Host-Interference Rejecting Spread-Spectrum Watermarking: Implementation and Implications

Chun-Hsiang Huang
Communications and Multimedia Laboratory,
Department of CSIE,
National Taiwan University
bh@cmlab.csie.ntu.edu.tw

Jin-Sin Liu
Communications and Multimedia Laboratory,
Department of CSIE,
National Taiwan University,
jerrick@cmlab.csie.ntu.edu.tw

Ja-Ling Wu
Graduate Institute of Networking and Multimedia,
National Taiwan University,
wjl@cmlab.csie.ntu.edu.tw

ABSTRACT

Blind-detection is a desirable characteristic of watermarking schemes so that embedded payloads can be readily extracted without referencing original cover works. However, most blind-detection watermarking methods, like the famous spread-spectrum approach, suffer from inherent performance loss caused by host-interference. In this manuscript, an iterative informed-embedding spread-spectrum approach is proposed and exploited to alleviate the host-interference problems. Geometric models, detailed algorithms as well as implementation issues are provided to illustrate the effectiveness of our iterative methodology. Furthermore, inherent performance losses caused by the assumption of conventional spread-spectrum watermarking schemes - watermark signals shall be added to most perceptually significant components - is pointed out. Finally, problems and practical applications of iterative informed-embedding watermarking schemes are discussed.

Keywords: digital watermarking, blind detection, informed embedding, iterative methodology

1. INTRODUCTION

With advances in content compression technologies, cost reduction of playback device manufacturing, as well as explosive growth of wideband networking infrastructures, consumption of high-quality digital contents has become an indispensable part in our daily life. However, due to the proliferation of low-cost replica devices and convenient global networks, data piracy has seriously infringed upon the rights of content providers and related industries. In the past decade, digital watermarking schemes were proposed to provide copyright protection for decrypted contents. Watermarked contents shall possess unobtrusive perceptual fidelity and, at the same time, the hidden watermark message should be able to be extracted to identify the rights ownership even the contents have undergone reasonable manipulations. A comprehensive survey to important watermarking concepts and schemes can be found in [1].

Performance of a digital watermarking scheme can be readily described by the three conflicting requirements – fidelity of marked content, robustness against attacks and the size of hidden payloads. Furthermore, blind detection – detecting hidden watermarks without using unmarked original works – is also an important requirement since in many important application scenarios of digital watermarking, such as DVD copy prevention, the original content is inherently unavailable. However, for most blind-detection watermarking schemes, host-interference – performance loss caused by the host signal during blind watermark detection - is still a serious problem for watermarking system design.

In the early days of digital watermarking, watermarking schemes are non-blind based on the assumption that original contents are readily available in the detection side. For example, [2] introduces the important spread-spectrum watermarking algorithm but the original is assumed to be available while performing payload detection. A spread spectrum steganography scheme is proposed in [3] by embedding messages in the spatial domain. The host interference is alleviated by estimating host signals based on filtered stego-images and introducing ECC to correct corrupted detected payloads. [4] provides a Gaussian-noise model based analysis of the blind-detection spread spectrum watermarking scheme, and proposes some variants of spread-spectrum watermarking that can achieve theoretical performance enhancement. However, the assumption that both the original signal and the attack signal are drawn from a Gaussian distribution is far from reality, and thus the theoretical analysis provided is only an approximation of practical cases. An iterative watermark embedding algorithm that applies a black-box channel model to a work to obtain robust stego-images is provided in [5]. Watermarking performance can be improved for cover works under the attacks considered in the channel model. On the other hand, quantization watermarking schemes are host-interference rejecting. [6] introduces the theoretical framework and implementations of quantization watermarking. Though quantization watermarking schemes can effectively eliminate the host-interference problem, they do suffer from signal scaling attacks.

In this paper, an iterative informed-embedding spread-spectrum approach, conceptually similar to [5], is exploited to alleviate the host-interference problems. Models, algorithms, as well as implementation details are provided to show the effectiveness of our iterative methodology. Furthermore, inherent performance losses caused by the assumption of conventional spread-spectrum watermarking schemes - watermark signals shall be added to most perceptually significant components - is pointed out. Finally, problems and practical applications of iterative informed-embedding watermarking schemes are discussed.

This paper is organized as follows. Section 2 reviews state-of-the art for the blind-detection spread-spectrum watermarking schemes. To improve its detection performance, the iterative informed-embedding methodology is introduced in Section 3. Geometric models, implementation details, and experimental results are also included. The inherent performance loss caused by some assumptions of conventional frequency-domain spread-spectrum watermarking schemes, as well as possible extensions and applications of informed-embedding schemes are investigated in Section 4. Section 5 concludes this paper.

2. BASIC BLIND-DETECTION SPREAD-SPECTRUM WATERMARKING SCHEME

The model of basic blind-detection spread watermarking scheme is introduced in [4] and can be illustrated in Fig. 1. $\boldsymbol{m}=\{b_1,\ldots,b_L\}$ is the payload bit-stream. Each payload bit b_i is embedded into a component of the original host signal, denoted as $\boldsymbol{c_i}$. $\boldsymbol{u_i}=\{u_{i1},\ldots,u_{iN}\}$ is the pseudo-random chip sequence whose value will be +1 or -1 and it will be modulated with each b_i. a is the weighing factor multiplied to the watermark sequence to control the watermark strength. The channel is often modeled as an additive noise $\boldsymbol{n}$. In the detection side, the correlation value of the distorted stego signal and the chip sequence $\boldsymbol{u_i}$ will be calculated and compared against a threshold value (0 in this scheme) to estimate the estimated payload bits. For the ease of explanation, we will focus on the host-interference only in this section, i.e. all the detection errors are caused by host-interference.

In our implementation, watermark signals are embedded into DCT coefficients of the test images. The embedding is done according to the zig-zag scan order from low-frequency coefficients to high frequency ones. This is an approximation of the original spread-spectrum watermarking assumption that watermarks shall be embedded into the DCT coefficients with largest magnitudes. Since in the blind-detection case, the detector cannot precisely find out which are those largest coefficients and their relative orders before watermark embedding, the approximation above is adopted. Fig. 2 shows the detection performance of different test images given L=100, N=500. The value of a is set from 1 to 10 respectively, in other words, each point in Fig. 2 represents the marking performance of some marked image using a certain a value. As expected, the stronger the watermark is, the lower the detection error rate will be and, therefore, the worse the quality of the marked image attains.

Figure 3 shows the watermarking performance of adopting different chip sequence lengths to embed single payload bit into the Lena image, other experimental settings are the same as the previous one. As the error estimation model provided in [4] states, the larger the chip sequence length is, the better the detection performance will be. Fig. 3 largely fits this assumption. Nevertheless, though increasing the length of chip sequence for each payload bit can enhance the detection performance against host-interference, total amount of available payload bits will be consequently reduced. Therefore, the length of chip sequence cannot be unlimitedly increased.

In fact, it is obvious that the basic blind-detection scheme did not utilize any available side-information about the host interference at all. Therefore, an informed-embedding approach for spread spectrum watermarking conceptually similar to the iterative scheme introduced in [5] is devised and tested.

3. ITERATIVE INFORMED-EMBEDDDING SPREAD-SPECTRUM WATERMARKING

(1) System Architecture and Algorithmic Description

The proposed iterative informed-embedder for the blind-detection spread-spectrum watermarking system is shown in Fig. 4. A watermark detector the same as the one in the basic scheme is now included in the embedder. In order to hide a payload stream m, the watermark strength a_i for each payload bit b_i will be gradually increased if the payload bit extracted is not the same as the one to be embedded. As long as a payload bit b_i is correctly extracted, the value of a_i stops increasing immediately. In other words, the more seriously a portion of the watermark sequence suffers from the host interference, the stronger the corresponding weighting factor will be set. The iterative process will be terminated after performing a certain number of iterations or an acceptable error rate is obtained (e.g. all payload bits can be correctly extracted). Note that the IDCT/DCT routines are still performed before sending the marked stego signal into the detector for simulating the actual host-interference the stego image may undergo, but here we don't depict them in Fig. 4 for the purpose of clear illustration. As for the effect of incorporating additional error model, shown as the dotted block in Fig. 4, a thorough discussion will be provided later. Fig.5 shows the detail description of the proposed algorithm.

(2) Experimental Results

Fig. 6 depicts the performance improvements of the proposed iterative approach, using different test images, against the basic blind spread-spectrum watermarking scheme. Experimental settings are given as follows: L=100, N=500, and Δa=1. The performances using the iterative approach obviously outperform those of the basic scheme.

To improve the performance of our iterative scheme further, the increasing step of the watermarking strength, denoted as Δa, is attenuated and tested. In other words, watermarking strengths are fine-tuned to avoid unnecessary coefficient modifications. Fig. 7 shows the detection performance with Δa of 0.1. As we expected, performance improvements achieved via iterative informed-embedding approach can be further improved by adopting finer iterative step of watermarking strength. In other words, watermarking performances are improved at the cost of performing more iteration, that is, more computation resources are involved to produce better marked works.

(3) Oscillations around Detection Threshold

It is worth noting that, during each iterative process, the performance improvement might not increase monotonically. For example, as shown in Fig. 6(b), the detection error rate of the fifth iteration is even higher than the fourth iteration. To clarify this seemingly unreasonable phenomenon, payload bits fail to be correctly extracted in different iterations (corresponding to Fig. 6 (b)) are listed in Table 1. In this case, the 25th payload bit that can be correctly marked in previous iterations fails to be extracted correctly during the 5th iteration. For iterative schemes using finer increasing step, such phenomenon occurs more frequently, as shown in Fig. 7.

We name this phenomenon as oscillations around the detection threshold. That is, the correlation values corresponding to some payload bits altered due to modifications in DCT coefficients in other positions, thus leading to wrongly detected payloads. Though, theoretically, DCT is a one-to-one transform, altering certain DCT coefficients may result in slight changes in other coefficients because of limited precisions and rounding errors of image representations. Since amounts of correlation values over the detection threshold based on the weakly-embedding nature of our iterative scheme are small, wrong detection results may be produced. Fortunately, this situation never lasts for a long time because the energy of corresponding watermark signals will be increased to successfully embed these bits and then the stability of the iterative scheme can be retained.

(4) Geometric Models of Iterative Informed-Embedding

Since spread-spectrum watermark detection is inherently correlation-based, geometric models introduced in [1] can be used to illustrate the proposed iterative spread-spectrum approach. Fig. 8 shows the geometric models of the basic spread-spectrum scheme; Fig. 9 illustrates the geometric model of the proposed iterative approach. In the basic schemes, using the same watermarking strength may result in detection failures or unnecessary modifications that lead to worse fidelity. On the other hand, the iterative informed-embedding approach only adds watermarks to an adequate degree so that the payload bits can be correctly extracted.

4. DISCUSSIONS

(1) Inherent Performance Loss Due to Assumptions of Conventional Spread-Spectrum Watermarking

In addition to that watermarking performance can benefit greatly from adopting the iterative informed-embedding methodology, some more implications are given by carefully examining the experimental results. Roughly speaking, embedding payload bits corresponding to low-frequency DCT coefficients requires larger watermarking strength (i.e., more iterative operations in the iterative approach).

This phenomenon is a reasonable result due to the characteristic of the chosen host signal – the zigzag ordered full-frame DCT coefficients. DCT coefficients possess a tendency of magnitude decreasing from low frequency regions to high frequency ones, therefore, the first few payload bits will be embedded into coefficients with larger magnitudes. This is the desirable property of the original non-blind spread spectrum watermarking schemes since coefficients with larger magnitude are more perceptually significant and will be more robust against media compression. However, without original contents in the detection side, the host signal now becomes a troublesome interfering noise. The larger the noise energy is, the worse the watermark detection will result, and consequently, stronger energy of the watermark signal shall be provided. However, since the fidelity of marked content depends highly on the energy of the signal, it is impractical to arbitrarily enlarge the watermark energy, especially for those perceptually significant low-frequency coefficients. Therefore, the embedding positions are shifted afterwards to alleviate this inherent performance loss. Fig. 10 shows the performance comparisons between the basic spread-spectrum scheme, the iterative informed-embedding approach, and the informed embedding-approach in with all embedding positions are shifted by an N-coefficient offset toward the higher-frequency coefficients. According to Fig. 10, the embedding-positions shifting approach does improve the watermarking performance further.

Though shifting the embedding positions into higher frequency can further improve the performance of watermark detection, unlimited backward shifting of embedding positions will undoubtedly suffers a significant loss in robustness. However, according to the experimental results, the assumption that traditional spread spectrum algorithm insist: watermarks shall be embedded into the most perceptually significant positions do introduce inherent loss of detection performance for the blind-detection scheme.

Frequency-domain watermarking schemes are notable for their better robustness and fidelity over spatial-domain schemes. However, the energy packing capability and larger dynamic ranges of digital transforms may result in strong host interference for blind-detection watermarking schemes, at least for schemes based on correlation values of pseudo-random signals such as the spread-spectrum watermarking.

(2) Concerns about Incorporating Attacking Models

As shown in Fig. 4, the iterative informed-embedding approach can be extended by incorporating error models that describe attacks the marked contents may undergo. However, all informed-embedding approaches face a difficult problem: precisely grasping all important attacks is far from reality. According to [5], obvious improvements in robustness can only be achieved for attacks being considered during watermark embedding only. Nevertheless, our host-interference scheme can successfully remove all host-interferences and be adopted to be basis of informed-embedding approaches taking certain attacks into considerations. Effects of attacks described by known error-models can be iteratively removed by the proposed iterative informed-embedding methodology.

(3) Appropriate Applications of Informed-Embedding Approaches

Though all informed-embedding schemes suffer from the difficulty of predicting all possible error models, they are appropriate for many feasible and important applications. For example, in a typical steganography scheme, guaranteed robustness is an important concern (e.g. a 100% correct extraction is required for decryption of the hidden payloads without adopting error-correcting codes), and all errors, including the host-interference and required lossy compression, are controllable during watermark embedding. Fig. 11 shows the architecture of a typical steganographic system. In addition, associating value-added metadata with lossy-compressed contents also possess similar system characteristics. Therefore, performance of such applications can benefit greatly from the (iterative) informed-embedding schemes.

5. CONCLUSIONS

In this paper, empirical detection performance of basic frequency-domain blind-detection spread spectrum watermarking is observed. Furthermore, possible enhancement through adopting iterative informed-embedding scheme is evaluated. Geometric modes, implementation details and possible applications are provided. Furthermore, the original assumption that original non-blind spread spectrum watermarking schemes followed to embed watermark signals into perceptually significant components may introduce loss of detection performance for the blind-detection cases, and a simple solution is discussed.

6. ACKNOWLEDGEMENT

This work is supported by NSC94-2752-E-002-006-PAE, “Multi-Modal Content Organization”, National Science Council, R. O. C.

7. Reference

[1] Cox, I. J., Miller, M. L., and Bloom, J. A. Digital Watermarking, Morgan Kaufmann Publishers, CA, 2002

[2] Cox, I. J., Kilian, J., Leighton F. T, and Shamoon T,.Secure Spread Spectrum Watermarking for Multimedia. IEEE Transactions on Image Processing, vol. 6, December, 1997

[3] Marvel L. M., Boncelet C. G., Retter C. T., Spread Spectrum Image Steganography, IEEE Transactions on Image Processing, vol. 8, no. 8, August 1999

[4] Marvar, H. S. and Florencio, A. F., Improved Spread Spectrum: A New Modulation Technique for Robust Watermarking, IEEE Transactions on Signal Processing, vol. 51, no. 4, April 2003

[5] Miller, M. L., Watermark Embedding for Black-box Channels, Proceeding of International Workshop on Digital Watermarking 2003 (IWDW 2003), Seoul, October, 2003

[6] Chen, B. and Wornell G. Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding, IEEE Transactions on Information Theory, vol. 47, pp1423-1443, May 2001

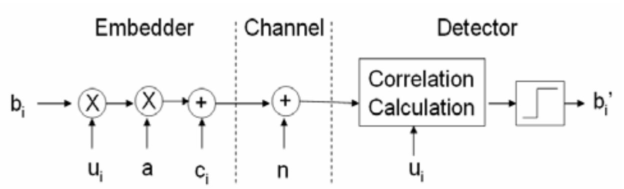


Fig. 1 The basic blind-detection spread-spectrum watermarking scheme.

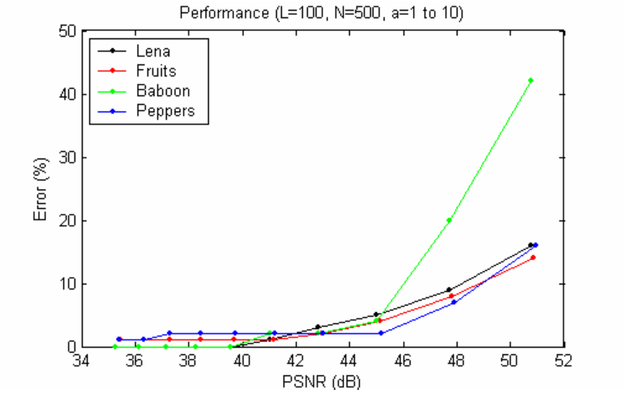


Fig. 2 Performances of the basic blind-detection spread-spectrum scheme using different test images and different watermarking strength, controlled by *a*, are illustrated.

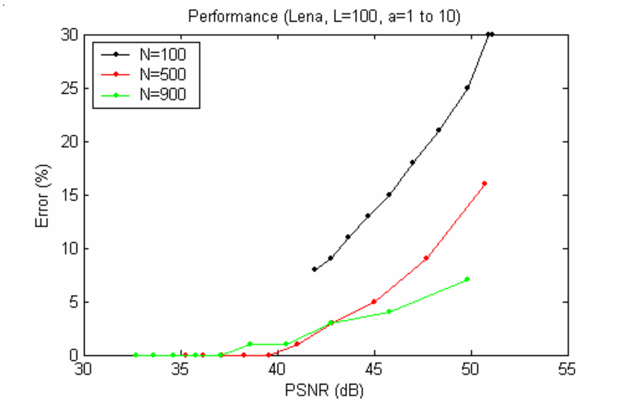


Fig. 3 Increasing the length of chip sequence, denoted as *N*, adequately can substantially improve the detection performance of the spread-spectrum watermarking system.

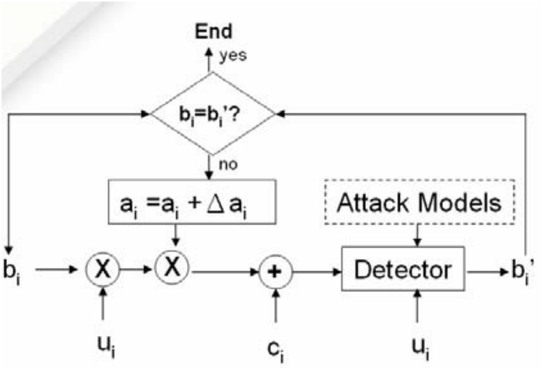
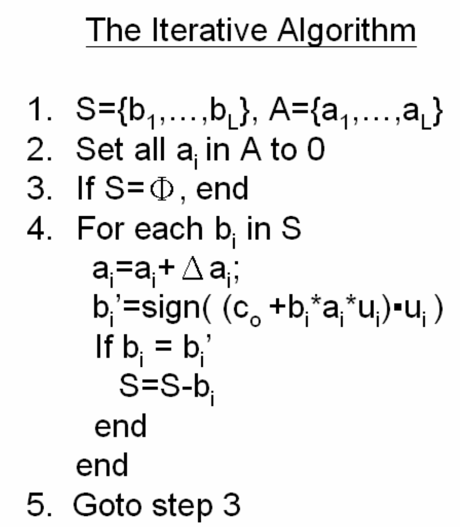


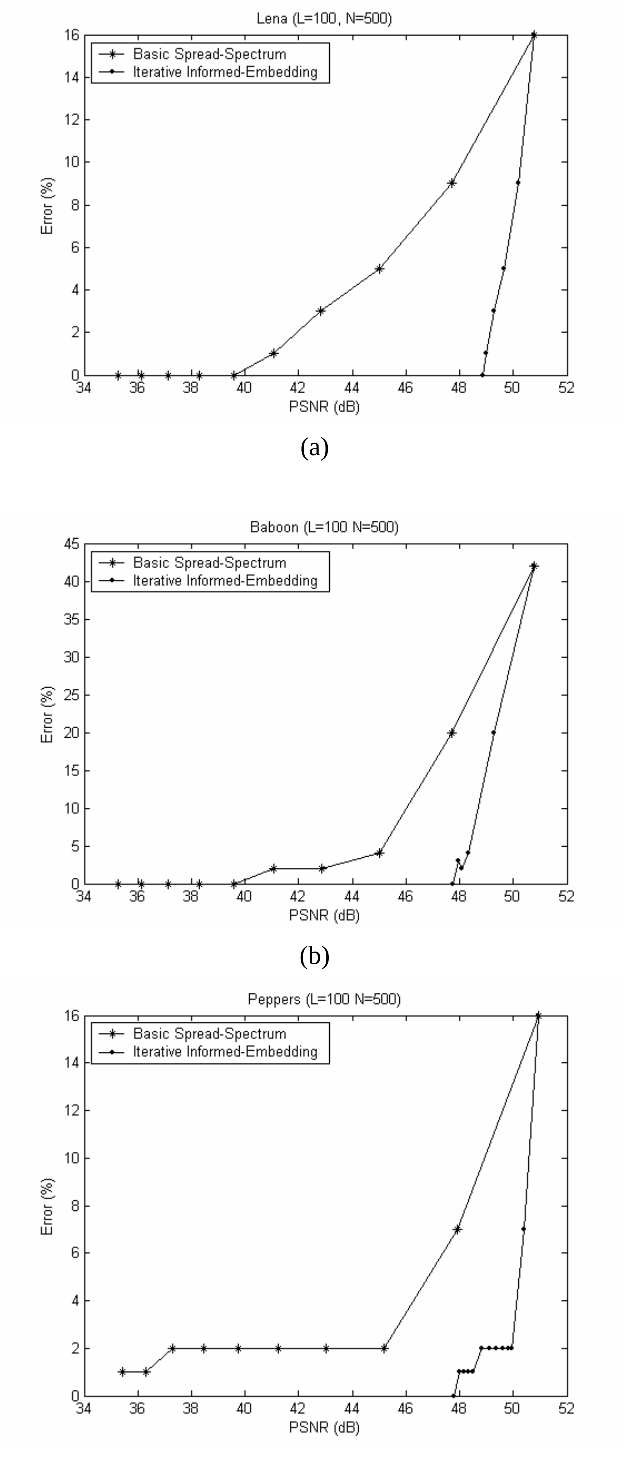
Fig. 4 In an iterative informed-embedder, the weighting factor for hiding each payload bit is iteratively increased until that payload bit can be correctly extracted



The Iterative Algorithm

```
1. $S=\{b_1,\ldots,b_L\}$, $A=\{a_1,\ldots,a_L\}$
2. Set all $a_i$ in A to 0
3. If $S=\Phi$, end
4. For each $b_i$ in S
     $a_i=a_i+\Delta a_i$;
     $b_i'=\mathrm{sign}((c_o+b_i*a_i*u_i)\cdot u_i)$
     If $b_i = b_i'$
       $S=S-b_i$
     end
   end
5. Goto step 3
```

Fig. 5 The algorithm of the proposed iterative informed-embedding approach



(a)

(b)

(c)

Fig. 6 Performance improvements obtained by the proposed iterative informed-embedding against the basic spread-spectrum system, using (a) Lena, (b) Baboon and (c) Peppers, are illustrated.

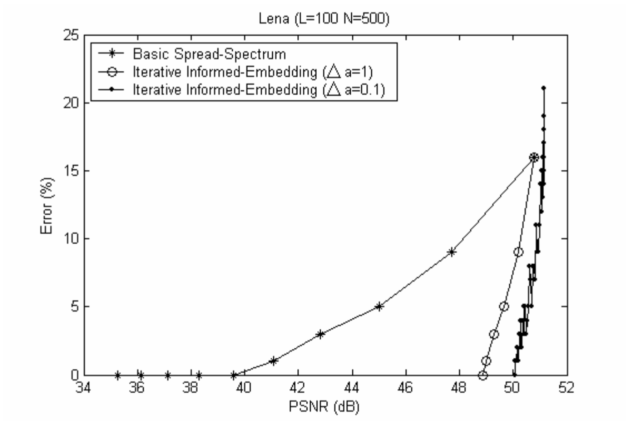
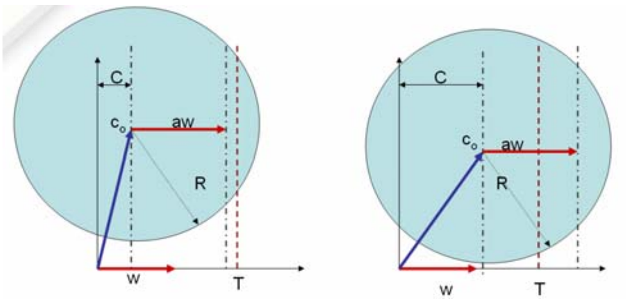
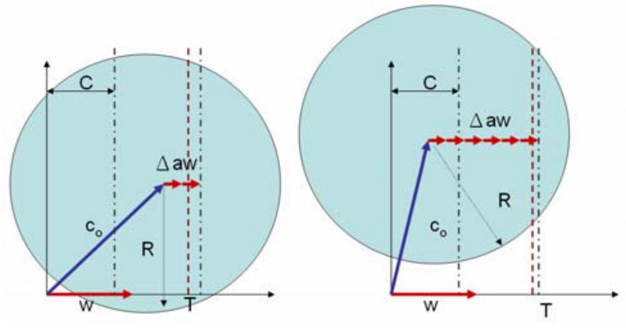


Fig. 7 Using finer increasing-step of watermarking strength can further increase the performance of iterative informed-embedding approach.



(a) (b)

Fig. 8 The geometric model of the basic spread-spectrum watermarking scheme. Fixed watermarking strength may result in (a) failure of detection or (b) more-than-required host-signal modification.



(a) (b)

Fig. 9 The geometric model of the iterative informed-embedding approach is shown. No matter the host-interference is large or small, adequate amount of watermarking strengths that lead to successful payload-bit detection and minimal coefficient modifications can be iteratively decided.

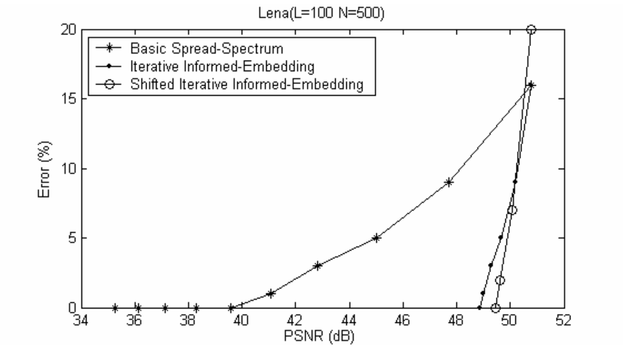


Fig. 10 The watermarking performance can be further improved by adequately shifting the embedding positions toward higher-frequency coefficients.

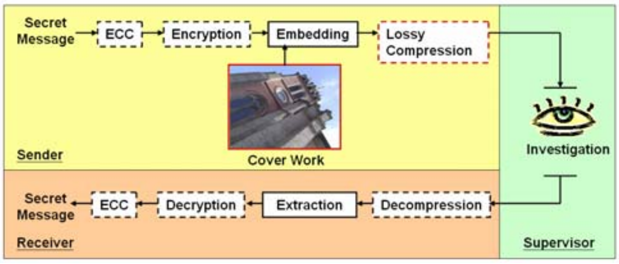


Fig. 11 The system architecture of a typical steganography system.

Table 1. Payload bits fail to be correctly extracted are listed

Iter. No.	Payload Bits failed to be extracted	# of Bits
1	1, 5, 6, 7, 9, 10, 13, 14, 15, 17, 18, 21, 22, 27, 28, 29, 31, 32, 36, 38, 39, 44, 49, 53, 56, 59, 60, 62, 71, 72, 74, 77, 79, 82, 85, 86, 89, 94, 95, 96, 98, 99	42
2	1, 5, 6, 9, 13, 14, 17, 18, 27, 28, 29, 32, 36, 38, 39, 59, 74, 89, 95, 98	20
3	1, 6, 18, 98	4
4	1, 6	2
5	1, 6, 25	3
6	0	0