ , STCs achieve a coding loss between 5% to 10% for various values of  $\{\varrho_i\}_{i=1}^n$  and arbitrary relative payloads  $\alpha \in [0, 1]$  [7]. Since the whole cover image can be processed at once, wet pixels can be handled as well. Although the algorithm described in [8] uses the costs  $\rho_i(x_i) = 0$ ,  $\rho_i(y_i) \geq 0$ , it in fact works without a modification with arbitrary costs  $\varrho_i > -\infty$ .

STCs can also be used for solving the distortion-limited sender (3) for given costs  $\{\varrho_i | i = 1, \dots, n\}$  and bound  $D_\epsilon$ . An optimal algorithm for this problem would send  $m_{\text{MAX}} = H(\pi)$  bits on average.<sup>3</sup> Although the number of communicated bits is a random variable in this problem (recall that now  $D_\epsilon$  is fixed), we fix the number of bits to  $m = m_{\text{MAX}}(1 - l')$  and let the Viterbi algorithm find the optimal solution as in problem (2). The parameter  $l'$  is the coding loss we expect the algorithm will achieve and is determined experimentally for a given constraint height  $h$ .

STCs were originally designed to handle payloads  $\alpha \leq 1/2$ . This assumption can be relaxed depending on the amount of wet pixels. Since wet pixels are not allowed to be changed, the maximum number of bits we can communicate

<sup>3</sup>The value of  $m_{\text{MAX}}$  can be found by a binary search for  $\lambda$  satisfying (3).

is  $n - \sum_{i=1}^n [x_i \text{ is wet}]$  (we define  $\tau = \sum_{i=1}^n [x_i \text{ is wet}]/n$  as the relative wetness). For this reason, the relative payload is commonly defined as  $\alpha = m/(n - \sum_{i=1}^n [x_i \text{ is wet}]) \in [0, 1]$ . When both  $\tau$  and  $\alpha$  are large, the Viterbi algorithm may need to change some wet elements due to the banded structure of  $\mathbb{H}$ . This may be acceptable if this number is small, say 5 out of  $10^6$ . Figure 1 shows the average number of wet elements out of  $n = 10^6$  required to be changed in order to solve (8) for STCs with  $h = 11$ . The exact value of  $\varrho_i$  is irrelevant in this experiment as long as it is finite. This experiment suggests that STCs can be used with arbitrary  $\tau$  as long as  $\alpha \leq 0.7$ .

#### 4. MULTI-LAYERED CONSTRUCTION

In this section, we introduce a multi-layered construction which has been largely motivated by [22] and can be considered as a generalization of this work. The main idea is to decompose the problems (2) and (3) with a non-binary embedding operation into a sequence of similar problems for a binary embedding operation and then use the results of Section 3.

Let  $|\mathcal{I}_i| = 2^L$  for some integer  $L \geq 0$  and let  $\mathcal{P}_1, \dots, \mathcal{P}_L$  be parity functions uniquely describing all  $2^L$  elements in  $\mathcal{I}_i$ , i.e.,  $(x_i \neq y_i) \Rightarrow \exists j, \mathcal{P}_j(x_i) \neq \mathcal{P}_j(y_i)$  for all  $x_i, y_i \in \mathcal{I}_i$  and all  $i \in \{1, \dots, n\}$ . For example,  $\mathcal{P}_j(x)$  can be defined as the  $j$ th LSB of  $x$ . The individual sets  $\mathcal{I}_i$  can be enlarged to satisfy the size constraint by setting the costs of added elements to  $\infty$ .

The optimal algorithm for (2) and (3) sends the stego symbols by sampling from the optimal distribution (4) with some  $\lambda$ . Let  $\mathbf{Y}_i$  be the random variable defined over  $\mathcal{I}_i$  representing the  $i$ th stego symbol. Due to the assigned parities,  $\mathbf{Y}_i$  can be represented as  $\mathbf{Y}_i = (Y_i^1, \dots, Y_i^L)$  with  $Y_i^j$  corresponding to the  $j$ th parity function. We construct the embedding algorithm by induction over  $L$ , the number of layers. By the chain rule, for each  $i$  the entropy  $H(\mathbf{Y}_i)$  can be decomposed into

$$H(\mathbf{Y}_i) = H(Y_i^1) + H(Y_i^2, \dots, Y_i^L | Y_i^1). \quad (9)$$

This tells us that  $H(Y_i^1)$  bits should be embedded by changing the first parity of the  $i$ th pixel. In fact, the parities should be distributed according to the marginal distribution  $P(Y_i^1)$ . Using the flipping lemma, this task is equivalent to a PLS, which can be realized in practice using STCs as reviewed in Section 3.1. To summarize, in the first step we embed  $m_1 = \sum_{i=1}^n H(Y_i^1)$  bits on average.

After the first layer is embedded, we obtain the parities  $\mathcal{P}_1(y_i)$  for all stego pixels. This allows us to calculate the conditional probability  $P(Y_i^2, \dots, Y_i^L | Y_i^1 = \mathcal{P}_1(y_i))$  and use the chain rule again, for example w.r.t  $Y_i^2$ . In the second layer, we embed  $m_2 = \sum_{i=1}^n H(Y_i^2 | Y_i^1 = \mathcal{P}_1(y_i))$  bits on average. In total, we have  $L$  such steps fixing one parity value at a time knowing the result of the previous parities. Finally, we send the values  $y_i$  corresponding to the obtained parities.

If all individual layers are implemented optimally, we send  $m = m_1 + \dots + m_L$  bits on average. By the chain rule, this is exactly  $H(\mathbf{Y}_i)$  bits in every pixel, which proves the optimality of this construction. In theory, the order in which the parities are being fixed can be arbitrary. As is shown in the example below, the order is important for practical realizations using STCs. In all our experiments, we start with the *most* significant bits (MSBs) and end with the LSBs.

In practice, the number of bits hidden in every layer,  $m_j$ , must be communicated to the receiver. The number  $m_j$  is used as a seed for a pseudo-random permutation of all bits in the  $j$ th layer. If, due to large payload and wetness, STCs cannot embed a given message, we try a different permutation by embedding a slightly different number of bits.

*Example ( $\pm 1$  embedding):* For simplicity, let  $x_i = 2$ ,  $\mathcal{I}_i = \{1, 2, 3\}$ ,  $\rho_i(1) = \rho_i(3) = 1$ , and  $\rho_i(2) = 0$  for  $i \in \{1, \dots, n\}$  and large  $n$ . For such ternary embedding, we use two LSBs as their parities. Suppose we want to solve the problem (2) with  $\alpha = 0.9217$ , which leads to  $\lambda = 2.08$ ,  $P(Y_i = 1) = P(Y_i = 3) = 0.1$ , and  $P(Y_i = 2) = 0.8$ . To make  $|\mathcal{I}_i|$  a power of two, we also include the symbol 0 and define  $\rho_i(0) = \infty$  which implies  $P(Y_i = 0) = 0$ . Let  $y_i = (y_i^2, y_i^1)$  be a binary representation of  $y_i \in \{0, \dots, 3\}$ , where  $y_i^1$  is the LSB of  $y_i$ .

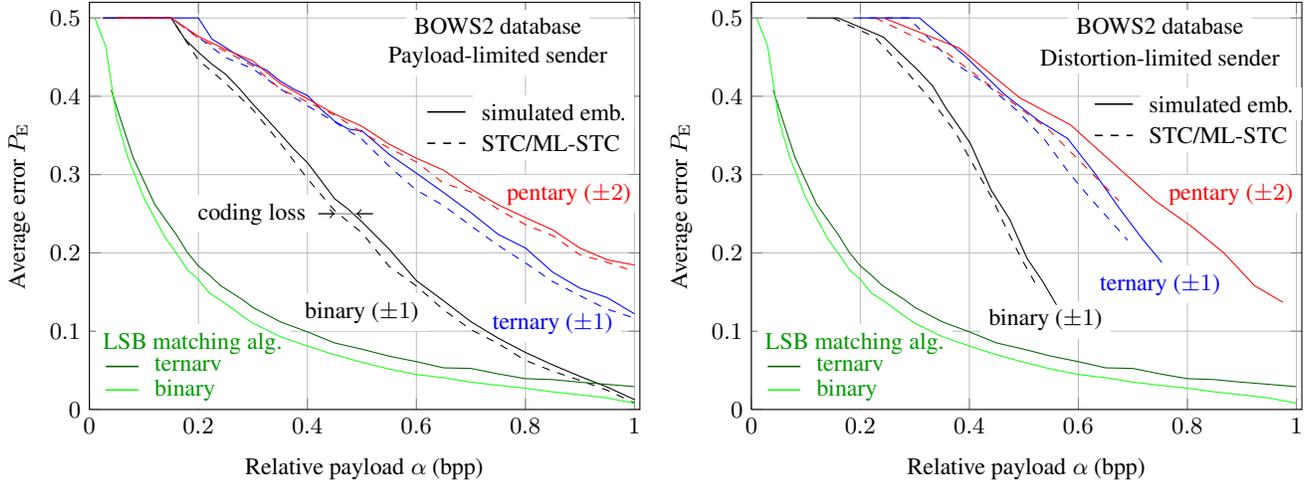
Starting from the LSBs as in [22], we obtain  $P(Y_i^1 = 0) = 0.8$ . If the LSB needs to be changed, then  $P(Y_i^2 = 0 | Y_i^1 = 1) = 0.5$  whereas  $P(Y_i^2 = 0 | Y_i^1 = 0) = 0$ . In practice, the first layer can be realized by any syndrome coding scheme minimizing the number of changes and embedding  $m_1 = n \cdot h(0.2)$  bits. The second layer can be implemented with wet paper codes [12], since we need to either embed one bit or leave the pixel unchanged (relative payload is 1).

If the weights of symbols 1 and 3 were slightly changed, however, we would have to use STCs in the second layer, which causes a problem due to the large relative payload ( $\alpha = 1$ ) combined with large wetness ( $\tau = 0.8$ ) (see Figure 1). The opposite decomposition starting with the MSB  $y_i^2$  will reveal that  $P(Y_i^2 = 0) = 0.1$ ,  $P(Y_i^1 = 0 | Y_i^2 = 0) = 0$ , and  $P(Y_i^1 = 0 | Y_i^2 = 1) = 0.8/0.9$ . Both layers can now be easily implemented by STCs since here the wetness is not as severe ( $\tau = 0.1$ ).

#### 5. PRACTICAL EMBEDDING CONSTRUCTION

We have implemented the above multi-layered construction based on STCs and present here a practical embedding scheme that was largely motivated by [16] and [6], which contain the justification and motivation of the design elements that appear below.

Let  $\mathbf{x} \in \{0, \dots, 255\}^{n_1 \times n_2}$  be an  $n_1 \times n_2$  grayscale cover image,  $n = n_1 n_2$ , represented in the spatial domain. Define the co-occurrence matrix computed from horizontal pixel differences  $D_{i,j}^{\rightarrow}(\mathbf{x}) = x_{i,j+1} - x_{i,j}$ ,  $i = 1, \dots, n_1$ ,



**Fig. 2.** Comparison of LSB matching with optimal binary and ternary coding with embedding algorithms based on the additive distortion measure (10) using embedding operations of three different cardinalities.

$j = 1, \dots, n_2 - 1$ :

$$A_{p,q,r}^{\rightarrow}(\mathbf{x}) = \frac{\sum_{i=1}^{n_1} \sum_{j=1}^{n_2-3} [(D_{i,j}^{\rightarrow}, D_{i,j+1}^{\rightarrow}, D_{i,j+2}^{\rightarrow})(\mathbf{x}) = (p, q, r)]}{n_1(n_2 - 3)},$$

where  $[(D_{i,j}^{\rightarrow}, D_{i,j+1}^{\rightarrow}, D_{i,j+2}^{\rightarrow})(\mathbf{x}) = (p, q, r)] = [(D_{i,j}^{\rightarrow}(\mathbf{x}) = p) \& (D_{i,j+1}^{\rightarrow}(\mathbf{x}) = q) \& (D_{i,j+2}^{\rightarrow}(\mathbf{x}) = r)]$ . Clearly,  $A_{p,q,r}^{\rightarrow}(\mathbf{x}) \in [0, 1]$  is the normalized count of neighboring quadruples of pixels  $\{x_{i,j}, x_{i,j+1}, x_{i,j+2}, x_{i,j+3}\}$  with differences  $x_{i,j+1} - x_{i,j} = p$ ,  $x_{i,j+2} - x_{i,j+1} = q$ , and  $x_{i,j+3} - x_{i,j+2} = r$  in the entire image. The superscript arrow “ $\rightarrow$ ” denotes the fact that the differences are computed by subtracting the left pixel from the right one. Similarly, we define matrices  $A_{p,q,r}^{\leftarrow}(\mathbf{x})$ ,  $A_{p,q,r}^{\uparrow}(\mathbf{x})$ , and  $A_{p,q,r}^{\searrow}(\mathbf{x})$ . Let  $y_{i,j} \mathbf{x}_{\sim i,j}$  be an image obtained from  $\mathbf{x}$  by replacing the  $(i, j)$ th pixel with value  $y_{i,j}$ . Finally, we define the distortion measure  $D(\mathbf{y}) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \rho_{i,j}(y_{i,j})$  as

$$\rho_{i,j}(y_{i,j}) = \sum_{\substack{p,q,r \in \{-255, \dots, 255\} \\ s \in \{\rightarrow, \leftarrow, \uparrow, \searrow\}}} w_{p,q,r} |A_{p,q,r}^s(\mathbf{x}) - A_{p,q,r}^s(y_{i,j} \mathbf{x}_{\sim i,j})|, \quad (10)$$

where  $w_{p,q,r} = 1/(1 + \sqrt{p^2 + q^2 + r^2})$  are heuristically chosen weights.

All tests were carried out on the BOWS2 database [1] containing approximately 10800 grayscale images with a fixed size of  $512 \times 512$  pixels coming from rescaled and cropped natural images of various sizes. Steganalysis was implemented using the second-order SPAM feature set with  $T = 3$  [15]. The image database was evenly divided into a training and a testing set of cover and stego images, respectively. A soft-margin support-vector machine was trained using the Gaussian kernel. The kernel width and the penalty parameter were determined using five-fold cross validation on the grid  $(C, \gamma) \in \{(10^k, 2^{j-d}) | k \in \{-3, \dots, 4\}, j \in \{-3, \dots, 3\}\}$ ,

where  $d$  is the binary logarithm of the number of features. We report the results using a measure frequently used in steganalysis – the minimum average classification error  $P_E = (P_{FA} + P_{MD})/2$ , where  $P_{FA}$  and  $P_{MD}$  are the false-alarm and missed-detection probabilities.

Figure 2 contains the comparison of embedding algorithms implementing the PLS and DLS with the costs (10). All algorithms are contrasted with LSB matching simulated on the binary and ternary bounds. To compare the effect of practical codes, we first simulated the embedding algorithm as if the best codes were available and then compared these results with algorithms implemented using STCs with  $h = 10$ . Both types of senders are implemented with binary, ternary ( $\mathcal{I}_i = \{x_i - 1, \dots, x_i + 1\}$ ), and pentary ( $\mathcal{I}_i = \{x_i - 2, \dots, x_i + 2\}$ ) embedding operations. Before embedding, the binary embedding operation was initialized to  $\mathcal{I}_i = \{x_i, y_i\}$  with  $y_i$  randomly chosen from  $\{x_i - 1, x_i + 1\}$ . The reported payload for the DLS with a fixed  $D_\epsilon$  was calculated as an average over the whole database after embedding.

The relative horizontal distance between the corresponding dashed and solid lines in Figure 2 is bounded by the coding loss. Most of the proposed algorithms are undetectable for relative payloads  $\alpha \leq 0.2$  bits per pixel (bpp). For payloads  $\alpha \leq 0.5$ , the DLS is more secure. For larger payloads, the distortion measure seems to fail to capture the statistical detectability correctly and thus the algorithms are more detectable than when implemented in the payload-limited regime. These results suggests that larger embedding changes are useful for steganography when placed adaptively.

## 6. CONCLUSION

This paper describes near-optimal codes for minimal-distortion steganography implemented in the following manner. Before

embedding in a given cover, the sender first specifies for each cover element  $x_i$  the set of values  $y_i \in \mathcal{I}_i$  to which  $x_i$  can change and the associated cost of making this modification,  $\rho_i(y_i)$ . The problem is to communicate a payload of a certain size with minimal expected embedding distortion obtained as a sum of individual pixel costs  $\rho_i$  (alternatively, embed the largest possible payload for a given bound on the expected distortion). The proposed approach works for an arbitrary cost assignment and arbitrary sets  $\mathcal{I}_i$ , which can even be different for every  $i$ . The method is a generalization of previously proposed syndrome-trellis codes and other special cases, including matrix embedding and wet paper codes.

The merit of the proposed method is demonstrated experimentally by implementing it for binary, ternary, and pentary embedding operations in spatial domain and showing an improvement in statistical detectability measured by a blind steganalyzer. This construction is not limited to embedding with larger amplitudes but can be used, e.g., for embedding in color images, where the LSBs of all three colors can be seen as 3-bit symbols on which the cost functions are defined. Applications outside the scope of digital images are possible as long as the costs can be meaningfully defined.

Matlab and C++ implementation of multi-layered STCs is available at <http://dde.binghamton.edu/download/syndrome/>.

## 7. REFERENCES

- [1] P. Bas and T. Furon. BOWS-2. <http://bows2.gipsa-lab.inpg.fr>, July 2007.
- [2] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. New York: John Wiley & Sons, Inc., 2006.
- [3] R. Crandall. Some notes on steganography. *Steganography Mailing List*, available from <http://os.inf.tu-dresden.de/westfeld/crandall.pdf>, 1998.
- [4] T. Filler and J. Fridrich. Wet ZZW construction for steganography. In *First IEEE Intern. Workshop on Information Forensics and Security*, London, UK, Dec. 2009.
- [5] *IEEE Trans. on Information Forensics and Security*.
- [6] T. Filler and J. Fridrich. Gibbs construction in steganography. [5], 2010. To appear in December issue.
- [7] T. Filler, J. Judas, and J. Fridrich. Minimizing additive distortion in steganography using syndrome-trellis codes. [5], 2010. Submitted.
- [8] T. Filler, J. Judas, and J. Fridrich. Minimizing embedding impact in steganography using trellis-coded quantization. In *Proceedings SPIE, Electronic Imaging*, volume 7541, pages 05–01–05–14, Jan. 17–21, 2010.
- [9] J. Fridrich. *Steganography in Digital Media: Principles, Algorithms, and Applications*. Cambridge University Press, 2009.
- [10] J. Fridrich and T. Filler. Practical methods for minimizing embedding impact in steganography. *Proceedings SPIE, Electronic Imaging*, volume 6505, pages 02–03, San Jose, CA, Jan. 29–Feb. 1, 2007.
- [11] J. Fridrich, M. Goljan, and D. Soukal. Wet paper codes with improved embedding efficiency. [5], 1(1):102–110, 2006.
- [12] J. Fridrich, M. Goljan, D. Soukal, and P. Lisoněk. Writing on wet paper. In *IEEE Transactions on Signal Processing, Special Issue on Media Security*, volume 53, pages 3923–3935, Oct. 2005. (journal version).
- [13] J. Fridrich and D. Soukal. Matrix embedding for large payloads. [5], 1(3):390–394, 2006.
- [14] Y. Kim, Z. Duric, and D. Richards. Modified matrix encoding technique for minimal distortion steganography. In *Information Hiding, 8th Intern. Workshop*, volume 4437 of Lecture Notes in Computer Science (LNCS), pages 314–327, Alexandria, VA, Jul. 10–12, 2006.
- [15] T. Pevný, P. Bas, and J. Fridrich. Steganalysis by subtractive pixel adjacency matrix. In *Proceedings of the 11th ACM Multimedia & Security Workshop*, pages 75–84, Princeton, NJ, Sep. 7–8, 2009.
- [16] T. Pevný, T. Filler, and P. Bas. Using high-dimensional image models to perform highly undetectable steganography. In *Inform. Hiding, 12th Intern. Conference*, LNCS, Calgary, Alberta, Canada, Jun. 28–30 2010.
- [17] D. Schönfeld and A. Winkler. Embedding with syndrome coding based on BCH codes. In *Proceedings of the 8th ACM Multimedia & Security Workshop*, pages 214–223, Geneva, Switzerland, Sep. 26–27, 2006.
- [18] C. E. Shannon. Coding theorems for a discrete source with a fidelity criterion. *IRE Nat. Conv. Rec.*, 4:142–163, 1959.
- [19] A. Westfeld. High capacity despite better steganalysis (F5 – a steganographic algorithm). In *Information Hiding, 4th International Workshop*, volume 2137 of LNCS, pages 289–302, Pittsburgh, PA, Apr. 25–27, 2001.
- [20] R. Zhang, V. Sachnev, and H. J. Kim. Fast BCH syndrome coding for steganography. In *Information Hiding, 11th International Workshop*, volume 5806 of LNCS, pages 31–47, Darmstadt, Germany, Jun. 7–10, 2009.
- [21] W. Zhang and X. Wang. Generalization of the ZZW embedding construction for steganography. [5], 4(3):564–569, Sep. 2009.
- [22] X. Zhang, W. Zhang, and S. Wang. Efficient double-layered steganographic embedding. *Electronics Letters*, 43:482–483, Apr. 2007.