

An Image Steganography Based on Multi-Layered Syndrome-Trellis Codes in DWT Domain

YANG Xiaoyuan^{1,2}, GUO Duntao², LI Jun^{1,2}

1. Key Laboratory of Network & Information Security under the Chinese Armed Police Force, Electronic Department, Engineering College of the Armed Police Force, Xi'an 710086, China
2. Institute of Network & Information Security under the Chinese Armed Police Force, Department of Electronic Technology, Engineering University of Armed Police Force, Xi'an 710086, China
E-mail: lijun9250lj@163.com

Abstract: A good method to improve the security of Steganography System is to embed messages by minimizing a sum of per-pixel distortions and the near-optimal codes for this minimization problem is the Syndrome-Trellis Codes (STCs). This paper will use the non-binary embedding operations with Multi-Layered Syndrome-Trellis Codes (ML-STCs) to design image steganography method instead of simply binary embedding operation with STCs. First, a distortion function in discrete wavelet transform domain was designed according to human visual system and integer lifting wavelet transform, which mainly concerned about the influences of frequency, luminance and texture masking factor to the cover distortion. And then, the Multi-Layered Syndrome-Trellis Codes was combined with the distortion function to propose the steganography method, so make sure that the embedding impact to the cover minimizing and centralize in un-sensitivity domain. Experiment results show that the proposed approach maintains a good visual quality of stego-image and have a high security against steganalysis in space and wavelet domain.

Key Words: Image Steganography; Multi-Layered Syndrome-Trellis Codes (ML-STCs); Discrete Wavelet Transform (DWT); Non-binary Embedding

1 Introduction

In steganography, the sender communicates with the receiver by hiding their messages in trusted media, Such as digital images, so that it is hard to distinguish between the cover image and the stego image. The message is typically hidden (embedded) in the cover image by slightly modifying the systems of secret communication; the senders embed the secret message into the cover so that it can achieve the purpose of security communication with the receivers. When the steganography algorithm modify the cover less, their safety is relatively higher, which is the problem of so called minimizing embedding impact [1]. Filler proposed the Syndrome trellis codes (STCs) [2-4] which is a practical near-optimal embedding method that can realize minimizing embedding impact embedding [5]. They proposed two distortion functions in spatial and discrete cosine transform domain. In this paper, we will study the important problem of selecting the embedding distortion at each cover element that minimizing distortion corresponds to minimizing statistical detectability and we will propose a new distortion function in the discrete wavelet transform domain and give a new steganography method combined with the ML-STCs (Multi-Layered Syndrome-Trellis Codes). The ML-STCs is proposed in paper [4], which is no longer only for embedding operations with binary, but also with ternary, pentary and so on.

2 Minimizing embedding impact

Take digital image cover for example, the cover and stego image can be expressed as $X = (x_1, \dots, x_n)$,

$Y = (y_1, \dots, y_n) \in \{0, 1\}^n$ respectively, the cost by the process of embedding is defined by follow:

$$D(X, Y) = \|X - Y\|_\rho = \sum_{i=1}^n \rho_i |x_i - y_i| \quad (1)$$

Where $0 \leq \rho_i \leq \infty$ is the cost of changing pixel x_i to y_i .

When $\rho_i = \infty$, it means that this pixel is the so called wet pixel which is not allowed to change^[6]. For the additive distortion function ρ in (1), paper [1] has the following separation principle:

Let $\rho = (\rho_i)_{i=1}^n$, $0 \leq \rho_i \leq \infty$ is the set of constants defining the additive distortion measure (1) for $i \in \{1, \dots, n\}$.

Let $0 \leq m \leq n$ be the number of bits we want to communicate by using a binary embedding operation. The minimal expected distortion has the following form:

$$D_{\min} = (m, n, \rho) = \sum_{i=1}^n p_i \rho_i \quad p_i = \frac{e^{-\lambda \rho_i}}{1 + e^{-\lambda \rho_i}} \quad (2)$$

Where p_i is the probability of changing the i th pixel. The parameter λ is obtained by solving following embedding capacity equation:

$$-\sum_{i=1}^n (p_i \log_2 p_i + (1 - p_i) \log_2 (1 - p_i)) = m \quad (3)$$

* This work is supported by Natural Science Basic Research Plan in Shaanxi Province of China (Program No. 2012JM8014), Foundation Fund of Engineering College of the Armed Police Force (Program No. wjy201123)

The importance of above principle is in the separation of the image model (needed for calculating ρ_i) and the coding algorithm used in a practical implementation. By virtue of this separation, better steganographic algorithms can be derived by using better coding or by using a better image model. One important consequence is that, in order to study the effect of the image model on steganographic security, no coding algorithm is needed at all. The optimal coding can be simulated by flipping each pixel with probability p_i as defined in (2), which is the so-called simulated optimal embedding method [5]. Based on the above considerations, design a reasonable measure of distortion ρ is very important, so that we will give a new ρ in the wavelet domain based on Human Visual System.

3 Multi-Layered Syndrome-trellis codes (ML-STCs)

The simulated optimal embedding method [5] mentioned above is based on the assumption that the probability of change i th pixel obey to (2), but in practice it is very difficult to obtain. In 2010, Filler, Judas, and Fridrich proposed the practical near-optimal embedding method syndrome-trellis codes [2~4]. So the designer of steganography can propose practical algorithms which are minimizing embedding impact with the use of STCs.

3.1 The Construction of STCs

The parity-check matrix of a binary syndrome-trellis code of length n and codimension m is obtained by placing a small sub matrix along the main diagonal as in Figure 1. The height h of the sub matrix (called the *constraint height*) is a design parameter that affects the algorithm speed and efficiency (typically, $6 \leq h \leq 15$). The width ω is rely on embedding capacity α : $1/(\omega+1) \leq \alpha \leq 1/\omega$.

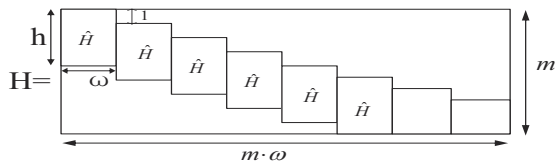


Fig. 1: Struction of H and $\hat{H}_{h \times \omega}$

3.2 The Details of STCs

Every codeword of STCs can be represented as a unique path through a graph called the syndrome trellis. An example of the syndrome trellis is shown in Figure 2, in which (a) is the H and $\hat{H}_{h \times \omega}$. More formally, the syndrome trellis is a graph consisting of b blocks (b is the number of \hat{H}), each containing $2^h \times (\omega+1)$ nodes. The nodes between two adjacent columns form a bipartite graph, i.e., all edges only connect nodes from two adjacent columns. Each block of the trellis represents one sub matrix.

Each $z \in \{0,1\}^n$ satisfying $H z = m$ represented as a path through the syndrome trellis and one small block can embed one bit message. Each path starts in the leftmost all-zero state in the trellis and extends to the right. The path

shows the step-by-step calculation of the syndrome using more and more bits of z . For example, the first two edges in Figure 2 (b), that connect the state 00 from column p_0 with states 11 and 00 in the next column, correspond to adding or not adding the first column of H to the syndrome, respectively. At the end of the first block, we terminate all paths, for which the first bit of the partial syndrome does not match m_1 . This way, we obtain a new column of the trellis, which will serve as the starting column of the next block. This operation is repeated at each block transition in the matrix H and guarantees that 2^h states are sufficient to represent the calculation of the partial syndrome throughout the whole syndrome trellis.

Each edge is assigned a weight that is the so-called embedding loss which respects the cost of modifying. The states of nodes in columns labeled l , $l \in \{1, \dots, n\}$, stand for the partial syndromes that can be obtained using the first l bits of the vector y . Fortunately, it is not necessary to enumerate all 2^l possible partial syndromes in each column. Instead, because each column of H has at most h nonzero values, we need to focus only on those h bits of the syndrome that can actually be changed by adding or not adding the l th column of H to the partial syndrome. Thus, each column in the trellis has only 2^h nodes.

Even though the number of paths through the trellis is exponential in n , the problem of finding the shortest Path can be efficiently solved by a form of dynamic programming called the Viterbi algorithm. This algorithm consists of two parts, the *forward* and the *backward* part. The *forward* part is from left to right, which finds the shortest codes. The black **bold** path in Figure 2(b) is just the shortest path when the cover X is 10110001 and the secret message m is 0111, whose total cost is 2; the *backward* part is get the stego Y according to the shortest path, in the above example, the stego Y is 00111001.

$$\hat{H} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \quad H = \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 \end{pmatrix}$$

(a) $\hat{H}_{h \times \omega}$ and H

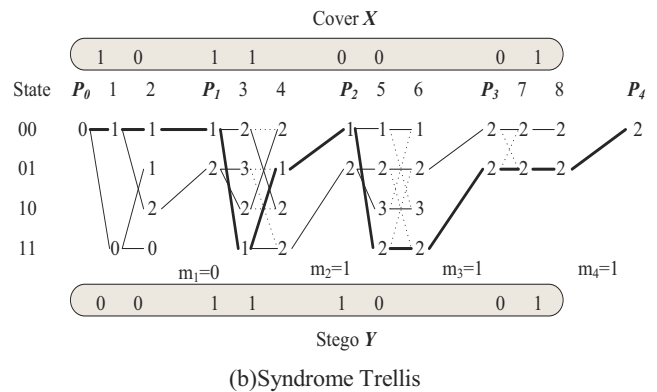


Fig. 2 An example of STCs

In This paper, we will set each carrier a reasonable pixel distortion value, so that we can get a stego Y which is closest to the cover X with the use of STCs.

3.3 Multi-Layered construction of Syndrome-trellis codes

In paper [4], Filler has introduced a multi-layered construction of syndrome-trellis codes which has been largely motivated by paper [13]. The main idea of ML-STCs is to decompose the problem of non-binary embedding operation into a sequence of similar problems for a binary embedding operation and then use the method of STCs. The details of multi-layered construction are show in paper [4] and [5].

4 Distortion function based on human visual system and integer lifting wavelet

For digital images, the following distortion functions are common in steganography: the *constant distortion function*, $\rho(x)=1$, all pixels have the same impact on detectability when changed, such as F5; the *linear distortion function*, $\rho(x)=2x$, when the distortion is related to a quantization error uniformly distributed on $[-Q/2, Q/2]$ for some quantization step $Q > 0$, such as QIM. Both of them are the basic distortion function, whose practice implication effects are not good. In paper [4] and [6], Filler proposed two distortion functions in spatial domain and DCT domain respectively, and also a high security steganographic algorithm called HuGO [8] based on spatial domain distortion function which is using co-occurrence matrix [8].

4.1 The Relationship between Distortion and JND Model

Based on human visual system (HVS) and integer lifting wavelet transform, this section we will propose a new distortion function ρ . HVS features include frequency sensitivity, luminance sensitivity, contrast masking, marginal characteristics, and texture masking, etc. These features play an important role in improving the steganography system's visual security, robustness and even statistical security. JND (Just Noticeable Difference) model is the max distortion derives from embedding that human eyes can't distinguish. As JND model depend on visual conditions and the content of the image itself, and the distortion function ρ represents the impact over the cover by the modification of a pixel, it is clear that ρ has inverse relationship with JND. Keeping this in mind, this section will give a distortion function in the base of JND Model. Lewis made use of the human visual system in the wavelet domain of images to compress redundant information [9], which established a wavelet JND threshold model. In the base of Lewis, Barni improved the Luminance masking and texture masking factor in JND model [10]. The JND model in wavelet domain mentioned in paper [13] and [14] is:

$$JND = A(r, i, j) * \Theta(r, \theta) * \Xi(r, i, j)^\delta \quad (4)$$

In this paper, we will improve the luminance masking factor $\Theta(r, \theta)$, the frequency masking factor $\Lambda(r, i, j)$ and the texture masking factor $\Xi(r, i, j)$ respectively. We define the distortion function as follow based on the new JND model:

$$\rho = \eta * \frac{1}{JND} = \eta * \frac{1}{A(r, i, j) * \Theta(r, \theta) * \Xi(r, i, j)^\delta} \quad (5)$$

Let the cover image divided into 16×16 nonoverlapping blocks, we can get coefficients I_r^θ from decomposing the blocks with 3 layers integer lifting wavelet. $\theta = \{0, 1, 2, 3\}$ means horizontal, vertical, diagonal details coefficients and approximation coefficients respectively. $r = \{1, 2, 3\}$ is the layer of coefficients.

4.2 Luminance masking factor

As the wavelet transforms itself in line with the HVS, the low-frequency approximation coefficients represent the brightness of small blocks. Lewis's model [9] argue that the lower the brightness, more sensitive the human eyes, which is inconsistent with the HVS. Barni's model [10] has pointed out that human eyes are less sensitive in low and high brightness area, and relative more sensitive in the middle brightness area, but it didn't notice that the very bright and very dark areas are differences between the sensitivity. In fact, the brightness sensitivity of the human eye is a non-linear model; very bright areas are more sensitive than very dark areas. Considering computational complexity, we use the following linear approximation function instead of non-linear luminance masking factors:

$$A(r, i, j) = \begin{cases} -2L(r, i, j) + 2 & L(r, i, j) \leq \frac{1}{2} \\ L(r, i, j) + \frac{1}{2} & L(r, i, j) > \frac{1}{2} \end{cases} \quad (6)$$

Where $L(r, i, j) = \frac{1}{2^n} \times I_3^3(r, i, j) = \frac{1}{2^n} \times I_3^3\left(\left\lceil \frac{i}{2^{3-r}} \right\rceil, \left\lceil \frac{j}{2^{3-r}} \right\rceil\right)$ are the values of wavelet coefficients which are normalized. $I_3^3(r, i, j)$ is the low-frequency approximation coefficients of 3rd layer. As we have no exactly accurate model for contrasting, we give the contrast of the provided model and the approximate model in figure 3, the difference of two is may be the deviation.

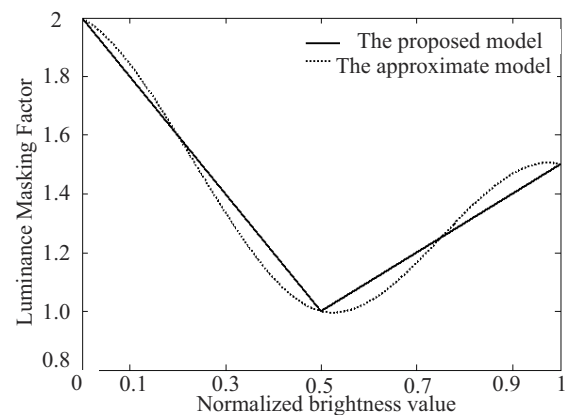


Fig. 3 The relationship between luminance masking factors and luminance value

4.3 Frequency masking factor

Human eyes have different sensitivities with different coefficients: the high-frequency diagonal coefficients have lowest sensitive; the low-frequency coefficients have the

highest sensitive, horizontal and vertical high-frequency coefficients are in the middle level. Paper [9] and [10] proposed a frequency masking factor. According to the experiments, as just only use three layers embedding, in this paper we let the frequency masking factor $\Theta(r, \theta)$ as Table 1 below:

Table 1: Frequency masking factor $\Theta(r, \theta)$

Direction θ layers r	horizontal 0	vertical 1	diagonal 2
1	1	1	$\sqrt{2}$
2	0.32	0.32	$\sqrt{2} \times 0.32$
3	0.16	0.16	$\sqrt{2} \times 0.16$

4.4 Texture masking factor

Texture masking is an important characteristic of HVS, human eyes is very sensitive to noises in smooth regions than strong texture regions. Edge information of the visual image is also important for JND model. Texture masking factor in paper [10] is:

$$\Xi(r, i, j) = \sum_{k=0}^{3-l} \frac{1}{16^{k+1}} \sum_{\theta=0}^2 \sum_{x=0}^1 \sum_{y=0}^1 \left[I_{k+l}^{\theta} \left(x + \left\lfloor \frac{i}{2^k} \right\rfloor, y + \left\lfloor \frac{j}{2^k} \right\rfloor \right) \right] \times \text{Var}\{I_3^3\} \quad (7)$$

It consists of two parts, first part is the local mean square value of the DWT coefficients in all detail sub bands, it represents the distance from the edges; while the second is the local variance of the low-pass sub bands, and it represents the texture. Both these parts are computed in a small 2×2 neighborhood corresponding to the location (i, j) of the pixel.

But after a test we found that the mean square value is not need at all. It is enough to represent the distance from the edges with the use of average absolute value of high-frequency coefficients instead of mean square value; an important consequence is that we can reduce the computational complexity evidently. In the other way, as the distortion values we want is a relatively relationship, it is reasonable that if the improved model is linear correlation with the original model. The relationship is show in Figure 4, it is clear that two model are perfectly correlated, so the improved model is reasonable. This improved texture masking factor is as below:

$$\Xi(r, i, j) = \sum_{k=0}^{3-l} \frac{1}{16^{k+1}} \sum_{\theta=0}^2 \sum_{x=0}^1 \sum_{y=0}^1 \left[\left| I_{k+l}^{\theta} \left(x + \left\lfloor \frac{i}{2^k} \right\rfloor, y + \left\lfloor \frac{j}{2^k} \right\rfloor \right) \right| \right] \times \text{Var}\{I_3^3\} \quad (8)$$

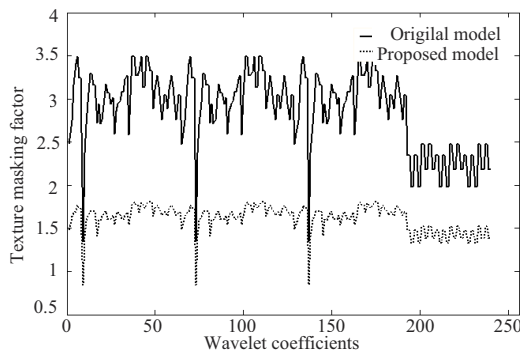


Fig. 4 The statistical model of two texture masking factors

5 Process of Embedding, Experiments and Analysis

5.1 The combination of proposed distortion function and ML-STCs

In section 2, we have gave a simple introduction of good underlying code ML-STCs, whose input parameters are Cover (such as pixel values, DCT coefficients), cost profiles ρ and binary secret message m , and output parameter is the Stego. The distortion between Stego and Cover will minimizing relying on distortion cost ρ . In section 3, we provided a distortion function based on HVS in wavelet domain that each DWT coefficients has a value on behalf of embedding impact. This section will give an image steganography based on ML-STCs and the proposed distortion function.

5.2 The process of Secret information embedding

1) Let the cover image with the size of $M \times N$ divided into 16×16 nonoverlapping blocks, each block is carried out using three layers integer lifting wavelet transform, and then get ρ_i according (5). It is noteworthy that, due to the cost by Modify the low-frequency approximation coefficients I_3^3 has greater impact on the image, in this scenario we will define it as wet pixels ^[6] ($\rho_i = \infty$), that means these coefficients are not allowed to embed secret bits.

2) In this steganography, we put the coefficients of DWT into a vector as the cover X , $X = \{x_1, \dots, x_n\}$, (length n is $\lfloor M/16 \rfloor \times \lfloor N/16 \rfloor \times 256$). We use the ternary embedding operations in ML-STCs, that is mean the cover x_i can be change to $y_i = \{x_i - 1, x_i, x_i + 1\}$ with the cost of $\{\rho_i, 0, \rho_i\}$ respectively, where ρ_i is come from (5). So we put the cover X , distortion values ρ and secret message m into the ML-STCs algorithm mentioned above, and we will get a vector with secret message which we called *stego* vector. *Stego* vector is still the coefficients values of DWT.

3) Using the *stego* vector replaces the original wavelet coefficient's values, and we can get an image with secret message according the inverse integer wavelet transform.

In the extraction of secret information, as long as the *stego* vector and the length of secret message m , ML-STCs algorithm can extract all message correctly without any information about distortion function ρ , integer lifting wavelet transform ensure the accuracy of the message extracted.

5.3 Experiment conditions

The images used for experiments are the normal Lena (size is 256×256 pixels, bit depth is 8) and UCID image database^[15], parameter η in (5) is 5, δ in JND model is 0.2, h in STCs is 10. The classifier used for steganalysis is support vector machine (SVM) with C-SVC, in which the kernel function is non-linear radial basis function (RBF). The chosen accuracy measure is the minimal average decision error under equal probability of cover and stego images, defined as

$P_E = (P_{F_p} + P_{F_n}) / 2$, where P_{F_p} and P_{F_n} stand for the probability of false alarm or false positive.

5.4 Experiment analysis

In Figure 5, (a) and (b) are the cover image and stego image about Lena, we can see that the stego image has good visual quality and human eyes can't notice the difference between cover image and stego image; (c) is the PSNR about stego image in which the PSNR values are more than 59 when the reciprocal embedding capacity change from 1/20 to 1/2.

As the proposed image steganography is first embedding secret message in DWT's coefficients and then transform back to spatial domain, in order to give a relatively complete steganalysis experiment we chose both spatial domain and wavelet domain feature for steganalysis. Prior to all experiments, the images were divided into two sets of equal size, one used exclusively for training, the other exclusively for evaluation of the accuracy. In following two experiment, we both give 1338 of images with 512×328 from UCID for experiments, one part for training and another for testing, the embedding capacity is from 0.1 to 0.9.



Fig.5 Imperceptibility of the proposed steganography method

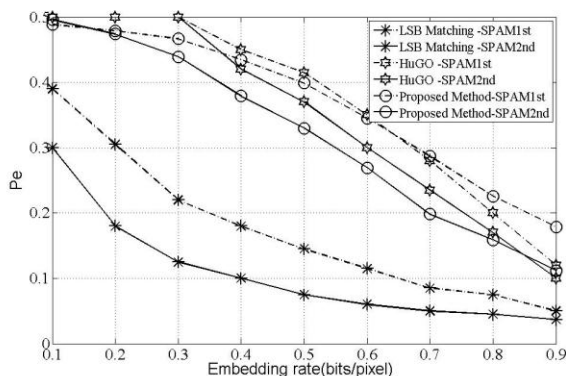


Fig.6 Security comparisons of 3 methods

Figure 6 is the experiments of security test with the feature of SPAM (Subtractive Pixel Adjacency Matrix) which is a good steganalysis feature for spatial and frequency domain in recent years^[7]. Steganalysis was implemented using the first-order SPAM feature set with $T=4$ and second-order SPAM feature set with $T=3$. As can be seen from Figure 6 that, if we regard 0.4 as the security capacity, the ordinary LSB matching's security embedding capacity is only 0.1bpp, but the proposed method's security embedding capacity is

almost 0.4bpp, in line with the current spatial safer steganography method namely HuGO^[8].

Table 2 is the steganalysis experiment result with the feature of Farid 216^[12] in wavelet domain. The feature is an effective feature for steganalysis in wavelet domain which is according to the image in the wavelet decomposition of high-level statistical regularities, as well as the linear correlation between the absolute value of wavelet coefficients to achieve the purpose of steganalysis. As can be seen from table 2, the security embedding capacity is about 0.4bpp which is a high level.

Table.2: Steganalysis against the proposed method using the feature of Farid 216

Embedding rate (bits/pixel)	P_{F_p}	P_{F_n}	P_E
0.1	0.499 8	0.519 8	0.5098
0.2	0.817 2	0.191 3	0.5042 5
0.3	0.621 8	0.344	0.4829
0.4	0.321 2	0.433 2	0.3772
0.5	0.240 7	0.429 1	0.3349
0.6	0.256 4	0.342 4	0.2994
0.7	0.233 3	0.243 1	0.2382
0.8	0.168 1	0.235 4	0.2017 5
0.9	0.103 2	0.212 1	0.1576 5

Accord to the result from above 2 experiments, the proposed method's security is relatively high.

6 Conclusion

This article has proposed an image steganography method in the discrete wavelet transform domain based on ML-STCs. The main idea of this article is the proposed of distortion function in DWT domain, and we are no longer use the simply STCs method any more but the ML-STCs. As the distortion function combined the HVS, wavelet transform and the Minimizing Embedding Impact, the method provides a good visual quality and statistical security. The ML-STCs can embed much more secret message than STCs. The next step we will research is improving this method with the use of non-additive distortion function proposed by Filler.

References

- [1] Fridrich J, Filler T. Practical methods for minimizing embedding impact in steganography [EB/OL]. <http://ws2.binghamton.edu/fridrich/publications.html#Steganography>, 2012-02-29.
- [2] Filler T, Judas J, Fridrich J. Minimizing Additive Distortion in Steganography using Syndrome-Trellis Codes [J]. IEEE Transactions on Information Forensics and Security, 2011, 6(3):920-935.
- [3] Filler T, Judas J, Fridrich J. Minimizing Embedding Impact in Steganography using Trellis-Coded Quantization. [EB/OL]. <http://ws2.binghamton.edu/fridrich/publications.html#Steganography>, 2012-02-29.
- [4] Filler T, Fridrich J. Minimizing Additive Distortion

- Functions With Non-Binary Embedding Operation in Steganography [C]//Proceedings of 2nd IEEE Workshop on Information Forensics and Security. Seattle, WA: IEEE Computer Society, 2010:1-6.
- [5] Filler T, Fridrich J. Gibbs construction in steganography [J]. IEEE Transactions on Information Forensics and Security, 2010, 5(4):705-720.
- [6] Fridrich J, Goljan M, Soukal D. Writing on wet paper [J]. IEEE Transactions on Signal Processing, 2005, 53(10):3923-3935.
- [7] Pevný T, Bas P, Fridrich J. Steganalysis by subtractive pixel adjacency matrix [J]. IEEE Transactions on Information Forensics and Security, 2010,5(2):215-224.
- [8] Pevný T, Filler T, Bas P. Using high-dimensional image models to perform highly undetectable steganography[C]//Proceedings of the 12th International Workshop on Information Hiding. Calgary, AB: Springer Verlag, 2010, 6387:161-177.
- [9] Lewis A S, Knowles G. Image compression using the 2-D wavelet transforms [J]. IEEE Transactions on Image Processing, 1992, 1(4): 244~250.
- [10] Barni M, Bartolini F, Piva A. Improved Wavelet-based Watermarking Through Pixel-wise Masking [J]. IEEE Transactions on Image Processing, 2001, 10(5): 783~791.
- [11] Schaefer G, Stich M. UCID—An uncompressed colour image database [R]. Nottingham, UK: School of Computing and Mathematics, Nottingham Trent University, 2003.
- [12] Farid H. Detecting Steganographic Messages in Digital Images [R]. TR2001-412, Hanover, NH, USA: Dartmouth College, 2001.
- [13] X. Zhang, W. Zhang, and S. Wang. Efficient double layered steganographic embedding. *Electronics Letters*, 43:482–483, Apr. 2007.