

BSS: Boosted steganography scheme with cover image preprocessing

Hedieh Sajedi *, Mansour Jamzad

Computer Engineering Department, Sharif University of Technology, Iran

ARTICLE INFO

Keywords:
Steganalysis
Steganography capacity
Support Vector Machine (SVM)
Ensemble methods

ABSTRACT

The existing powerful steganalyzers can find out the presence of secret information in images with high accuracy. Increasing the embedding capacity of cover images reduces the detection risk of stego images. In this respect, we introduce boosted steganography scheme (BSS) that has a preprocessing stage before applying steganography methods. The goal of BSS is increasing the undetectability of stego images. Due to the dependence of embedding capacity of images to their content, we proposed an ensemble steganalyzer to estimate the embedding capacity of each cover image. Since the content of cover images has less significance in steganography, therefore to have more security, the steganographer can select a cover image from a database to achieve higher security and satisfactory embedding capacity. We present several experiments that show the effectiveness of boosted steganography scheme in improving the security of stego images. The experimental results demonstrate that considering a preprocessing stage can significantly improve the steganography security.

© 2010 Elsevier Ltd. All rights reserved.

1. Introduction

Steganography is the art and science of writing secret data in such a way that no one, except the intended receiver, knows of the existence of the data (Marvel, 1999). It utilizes the typical digital media such as text, image, audio, video, and multimedia as a carrier for secret data. Successful steganography depends upon the carrier medium not to attract attention.

The goal of the steganalysis methods is to sense the existence of hidden data from observed data. Steganalyzers have taken great progress in the past few years and a number of powerful steganalysis techniques have been proposed.

Different image steganography methods have been proposed in the literature. Embedding in various steganography methods like F5 (Westfeld, 2001), Model-based (MB) (Sallee, 2003), Perturbed Quantization (PQ) (Fridrich, Goljan, & Soukal, 2004), YASS (Solanki, Sarkar, & Manjunath, 2007), and Contourlet-based steganography (Sajedi & Jamzad, 2008) is done by modifications of properly selected coefficients in a transform domain.

The key performance measure to evaluate different data embedding algorithms is the embedding capacity. In general sense, it is the maximum secret data size that can be securely embedded in an image with respect to the certain constraints. The embedding capacity in most of the steganography methods are given in terms of non-zero DCT coefficients of images. Generally, an image with

high details has more non-zero DCT coefficients and therefore it has a higher capacity.

Due to the complexities in steganography and progressive strength of steganalysis algorithms, it has become a challenge to systematically develop techniques with much better performance. In this paper, we aim to cope with this problem by proposing a scheme to increase the embedding capacity of images. Some images do not have enough varieties to hide a secret data securely. For this reason, we use the lucrative effects of image preprocessing methods to amplify the details of cover images.

Most of the image preprocessing methods strongly depend on the intensity value of image pixels. When these methods are applied to images, they can make more variation in image intensities and the images will be more suitable to host secret data.

In passive steganalysis methods it is assumed that the original cover image is not available. Therefore, it should be difficult to find out if a preprocessing method has been applied on a cover image. Consequently, in steganography methods a cover image can be processed in two stages, preprocessing and embedding. The output of the first stage is passed to the embedding stage in order to hide data. The main idea behind the first stage is to impose more variation in pixel intensities of cover images compared to the original ones. The results of our experiments illustrate that the stego images, which are achieved by hiding secret data in cover images with more variation in their pixel intensities, are less detectable by the statistical examinations of steganalyzers.

To determine the effects of different preprocessing techniques on steganography methods, the PQ steganography technique (Fridrich et al., 2004) has been taken as the baseline test system.

* Corresponding author.

E-mail addresses: a_sajedi@ce.sharif.edu (H. Sajedi), jamzad@sharif.edu (M. Jamzad).

In this paper, we append a new stage to steganography methods to improve the security level of stego images. This improvement is due to the properties of contrast enhancement methods and Successive Mean Quantization Transform (SMQT) enhancement (Nilsson, Dahl, & Claesson, 2005a) on cover images. The SMQT transform has properties that reveal the underlying organization or structure of data. This transform extracts the structure of data in a robust manner, which makes it insensitive to changes in bias and gain in the signal (Nilsson et al., 2005a).

Until now, the embedding capacity of a steganography method is determined based on non-zero DCT coefficients. However, images with different contents may have unequal number of non-zero DCT coefficients. In this state, there is no guaranty that these images would have similar undetectability rate after embedding the same secret data in them. Therefore, the embedding capacity may not be associated to a steganography method rather it depends on the content of cover image. To estimate the embedding capacity of a cover image we use an ensemble steganalysis structure. An ensemble classifier is often used for boosting weak classifiers, such as decision tree and neural networks (Dong & Han, 2005). In this paper, each weak classifier is a steganalyzer and our intent is to discriminate between the secure and non-secure limits of embedding rate in an image. The combination of the vote of all the steganalyzers in the ensemble defines the embedding capacity of a cover image. If the steganalyzers agree with each other that an image is a stego image, the goal of the steganography is obtained. Therefore, we can embed in an image until the distortion of image features does not overrun a safety threshold.

Our experimental results indicate that the embedding capacity increases significantly, if the cover images are processed beforehand. In this regard, SMQT transform performs better than other preprocessing techniques such as linear stretching and histogram equalization.

The rest of this paper is organized as follows: Section 2 proposed boosted steganography scheme. In Section 3, we introduce an approach to estimate the embedding capacity of images. Some image preprocessing methods are reviewed in Section 4. Cover selection steganography based on image properties is discussed in Section 5. Performance of the proposed two-stage steganography method is analyzed in Section 6 and finally, conclusion is given in Section 7.

2. Boosted steganography scheme

In boosted steganography scheme (BSS), a cover image is processed in two stages starting with preprocessing stage, followed by embedding stage. In this approach, preprocessed images with more variation can cover a secret data better than their original version. Such preprocessed images are more proper for visual and statistical examination of steganalysis methods. In the second stage, any of the existing steganography techniques can be applied. Fig. 1 shows the structure of the proposed boosted steganography

scheme, in which a cover image can be determined previously by the steganographer or he may select a proper image from the database according to a selection measure.

In preprocessing stage of BSS, different image processing techniques like contrast enhancement methods, Successive Mean Quantization Transform (SMQT) enhancement and some other manipulation methods can be applied on cover images.

3. Embedding capacity estimation

Steganalysis methods analyze images to decide whether a secret data has been embedded in them. So, a steganalysis method can be seen as a two-class classification problem (Martin, Sapiro, & Seroussi, 2004). The core of each steganalyzer is a classifier, which given an image feature vector, decides whether the image contains any secret data.

In the area of machine learning, the concept of combining classifiers is proposed as a means to improve the performance of individual classifiers. Ensemble learning refers to a collection of methods that learn a target function by training a number of individual learners and combining their predictions. The objective of classification integration algorithms is to generate more certain, precise, and accurate results (Dietterich, 2001). In addition, uncorrelated errors of individual classifiers can be eliminated.

A classification task usually involves with train and test dataset, which consist of some data instances. Each instance in the train set contains one class label and several features. The objective of Support Vector Machine (SVM) is to produce a model, which predicts class label of data instances (i.e. given only by their features (Meyer, Leisch, & Hornik, 2003) in the test set. Here, in steganalyzer case, training vectors X_i are mapped into a higher dimensional space. Then SVM finds a linear separating hyperplane with the maximal margin in this higher dimensional space. In two-class namely, stego and cover image classification, using a SVM, the result of the steganalyzer is obtained as Eq. (1).

$$decision = \begin{cases} I \in stego, & d_j(I) > 0 \\ no-decision, & d_j(I) = 0 \\ I \in cover, & d_j(I) < 0 \end{cases} \quad (1)$$

where $d_j(I)$ is the distance of image I in feature space from decision hyperplane j that separates clean and stego images. In availability of different steganalyzers, each one makes mistakes independently of the rest. As a solution, we investigate cooperation of steganalyzers with the help of ensemble learning methods.

To determine the security of stego images, first, we compose a 'steganalyzer unit', which is a cascade classifier. Then, we consider the maximum result of all steganalyzers as the result of the whole steganalyzer unit as the following:

$$d = \max(d_j(X_i)) \quad (2)$$

A secure upper bound for embedding in an image is determined with regarding to the maximum distance of the image from all steganalyzer discriminant hyperplanes. The distance shows the

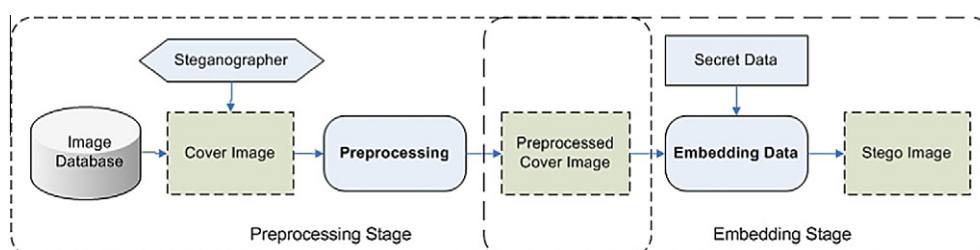


Fig. 1. Boosted steganography scheme (BSS).

closeness of an image to the unsafe region in feature space (stego space). We combine some moderately inaccurate base classifiers (steganalyzers) into a combined predictor to determine the upper bound of embedding capacity of an image. Combining different classifiers to make an ensemble, we can benefit from better classification performance and more resistance to noise than individual classifiers. Each vote (detection result) is the confidence of a classifier on being an image clean or stego.

Embedding capacity of an image may be different using different steganography methods. Therefore, each steganography method can embed the secret data in an image incrementally until it is not threaten to be detected by steganalysis algorithms. The structure for estimation of embedding capacity is constructed based on the following efficient and famous steganalysis methods in the literature:

- (1) Wavelet-based steganalysis method (WBS) proposed by [Lyu and Farid \(2002\)](#). In the feature extraction part of this method, statistics such as mean, variance, skewness, and kurtosis are calculated from each wavelet decomposition subbands.
- (2) Markov-DCT based steganalysis method (274-dim) has a 274-dimensional feature vector that merges Markov and DCT features ([Pevny & Fridrich, 2007](#)).
- (3) 324-dimensional feature vector steganalysis method (324-dim) proposed by [Chen, Shi, Chen, and Xuan, \(2006\)](#) is based on statistical moments of wavelet characteristic functions.

[Fig. 2](#) shows the structure of the ensemble steganalyzer and the mechanism to estimate the embedding capacity of an image. Using this structure, we estimate the embedding capacity of images. This combination of steganalyzers fills some gaps between feature spaces and achieves a good estimation of embedding capacity regarding to the advantages of steganalysis methods.

In presence of a cascade classifier in the steganalyzer unit, each one is a discriminator between certain payload stego images and clean images. The distance of an image in feature space from a SVM classifier represents the confidence of the vote of SVM. The vote is positive if it is recognized as a cover image; otherwise, it is negative. Therefore, threshold τ with value of zero can determine the result of each classifier. We use Radial Basis Function (RBF) kernel that nonlinearly maps samples into a higher dimensional space, so unlike the linear kernel, it can handle the case when the relation between class labels and attributes is nonlinear.

To construct each steganalyzer unit we quantized the range between 0 to 10,000 bits of secret data to ten equal parts. Since when a steganalyzer detects a stego image, the steganography purpose gets broken, the steganalyzer unit checks each classifier (steganalyzer) in an ascending payload order. If any one detects the stego image, the unit stops and reports the result without checking other

classifiers. [Fig. 3](#) shows the arrangement of a steganalyzer unit. All the steganalyzers in a steganalyzer unit are from the same type (WBS, 324-dim or 274-dim). Steganalyzer_i , $i = 1, \dots, n$, in [Fig. 3](#) is trained to detect the stego images that hide a certain payload.

It seems that if a steganalyzer classifies stego and clean images with payload of 1000 bits, this classifier can correctly detect stego images with higher payloads as well. To verify this assumption, we did some experiments and realized that this is correct. However if a classifier has been trained for each payload, the detection accuracy is higher. Consequently, to have higher detection accuracy we train one classifier for each quantized payload and let a cascade classifier to detect stego images. We call each cascade classifier a steganalyzer unit. The result of a unit is the vote of confidence the unit gives to an image.

The most secure state for a stego image is when all the units in the ensemble steganalyzer announce that the image is clean. To determine the embedding capacity of an image, we increase embedding rate until the maximum distance of the image from discriminants in feature space reaches zero. While the following relation remains true, it implies that the image is a clean.

$$\max(d_{ui}) \leq 0 \quad (3)$$

where $d_{ui} \in [-1, 1]$ is the distance between the result of i th steganalyzer unit and the safe embedding threshold τ . To have secure steganography, one distinct classifier is considered for each payload and each steganalysis method is used to compose one steganalyzer unit. Our experimental results illustrated that distinct classifier for each quantized payload provides more accuracy for steganalyzer. Using our proposed approach, we can guarantee the security of a stego image against the used steganalyzers if it hides a secret data with size of equal or smaller than its embedding capacity.

4. Image preprocessing

Image preprocessing specially enhancement has been an area of active research for decades. Most studies were aimed at improving the quality of image display for better visualization. Yet few studies have been conducted to investigate the impact of image enhancement on some pattern recognition problems ([Wu, Wang, Liu, & Chen, 2002](#)). Some primitive enhancement methods are available in some toolboxes and numerous image enhancement techniques have been developed and published in the literature ([Chang & Wu, 1998](#); [Wang, Wu, Castelman, & Xiong, 2001](#)). Such techniques are used to improve an image, where improve is sometimes defined objectively (e.g., increase the signal-to-noise ratio), and sometimes subjectively (e.g., make certain features easier to see by modifying the colors or intensities). However, the aim of image enhancement is to improve the interpretability or perception of information in images for human viewers, or to provide better input for other automated image processing techniques.

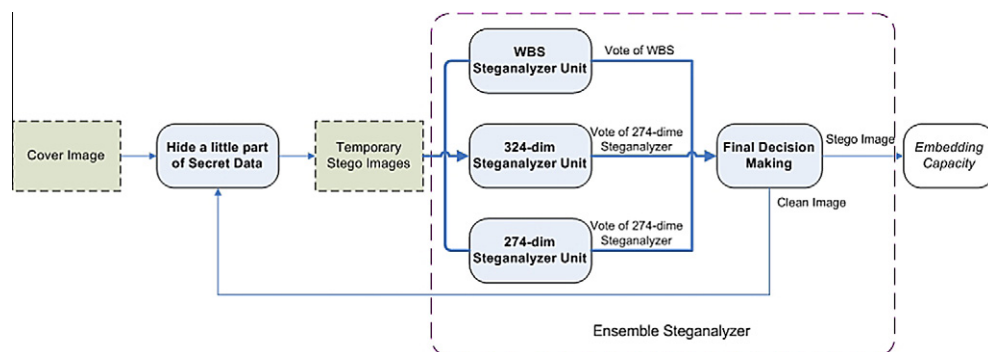


Fig. 2. Structure of the ensemble steganalyzer and the mechanism to estimate the embedding capacity of an image.

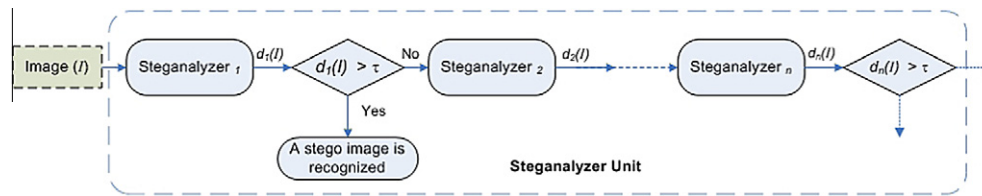


Fig. 3. Structure of a steganalyzer unit.

Most studies were aimed at improving the quality of image display for visualization by increasing the visibility of low-contrast images. Image preprocessing techniques can be divided into two broad categories: first, spatial domain methods, which operate directly on pixels, and the second, frequency domain methods, which operate on the Fourier transform of an image. Unfortunately, there is no general theory for determining what is a good image preprocessing method when it comes to human perception. However, when image enhancement techniques are used as preprocessing stage for other image processing techniques, then quantitative measures can determine which techniques are most appropriate. There are many enhancement techniques that can be selected according to the need. In the following we give a short description on a few such techniques that were used in this research.

4.1. Contrast adjustment and histogram equalization

Contrast adjustment is an image enhancement technique that maps an image intensity values to a new range. The result is a linear mapping of a subset of pixel values to the entire range of grays, producing an image of higher contrast.

Histogram equalization (HE) is one of the most popular enhancement techniques. The HE enhances the contrast of images by transforming the intensity values of an image so that the histogram of the output image is distributed over the entire allowable range of gray levels.

4.2. SMQT transform enhancement

Producing digital images that render contrast and details well is a strong requirement in several areas, such as remote sensing, biomedical image analysis and fault detection (Dong & Han, 2005). Performing these tasks automatically without human intervention is a particularly hard task. Different approaches and techniques have been suggested for this problem (Nilsson et al., 2005a, Nilsson, Dahl, & Claesson, 2005b). SMQT uses an approach that performs an automatic structural breakdown of information. This operation can be seen as a progressive focus on the details in an image. SMQT transform reveals the organization or structure of the data and removes properties such as gain and bias. SMQT can be used to extend the structure representation to an arbitrary predefined number of bits on arbitrary dimensional data. In image processing, SMQT transform is applied in automatic image enhancement and dynamic range compression (Marvel, 1999). Fig. 4 shows the enhanced version of some images by histogram equalization (HE) and SMQT transform. As it is shown, the SMQT enhances images more indigenously than HE. However, if HE is applied to images with a very light background, it may produce very dark regions and vice versa, which can cause security of steganography to fail.

4.3. Brightening and darkening

Brightness refers to the overall lightness or darkness of an image. To have a brightened or darkened image we multiply or divide respectively all the intensities by a constant greater than 1.

4.4. Blurring and sharpening

Blurring is the opponent of contrast enhancement. In this paper, a two-dimensional Gaussian low-pass filter of size 3×3 and standard deviation of size 0.5 is applied to make a blurred image.

To sharpen an image, it is needed to make the intensity transitions more acute. To do this, a highpass filter or unsharp contrast enhancement filter can be applied. In this process, an image is sharpened by subtracting a blurred (unsharp) version of the image from itself.

5. Cover selection based on cover image properties

The cover object in steganography acts only as a carrier for secret messages. Therefore, the embedder is allowed to choose any cover images from the database using a cover selection steganography. A cover selection technique for hiding a secret image in a cover image was first introduced in Kermani and Jamzad (2005). This method operates based on image texture similarity and replaces some blocks of a cover image with similar secret image blocks; then, indices of secret image blocks are stored in the cover image. In this cover selection method, the blocks of the secret image are compared with the blocks of a set of cover images and the image with most similar blocks to those of the secret image is selected as the best candidate to carry the secret image. An improvement on this method is proposed in Sajedi and Jamzad (2008) that uses statistical features of image blocks and their neighborhoods. Using block neighborhood information prevents appearance of virtual edges in the sides and corners of the replaced blocks. In Kharrazi, Sencar, and Memon (2006), the cover selection problem was studied by investigating three scenarios in which the embedder has either no knowledge, partial knowledge, or complete knowledge of the steganalysis method.

In cover selection methods, a batch process determines the value of cover selection measure for each image in a database and the results are stored in a measure value database. When the steganographer wants to select a cover image, he can refer to this database and choose a proper image to hold his secret data. Contrast, brightness, and darkness can be used as a measure for cover selection. In this way, to have a secure covert communication one can select a cover image with high contrast, brightness, or darkness from the database. The reason behind such selection, as will be explained in the following section, is that, in general, images with higher contrast, brightness, and darkness provide more embedding capacity than other enhancement techniques.

Applying classical mathematical theory to steganography problem would require estimating the probability distribution function (pdf) of cover and stego images. It is hard to calculate such estimation, because the feature spaces that result in complete and practical models for images are still relatively high-dimensional. Consequently, to obtain accurate parametric or non-parametric models, in practice the problem of estimating a pdf is replaced with a simpler problem of classification. A classifier can be trained on features derived from a database (Pevny & Fridrich, 2008). In our work the pdf of cover and stego images are estimated with SVM classifiers. With the proposed cover selection approach, we

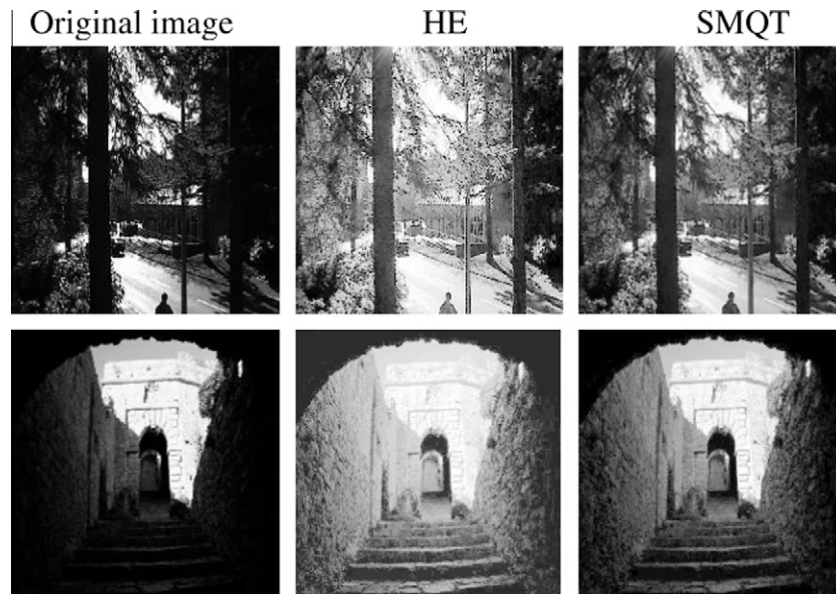


Fig. 4. Results of applying histogram equalization and SMQT transform on some images.

select from an image database the cover images that their pdfs do not change a lot after data embedding. Our experiments showed that the steganalyzers described in Section 3 were unable to detect such stego images.

6. Experiments

To evaluate the proposed boosted steganography scheme different experiments were done. Comparison experiments were conducted with 1000 JPEG images from Washington University image database (<http://www.cs.washington.edu/research/image-database>) and some other images. All the images were converted to grayscale images. Then we cropped them to images of size 512×512 . To make stego datasets, random binary data is embedded into images by using PQ steganography method. The parameters of PQ method were set to different amounts to construct 10 stego datasets with different payloads. Each stego dataset with certain payload has 1000 stego images. To construct a certain payload dataset, for example, 2000-bit-payload stego dataset, those images that have the payload between 1500 and 2500 bits reside in it. Input images and output stego images of PQ method are in JPEG format with the quality factor of 85 and 70, respectively. Each steganalyzer that is used to compute the embedding capacity is achieved as the following. From the database, 2000 images (1000 clean and 1000 stego) were used for training a SVM. The SVM was trained with the statistics that are extracted from the training image subset.

6.1. Relation between complexity and embedding capacity of images

This experiment is arranged to investigate the relation between the complexity of cover images and detectability of the produced stego images. For this purpose, first we group all images in the database according to their complexity. Then PQ steganography method is used for evaluation of embedding capacity of each group. We apply Uniformity as a complexity measure computed by using co-occurrence matrix (Haller, 1970). We calculate the logarithm of image complexity values and divide the range of the results to five equal intervals namely, very low, low, middle, high and very high image complexity. Then the embedding capacity of each image is computed and at last, the average value and the standard

deviation of embedding capacities in each group are achieved. The results show that to have a high capacity cover image, it is preferred to select low, middle, and high complexity images. In contrast, very high and very low complexity images do not have a high embedding capacity.

6.2. Embedding capacity of smooth and non-smooth regions

The results of previous experiment show that middle complex images have the most embedding capacity among other images. On the contrary, very high and very low complex images do not have a high embedding capacity. We can induct that for a stego image to be undetectable it should have both smooth (very simple) and non-smooth (very complex) regions. To investigate this observation more precisely, in this experiment, we calculate the embedding capacity of three groups of images that have only smooth regions, non-smooth regions or a combination of both. Fig. 5 shows some examples of these images and their embedding capacities. The first and the second rows of Fig. 5 show very simple and very complex images, respectively. The images in the third row are a combination of both very simple and very complex images in the same column. As we see, the combination of both types of images has more embedding capacity than the original images. The last row shows some very complex images that are framed by a very simple border. High embedding capacity of these images compared to both very complex and very simple images, indicates that to increase the embedding capacity and reduce the detectability of a stego image we can add a simple (smooth and very low complex pattern) full frame around a very complex cover image or vice versa. Instead of framing manipulation, the steganographer can select a cover image from the database that has both very simple and very complex regions.

6.3. Effect of cover image preprocessing on embedding capacity

Some random images are selected from the database and the embedding capacities of them using PQ steganography method are determined. Then we preprocessed them by employing some preprocessing methods and calculated the embedding capacity of them again. The results are shown in Fig. 6. The figure demonstrates that in most of the images, contrast enhancement increases

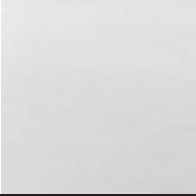
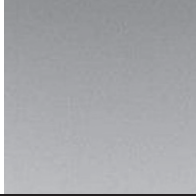
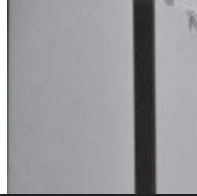

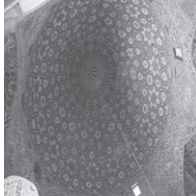


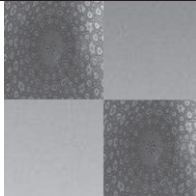
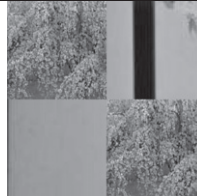

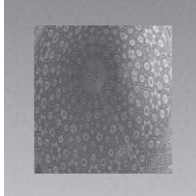

very simple cover image			
Embedding capacity (bits)	1051	727	716
very complex cover image			
Embedding capacity (bits)	1587	1652	1525
combination of very simple and very complex images			
Embedding capacity (bits)	3811	5135	6993
full framing very complex images with very simple images			
Embedding capacity (bits)	7055	4824	7209

Fig. 5. Embedding capacity (in bits) of some very simple and very complex cover images and their combinations.

the embedding capacity largely, but in few cases enhancement decreases the embedding capacity a little.

The effects of preprocessing on all the images with different levels of complexity in the database and their embedding capacity are shown in Fig. 7. The values are the mean of embedding capacity of images before and after preprocessing. The results indicate that, contrast adjustment, HE, sharpening and SMQT transform increase the embedding capacity and among these operations, sharpening is more effective. However, HE, darkening, brightening, and sharpening may visually be detectable if the effect of the modification on the images is great but SMQT enhances the images locally and is less detectable than other preprocessing methods. Despite others, SMQT maintains the natural state of an image while increases its

capacity. It may be thought that the images with lower quality have superior ability to carry embedded data and the visual artifact caused by the steganography methods is less detectable. However, we see in Fig. 7 that blurring (that produces a low quality image) that has the inverse effect of contrast enhancement decreases the embedding capacity and the statistical artifacts produced in these images are more detectable.

To investigate that if the steganalyzers detect the preprocessed images as stego images, we arranged an experiment in which 1000 clean images and 1000 preprocessed images (with SMQT transform) are used. The detectability of these images is evaluated against some steganalyzers. In this experiment, 600 clean images and 600 preprocessed images are used to train a SVM classifier







Original image						
Embedding Capacity of Original image	2177	2575	5666	4700	1801	1910
Contrast adjustment	10116	9764	9765	4650	1789	10100
Histogram equalization	9511	727	9575	9887	1849	9824
SMQT enhancement	10129	9659	9618	9845	1827	9848
Darkening	2796	7690	9821	4701	2124	6898
Brightening	2101	4442	4632	7645	1624	9226
Blurring	2083	2562	2778	4640	1679	1839
Sharpening	8421	9912	10074	9896	2415	8159

Fig. 6. Embedding capacity (in bits) of some cover images before and after cover image preprocessing.

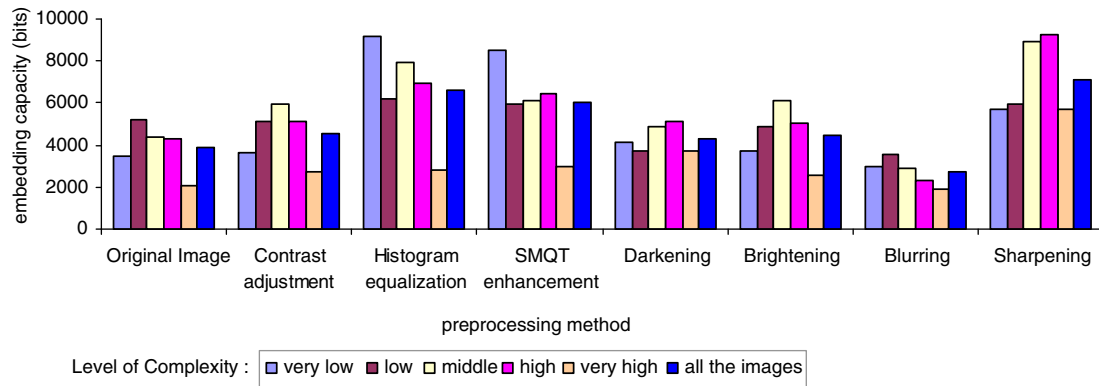


Fig. 7. The mean value of embedding capacity of cover images (in bits) with different levels of complexity before and after the cover image preprocessing.

Table 1

Detection accuracy (%) of the steganalyzers on the preprocessed images using SMQT transform.

	Preprocessed images using SMQT transform		
	WBS	274-dim	324-dim
Detection accuracy (%)	53	57	54

for each steganalyzer and 800 remaining images are used to test. The results are shown in Table 1. The results demonstrate that the steganalyzers cannot discriminate between clean and preprocessed images reliably. Therefore, the preprocessing stage increases the embedding capacity of images without threatening the security of them.

6.4. Cover selection based on image properties

In this experiment, contrast, brightness, and darkness are considered as the measures for cover selection. Fig. 8 shows the result of cover selection based on these measures. As we see, embedding

in higher contrast images results in greater embedding capacities. It is deduced from the experimental results that in each complexity group, the images with higher contrast have higher embedding capacities. In the first row of the figure, the images are the ones that in five complexity levels have the highest contrast. Contrast, brightness, and darkness of an image are computed using the functions that exist in image processing toolbox of Matlab 2007.

6.5. Evaluating the security of the proposed approach

To evaluate the security of stego images, we test the proposed approach with 500 images, which are collected from some sets of images that are taken with three cameras with different resolutions. We cropped the images to size of 512×512 and converted them to grayscale. These images have not shown to the steganalyzers that determine the embedding capacity of images. Then to construct stego image dataset, a secret data (random binary string) is embedded in each image in the dataset so as the size of secret data is equal or smaller than the embedding capacity of the cover image. Then the detectability of these images is evaluated against the state-of-the-art steganalyzers. Table 2 shows the results.

Complexity		very low	low	middle	high	very high
Selection Measure	Contrast					
	Embedding capacity (bits)	4400	5619	9172	4619	2890
Darkness						
	Embedding capacity (bits)	1591	4646	5125	4764	1980
Brightness						
	Embedding capacity (bits)	4000	4794	5367	6885	2139

Fig. 8. Embedding capacity (in bits) of selected covers based on contrast, darkness, and brightness.

Table 2

Detection accuracy (%) of the steganalyzers on the stego images, which are produced by our proposed approach.

Payload	Classical PQ steganography method Steganalyzer			Proposed approach using PQ steganography method		
	WBS	274-dim	324-dim	WBS	274-dim	324-dim
2000 bits	72	74	57	54	58	56
6000 bits	76	77	83	55	53	57
10000 bits	79	79	91	59	62	60

As the outcomes show, that preprocessed images can carry the secret data with high security without attracting the attention of the steganalyzers.

7. Conclusion

Because of the complication of the steganography problem and progressive power of steganalysis algorithms, it is a hard problem to build up techniques with considerable better performance. To improve the security of the existing steganography methods, in this paper, we proposed a two-stage boosted steganography scheme that processes a cover image before embedding the secret data in it. The experimental results illustrate that cover image pre-processing is very valuable for increasing the embedding capacity of images, especially for images acquired under unconstrained illumination conditions. However, the effects on embedding capacity of images vary using different preprocessing methods. Sharpening and HE provide more embedding capacity than SMQT transform. However, SMQT enhancement results in more normal images (e.g. according to its appearance) and the enhancement is not apparent visually. Sharpening and HE enhancement may be detectable with some simple processing but SMQT that has a local enhancement is not detectable easily. The image that its quality is enhanced by SMQT transform is usually perceptually indistinguishable from the original image. The results of our experiments show that the stego images produced by our proposed technique are not detected reliably by the state-of-the-art-steganalyzers.

Acknowledgment

We would like to thank Iran Telecommunication Research Center for their financial support.

References

- Chang, D. C., & Wu, W. R. (1998). Image contrast enhancement based on a histogram transformation of local standard deviation. *IEEE Transactions on Medical Imaging*, 17(4), 518–531.
- Chen, C., Shi, Y. Q., Chen, W., & Xuan, G. (2006). Statistical moments based universal steganalysis using JPEG-2D array and 2-D characteristic function. In *Proceeding of ICIP, Atlanta, GA, USA* (pp. 105–108).

- Dietterich, T.G. (2001). Ensemble methods in machine learning. Multiple classifier systems. LNCS (Vol. 1857, pp. 1–15). Springer.
- Dong, Y., & Han, K. (2005). Boosting SVM classifiers by ensemble. In *Proceeding of 14th international ACM conference on World Wide Web* (pp. 1072–1073).
- Fridrich, J., Goljan, M., & Soukal, D. (2004). Perturbed quantization steganography with wet paper codes. In *Proceeding of ACM multimedia workshop, Germany*.
- Haller, R. S. (1970). Complexity of real images evaluated by densitometric analysis and by psychophysical scaling. MSc thesis, University of Arizona.
- Kermani, Z., & Jamzad, M. (2005). A robust steganography algorithm based on texture similarity using gabor filter. In *Proceeding of IEEE international symposium signal processing and information technology* (pp. 578–582).
- Kharrazi, M., Sencar, H., & Memon, N. (2006). Cover selection for steganographic embedding. In *Proceeding of ICIP* (pp. 117–121).
- Lyu, S., & Farid, H. (2002). Detecting hidden messages using higher-order statistics and support vector machines. In *Proceeding of 5th international workshop on information hiding*.
- Martin, A., Sapiro, G., & Seroussi, G. (2004). Is image steganography natural? Technical Report, Information Theory Research Group, HP Laboratories, Palo Alto.
- Marvel, L. M. (1999). Spread spectrum image steganography. *IEEE Transactions on Image Processing*, 1075–1083.
- Meyer, D., Leisch, F., & Hornik, K. (2003). The support vector machine under test. *Neurocomputing*, 55, 169–186.
- Nilsson, M., Dahl, M., & Claesson, I. (2005). The successive mean quantization transform. In *Proceeding of international conference on acoustics, speech, and signal processing* (pp. 429–432).
- Nilsson, M., Dahl, M., & Claesson, I. (2005). Gray-scale image enhancement using the SMQT. In *Proceeding of international conference on image processing*.
- Pevny, T., & Fridrich, J. (2007). Merging Markov and DCT features for multi-class JPEG steganalysis. In *Proceeding of SPIE, San Jose, CA*.
- Pevny, T., & Fridrich, J. (2008). Novelty detection in blind steganalysis. In *Proceeding of ACM multimedia and security workshop, Oxford, UK, September 22–23* (pp. 167–176).
- Sajedi, H., & Jamzad, M. (2008). Cover selection steganography method based on similarity of image blocks. In *Proceeding of IEEE 8th CIT conference, Sydney, Australia* (pp. 379–384).
- Sajedi, H., & Jamzad, M. (2008). Adaptive steganography method based on contourlet transform. In *Proceeding of 9th international conference on signal processing*.
- Sallee, P. (2003). Model-based steganography. In *Proceeding of international workshop on digital watermarking, Seoul, Korea*.
- Solanki, K., Sarkar, A., & Manjunath, B. S. (2007). YASS: Yet another steganographic scheme that resists blind steganalysis. In *Proceeding of 9th international workshop on information hiding*.
- Wang, Y., Wu, Q., Castelman, K. R., & Xiong, Z. (2001). Image enhancement using multiscale differential operators. In *Proceeding of international conf. on acoustics, speech, and signal processing*.
- Westfeld, A. (2001). F5 – a steganographic algorithm: high capacity despite better steganalysis. In *Proceeding of 4th international workshop on info. hiding*.
- Wu, Q., Wang, Y., Liu, Z., & Chen, T. (2002). The effect of image enhancement on biomedical pattern recognition. In *Proceeding of EMBS/BMES conf., Houston*.