EURASIP Journal on Image
and Video Processing

RESEARCH Open Access

Efficient cover image selection based on spatial block analysis and DCT embedding

Sa'ed Abed*, Suood Abdulaziz Al-Roomi and Mohammad Al-Shayeji

Abstract

In steganography, the cover medium is widely treated as a mere container for the embedded information, even though it affects the stego-image quality, security, and robustness. In addition, there is no consensus on the characteristics of a suitable cover image. In this work, we introduce and practically prove the most suitable cover image (MSCI) framework to automatically select a cover image for a given secret image. This paper proposes choosing the most suitable cover from a set of images based on two steps. First, a set of cover images is filtered based on relative entropy and a histogram in order to identify the most suitable candidates. Second, the local block pixel intensity features of the candidates are analyzed to select the most suitable cover image. Furthermore, cover image local blocks were optimized, using rotation and flipping, during the embedding process to further improve stego-image representation. The proposed framework demonstrated high visual image quality when compared with existing solutions. Steganalysis tests indicated that the proposed solution for cover selection provided an increased resistance to modern steganalyzers with up to 30% lowered detection rate, which improved security.

Keywords: Steganography, DCT, Local features, Cover selection

1 Introduction

Intelligent methods for information analysis are rapidly growing. One of the most important and popular methods is secure digital communication. This field is closely related to steganography, which is known as "the art of data hiding". In general, such a process embeds a secret message (text, image, sound, etc.) into a media container (sound, image, video, etc.) for hidden communication.

There have been a considerable number of steganographic techniques developed in the literature [1]. A major proportion of these techniques are devoted to image steganography, specifically. Images, in general, are excellent candidates for data hiding, due to high capacity and redundancy. In addition, visual information is widely used in public domains and networks. This paper focuses on the efficient embedding of visual content into a cover image to effectively hide the secret data.

The challenge is that there are two criteria for the optimization of secret data embedding: visual quality and security [1]. The former is determined as a capability of the human visual system (HVS) to perceive the visual changes caused by steganographic embedding. The latter is a performance of so-called steganalysis tools [1], which try to detect the existence of hidden information in media. These instruments include various algorithms for the statistical analysis of the local structure of steganographic containers. Since the sender chooses a steganographic container, a key question is how such a choice affects the visual quality and security of the stego-image.

In this paper, a novel automatic cover selection framework called the most suitable cover image (MSCI) is proposed. The high-level diagram of the proposed cover selection is illustrated in Fig. 1. We comprehensively analyze the effect of the cover image selection on the steganography outputs. The problem is considered on two scales. First, the developed framework uses global-level filtering to reject unsuitable cover images. A combination of image histogram and entropy is proposed for the global filtering of a cover image database. Next, an analysis of block-level similarity is performed on the secret image and candidate covers to determine the most suitable cover image. In contrast to [2, 3], a different local feature set and cover image similarity measure is proposed. Also, pixel intensities are used directly to form the feature vector of local blocks. Additional geometric

* Correspondence: s.abed@ku.edu.kw
Computer Engineering Department, College of Engineering and Petroleum, Kuwait University, Kuwait, Kuwait

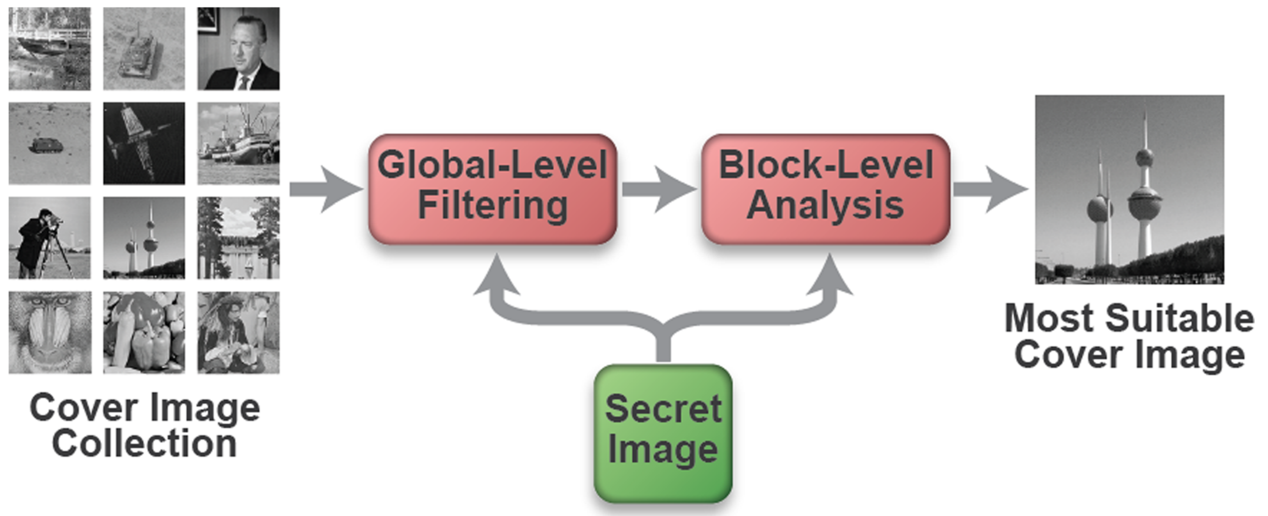

Fig. 1 High-level illustration of the proposed approach: most suitable cover image (MSCI)

manipulations to the local blocks of the secret image are introduced as a visual quality improvement mechanism. The proposed framework was tested, evaluated, and compared with existing solutions. The framework was practically proven and shown that random cover selection is much worse than adaptive selection of a proper image container. Steganalysis tests indicated that the proposed solution demonstrated high visual image quality and provided an increased resistance to modern steganalyzers with up to 30% lowered detection rate, and hence improved the security.

The contributions of the paper are as follows:

- Analyze the effect of the cover image selection on the output of steganographic scheme.
- Conduct a comprehensive analysis of all aspects of cover medium and its effects on the steganography metrics.
- Propose a novel approach to choose a single suitable cover image from a collection of available covers that achieves high quality and security when hiding secret content.
- The proposed framework demonstrated superior image quality values compared with the current work in literature.

The paper is organized as follows. Section 2 contains a brief review of existing steganographic techniques, and a description of key issues related to discrete cosine transform (DCT) steganography, state-of-the-art results in cover selection, the specifics of cover image analysis, and the MSCI framework structure are described in Section 2. Experimental results using the BOSS database are discussed in Section 3, where both the visual quality and security levels are analyzed. Finally, Section 4 summarizes the key findings.

2 Materials and methods

2.1 A review of steganography embedding techniques and principles of DCT domain steganography

There are two types of information embedding used in steganography: embedding in the spatial domain and in the frequency domain (transform domain) [1, 4]. Spatial domain techniques address the intensity of image pixels directly to encode the bits of the secret message. Among these methods, the least-significant bit (LSB) [1] is the oldest and the most popular, due to its simplicity in implementation. A wide variety of spatial domain steganography methods has been developed. A recent example uses texture synthesis to create a cover for spatial embedding [5]. The main weakness of spatial techniques is the sensitivity to common operations, such as transformation and compression.

In the case of frequency (transform) domain steganography, the cover medium is first transformed to another domain. In image processing, the frequency is typically related to the Fourier transform, but in steganography, it is related to DCT. Proper use of DCT coefficients as information containers provides good visual quality and security [1]. The development of a uniform embedding distortion function to find a codeword with the lowest distortion is an example of such techniques [6]. Spreading the embedded information between DCT coefficients leads to fewer image distortions and lower detection accuracy [6, 7].

Both types are combined in adaptive (or model-based) steganography. Adaptive steganography is based on both the spatial and frequency domains with an additional layer of a mathematical model [4]. Here, data hiding may be accomplished in different domains providing less disturbance to the cover image. Embedded regions of the cover image are determined based on a special condition. For example, in [8], the edge regions of the image were detected and used for spatial embedding. In [9], the parameters of the edge detector were automatically determined, making spatial embedding more flexible. Similarly, local image complexity was analyzed in [10], highlighting the most suitable regions for embedding. In [11], an appropriate treatment of image pixels improved the steganographic security.

A recently developed example of adaptive steganography is a technique based on the curvelet transform [12]. Low-frequency curvelet coefficients were used to provide high quality stego-objects. In [13], the authors proposed a steganographic technique resistant to image compression. This was achieved by carefully analyzing the DCT coefficient relationship. In [14], the concept of using two steganographic containers simultaneously was presented. Authors claimed that the security level was improved by analyzing two images of the same scene. Analyses of local image patches and different embedding strategies were performed in [15, 16].

The principle of DCT embedding is quite simple. DCT steganography starts with partitioning the cover image into 8 × 8 pixel blocks (Fig. 2a). For each block, the pixels are transferred into the frequency domain. This results in an 8 × 8-frequency energy matrix (also known as a coefficient matrix) that describes the block (Fig. 2b). The frequency increases from the top left corner to the bottom right corner of the matrix (Fig. 2c). The top-left coefficient is referred to as the zero frequency, and it contains the average intensity of the block. HVS is very sensitive to low frequencies and associated distortions. That is why lossy compression techniques usually neglect the information stored in high frequencies. Thus, to avoid compression of the secret message and to reduce its effect on visual quality, the information may be encoded using the relative values of the DCT coefficients, corresponding to middle frequencies. The DCT domain steganography techniques, in general, are more robust and less detectable than spatial techniques.

2.2 Cover image analysis

In this subsection, an analysis of the cover selection problem was performed. Novel procedures for the efficient selection of the best image container are described, including using the global image characteristics to pre-filter the cover image candidates, manipulations with local spatial blocks, new features, and distance metrics for block matching.

2.2.1 Global image features and complexity measures

Since the problem of cover selection requires significant computational burden, a good idea is to perform initial filtering on the cover image database. To do that, a set of image complexity measures and global characteristics can be evaluated. One of the commonly applied characteristics is the DCT complexity measure [1], shown in Eq. (1):

$$\mathrm{IC}_{\mathrm{DCT}} = \sum_{(i,j)\in A} \sum | \mathrm{DCT}(i,j) | \quad (1)$$

where set A corresponds to the lowest frequency DCT cells [1]. A more sophisticated metric was proposed in [17]. The idea is to split the input image into blocks based on a predefined condition. Such an approach is called quad-tree. It was demonstrated that the most efficient measure of the image complexity in this case could be evaluated as shown in Eq. (2):

$$\mathrm{IC}_{\mathrm{quad-tree}} = \sum_{i=1}^{n} (2x_i)^i \quad (2)$$

where n is the number of tree levels, and x_i is the number of pixels on each level. Other commonly applied complexity measures include homogeneity (a metric related to the high-frequency content of an image), the number of corners and edges [18], and uniformity [19].

Based on the values of the image complexity measures, the most complex cover images may be filtered and extracted for further analysis. However, in the case of a uniform secret image with a low number of gray levels,

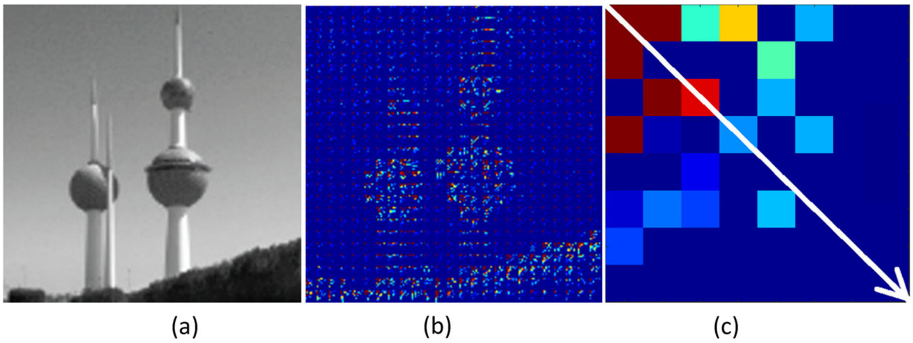

Fig. 2 a Cover image, (**b**) its DCT representation, and (**c**) a single DCT block filtering based on global complexity measures is ineffective [20]. For instance, in the case of the input secret image having quite uniform content with a moderate amount of gray levels, filtering with respect to the image is not optimal.

To make the cover image selection more generic, we propose choosing the cover image that has an entropy [1] value higher than the secret image, as shown in Eq. (3). The entropy of an image may target the difference between neighbor pixels. The highest will be an image with greatest entropy:

$$\Delta E_I = E_1 - E_2 \geq 0, \quad E_i = -\sum_i p_i \log p_i \tag{3}$$

where p_i represents the bins of the image histogram. In this case, Eq. (3) implies that the amount of information required to encode the cover image is larger than for the secret image. For more stringent filtering of the cover candidates, we propose analyzing the histogram bins as well. The following condition is checked in this case:

$$HF_i(I_1, I_2) = \left(H(I_1)_i - H(I_2)_i\right) \geq 0 \tag{4}$$

where H_i is the ith histogram bin and I_1, I_2 are the images to be compared. In our consideration, these are the cover and secret images. Equation (4) implies that the number of pixels for each histogram gray level is sufficient for encoding the secret image pixels for the same gray level.

The proposed filtering algorithm based on image entropy and histograms was compared with the DCT complexity and the measure based on a quad-tree. For this purpose, 1000 images from the BOSS image database [21] were used in this experiment. Using one of the image quality measures, 100 best covers were found. Three secret images with different average brightness and complexity level were embedded in two steps: local block embedding in the spatial domain and hiding the positions in the DCT domain. The embedding principle was described in Section 2.1. The peak signal-to-noise ratio (PSNR) was used to estimate the quality of the obtained stego-images as stated in Eq. (5):

$$\mathrm{PSNR} = 10\ \lg \frac{L^2}{\mathrm{MSE}} \tag{5}$$

$$\mathrm{MSE} = \frac{1}{N_x \cdot N_y} \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} \left(C(i,j) - S(i,j)\right)^2 \tag{6}$$

where MSE in Eq. (6) represents the mean square error, N_x, N_y describe the image size, L is the maximal gray level of the image, $C(i,j)$ is the value of the cover image pixels, and $S(i,j)$ is the value of the stego-image pixels.

Figure 3 shows the time-series of the obtained PSNR after embedding the secret image blocks for each of the three secret images using the three filters. One can

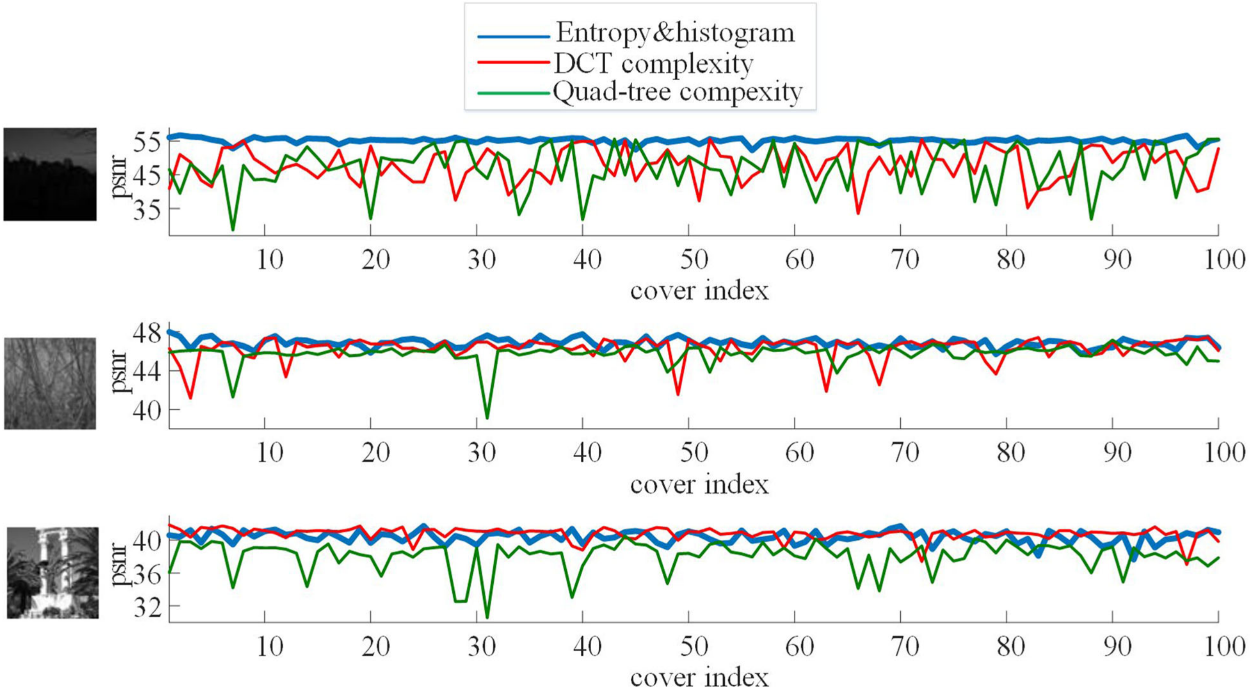

Fig. 3 PSNR for three different secret images and top 100 best covers, chosen by entropy and histogram matching, DCT complexity, and quad-tree complexity analysis

observe that using a gray levels distribution based on the entropy and histogram leads to good and stable results. In contrast, the image complexity measures are sensitive to the type of the input secret image. For example, the DCT metric demonstrates the comparable visual quality results for the most complex secret image (third image) but does not provide stable results for the rest of secret images. The quad-tree complexity measure was suitable for secret images with a large amount of small-scale details. Thus, the proposed entropy and histogram filtering ensures that the amount of information required to encode the cover image is larger than for the secret image and results in a high PSNR level at the same time.

2.2.2 Local spatial block analysis

Selection of features evaluated in local blocks is important for the overall cover analysis procedure. One of the goals is to ensure that the PSNR is quite high after the embedding process. According to Eq. (5), maximization of the PSNR corresponds to minimization of MSE Eq. (6). Since data embedding in the spatial domain is performed by local block replacement, the contribution to MSE is determined by the pixel intensities from these blocks. Thus, Eq. (6) for MSE may be simplified as:

$$\mathrm{MSE} = \frac{1}{N_{\mathrm{blocks}}^{\mathrm{cover}}} \sum_{iB=1}^{N_{\mathrm{blocks}}^{\mathrm{secret}}} \mathrm{MSE}_{iB}, \tag{7}$$

$$\mathrm{MSE}_{iB} = \frac{1}{N_b^2} \sum_{i=1}^{N_b} \sum_{j=1}^{N_b} \left(C\left(i + i_{i\mathrm{CB}}, j + j_{i\mathrm{CB}}\right) - S\left(i + i_{i\mathrm{SB}}, j + j_{i\mathrm{SB}}\right) \right)^2, \tag{8}$$

where $N_{blocks}^{secret} = N_{pixels}^{secret} / N_b^2, N_{blocks}^{cover} = N_{pixels}^{cover} / N_b^2$, N_{pixels}^{secret} and N_{pixels}^{cover} are the number of pixels in the secret and cover images, N_b is the block size, MSE_{iB} is the local block MSE (local MSE), and $(i_{i\mathrm{CB}}, j_{i\mathrm{CB}})$ and $(i_{i\mathrm{SB}}, j_{i\mathrm{SB}})$ are the coordinates of the local block corner for the cover and the secret images, respectively. The definition of the local MSE in Eq. (8) led to directly using pixel intensities as components of the local feature vector. Minimizing the Euclidean distance between the feature vectors of the cover and secret blocks maximized the PSNR value in this case.

Such a "direct" approach is quite different from that proposed in [3], where the authors used the mean, variance, and skewness in 2 × 2 sub-blocks of the 4 × 4 local block to form the feature vector. In the experimental section, both techniques are compared.

In order to improve the visual quality of the stego-image, the local blocks of the secret image were rotated and flipped before embedding. The orientation that provided minimal local MSE was chosen. Possible image manipulations are defined by the following expressions Eqs. (9, 10, 11, 12, 13, 14,15):

$$\mathrm{dst}_{i,j} = \mathrm{src}_{N_b - i, j} - \text{vertical flip} \tag{9}$$

$$\mathrm{dst}_{i,j} = \mathrm{src}_{i, N_b - j} - \text{horizontal flip} \tag{10}$$

$$\mathrm{dst}_{i,j} = \mathrm{src}_{j,i} - \text{flip over the main diagonal} \tag{11}$$

$$\mathrm{dst}_{i,j} = \mathrm{src}_{N_b - j, N_b - i} - \text{flip over anti–diagonal} \tag{12}$$

$$\mathrm{dst}_{i,j} = \mathrm{src}_{N_b - j, i} - 90^\circ\text{rotation} \tag{13}$$

$$\mathrm{dst}_{i,j} = \mathrm{src}_{N_b - i, N_b - j} - 180^\circ\text{rotation} \tag{14}$$

$$\mathrm{dst}_{i,j} = \mathrm{src}_{j, N_b - i} - 270^\circ\text{rotation} \tag{15}$$

here, src and dst are the initial and the resulting blocks, respectively, and *i*, *j* is the pixel index inside the block.

The important point here is that the orientation of each local block is coded in the DCT domain together with its position. Thus, decreasing the PSNR with improved spatial embedding is followed by increasing the amount of hidden information (and thus the distortions) in the DCT domain. The required capacity (in bits) for this case is determined as in Eq. (1):

$$N_{bits} = \log_2\left(N_{blocks}^{cover}\right) N_{blocks}^{secret} * \left(1 + \frac{1}{floor\left(\log_{10}\left(N_{blocks}^{cover}\right)\right)}\right) * K_H, \tag{16}$$

where K_H is the Hamming coding multiplier. It is calculated through the codeword length and the block message length. In this study, $K_H = 7/4$.

The second term in the brackets of Eq. (16) is related to image block manipulation. The descriptors of rotation/flipping of several blocks are grouped into one decimal number of the same bit length as the numbers required for coding hidden block indices ($\log_2\left(N_{\mathrm{blocks}}^{\mathrm{cover}}\right)$). The number of blocks with grouped orientation descriptors is determined by the denominator of the second term in the brackets of Eq. (16) where the floor function rounds a number to the nearest integer, if necessary. The number of DCT coefficient pairs required for embedding is easily calculated from Eq. (16) for a given cover image size.

The effect of image block manipulation on the PSNR of the stego-image was illustrated with an experiment on 50 cover images taken from the BOSS database. The images were chosen by filtering 1000 images with the image entropy and histogram analysis (Section 2.2.1). The image demonstrated in Fig. 2a (resized to 32 × 32 pixels) was used as the secret image. Figure 4 illustrates the PSNR before and after embedding in the DCT domain with and without image block manipulations. The block size was fixed to 4 × 4. The PSNR before DCT embedding for the case using rotation/flipping of the local blocks (green bars) was always the highest value. For

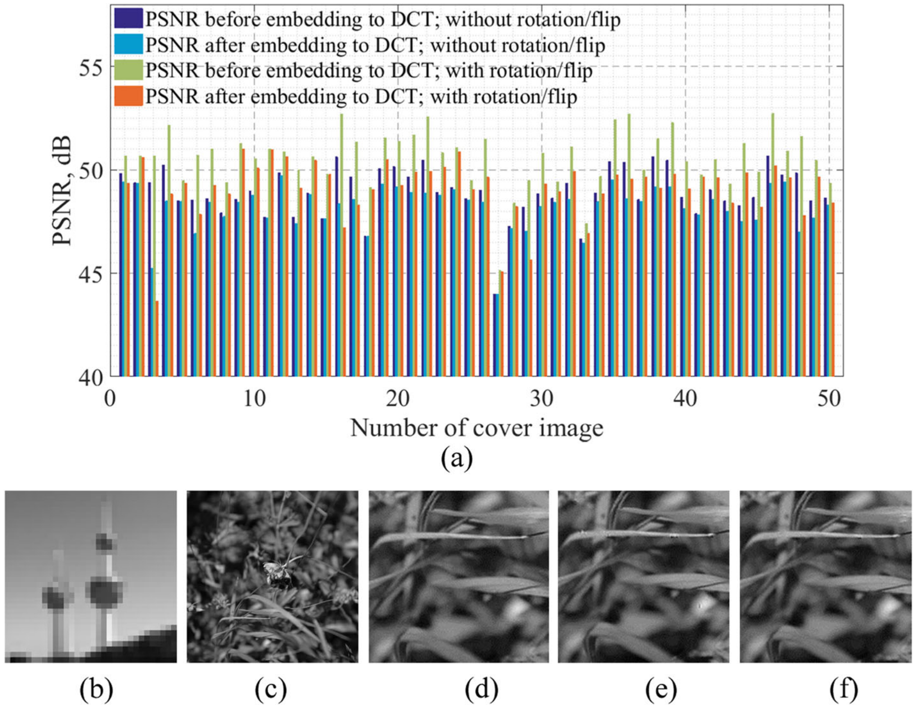

Fig. 4 a PSNR before and after embedding in the DCT domain for the cases with and without manipulations using image local blocks, (**b**) the secret image, 32 × 32, (**c**) an example of a cover image, (**d**) a part of the cover image and (**e**) the same part of the image after spatial embedding without local block manipulations, and (**f**) with local block manipulation

most cover images (92%), PSNR after DCT embedding with image block manipulation (red bars) was also higher than without (cyan bars). Visual analysis (comparing Fig. 4 e and f with Fig. 4d) also showed noticeable improvement. This result was used to propose the manipulations with local blocks as a stable improvement of secret image embedding using a combined technique.

The block size N_b also had a noticeable impact on the embedding quality. The visual quality of the stego-image after spatial embedding depended on the block size. In Eq. (16), $N_{\text{blocks}}^{\text{secret}}$ and $N_{\text{blocks}}^{\text{cover}}$ are dependent on N_b, so the amount of information to be encoded in the DCT domain is also determined by the block size. Similar to the case of image block manipulation, two factors have the opposite influence on the PSNR. Thus, an experiment, similar to the PSNR experiment above, was conducted with 100 cover images randomly collected from the BOSS database and a single secret image (Fig. 2a) of size 32 × 32. Blocks of 2 × 2, 4 × 4, and 8 × 8 were tested. Figure 5a demonstrates that the PSNR values after embedding decreased only in the spatial domain as N_b increased. However, after embedding in the DCT domain, the situation was not as obvious (Fig. 5b). Comparing the mean values and standard deviations (Fig. 5), the 2 × 2 size demonstrated the worst performance, as most information was encoded in the DCT domain. In the case of the 4 × 4 blocks, higher mean PSNR values were obtained along with a higher variance. For the 8 × 8 blocks, the variance was lower and the mean PSNR value was lower as well. The 4 × 4 blocks were used for further analysis in order to compare with [3].

Now, the actual embedding procedure will be described. The distance in Eq. (8) was calculated using different local block manipulations for all the blocks in the cover and secret images in search of the minimum. The coordinates of the local block for embedding were obtained along with the orientation of each block of the secret image. Spatial embedding was performed using these data, and the position and orientation of all the blocks were then coded in the DCT domain [1–3]. The number of bits required for coding the index of each block in the cover image was calculated as $\log_2(N_{\text{blocks}}^{cover})$. The rotation and flipping indices were grouped together in order to be described with the same number of bits (see the description in Eq. 16).

2.2.3 Similarity measure of local blocks

Local block similarity is calculated based on the distance between feature vectors. A commonly used metric is the Euclidean distance as stated in Eq. (17):

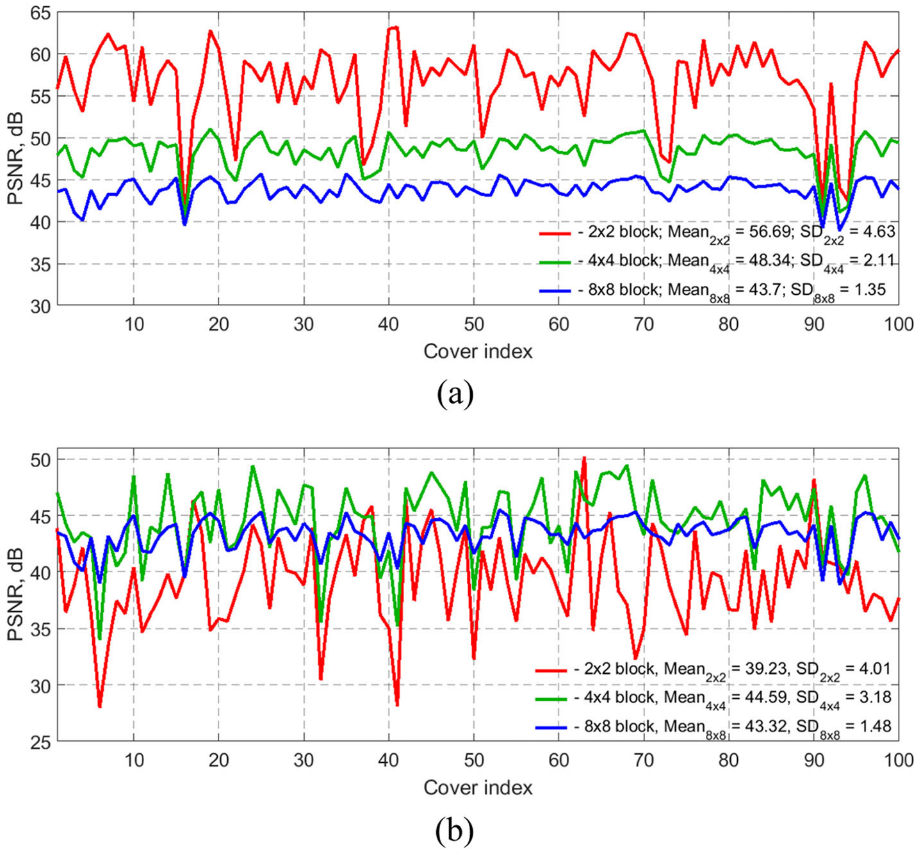

Fig. 5 a PSNR values for the differently sized local blocks before embedding in the DCT domain and (**b**) after embedding in the DCT domain

$$d_{ij} = \left[\sum_{l} (u_l - v_l)^2\right]^{1/2}, \qquad (17)$$

where i, j are the indices of the block in the cover and the secret image, u, v are the corresponding feature vectors, and l is the index of the feature vector component. In [3], the authors used statistical moments of subblocks to form the local block's feature vector. The most suitable container was chosen by the maximum number of most similar blocks. However, such an approach does not guarantee that all blocks of the secret image have similar blocks in the chosen cover image. Thus, the PSNR may become quite low in some situations. Instead of calculating the number of the most similar blocks, a novel distance metric is proposed as in Eq. (18):

$$\overline{d}_k = \frac{1}{N^{\text{secret}}_{\text{blocks}}} \sum_j d^j_{\min} \qquad (18)$$

where $d^j_{\min}$ is the minimal distance evaluated for the jth block of the secret image in the analyzed cover image k. Equation (18) calculates the distance to all the blocks of the secret image, and it provides more stable results.

The proposed measure in Eq. (18) directly minimizes the PSNR value for spatial embedding (compared with Eqs. (18) and (8)) but does not consider DCT embedding. To overcome this, the following algorithm was proposed. $d^j_{\min}$ provides the corresponding block index i of the current cover k. Having all indices, the bit sequence for embedding in the DCT domain can be formed. This will facilitate the embedding and calculate the mean MSE for all the DCT blocks used for embedding represented by $S^{-}{}_{k}$. The resulting complex measure is a multiplication of $d^{-}{}_{k}$ and $S^{-}{}_{k}$:

$$D_k = \overline{d}_k \overline{S}_k \qquad (19)$$

The effectiveness of the proposed measure was confirmed with an experiment. A set of 100 cover image candidates was used from the BOSS database along with one secret image of size 64 × 6. Cover selection was performed using the algorithm based on the complex measure Eq. (19) (Fig. 6a), using the algorithm based on the maximum number of most similar blocks [2, 3] (Fig. 6b), and using random selection (Fig. 6c). The results are presented in the form of distance Eq. (17) from all the blocks of the secret image. To provide a simple comparison,

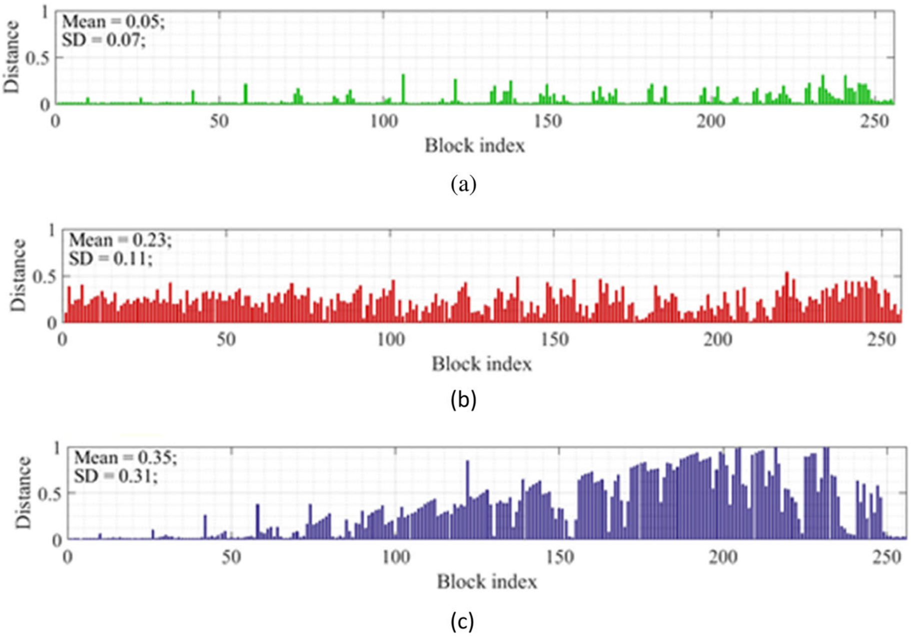

Fig. 6 a The distance between local blocks of the cover and the secret image for the complex measure, (**b**) the maximum number of most similar blocks, and (**c**) random selection used to choose the cover image

the distributions were normalized on the maximum distance determined from all three distributions (found for the randomly chosen cover, see Fig. 6c). Comparison of Fig. 6 a and b revealed that the proposed measure, based on the mean distance and the PSNR, provided a smoother distribution with fewer peaks.

2.2.4 Most suitable cover image (MSCI)

The above has been combined into a cover selection framework called the most suitable cover image (MSCI). According to its name, the framework adaptively (for a given secret image) picks the best cover image from an image database. Figure 7 illustrates the main components and the flow of MSCI. The database processing step is carried out only once for each database image in order to minimize the execution time. All necessary features (global and local) are extracted from all the cover candidates and saved in the feature database. Global features are represented by the entropy and the histogram (Section 2.2.1). The pixel intensities of each block are utilized as local features (Section 2.2.2).

When a secret image is embedded, the same features are extracted from it. Global filtering is accomplished by removing the images with negative values in the histogram and the entropy metrics (see Sections 2.2.3 and 2.2.4). Cover images that pass this step are referred further as cover candidates. The blocks are analyzed using the proposed intensity-based local features. The cover that provides the lowest complex measure Eq. (19) (Section 2.2.4) is chosen to be the most suitable cover image.

The computational complexity of the MSCI can be estimated as follows: local block matching involves $2N_b^2$ operations. The total amount of operations required for matching the given secret image with a particular cover container is determined as shown in Eq. (20):

$$\begin{aligned} N_{\text{operations}} &= \sum_k 2N_b^2 N_{\text{blocks}}^{\text{secret}} N_{\text{blocks}}^{k-th\ \text{cover}} \\ &= \frac{2N_{\text{pixels}}^{\text{secret}}}{N_b^2} \sum_k N_{\text{pixels}}^{k-th\ \text{cover}} \end{aligned} \qquad (20)$$

here, $2N_b^2$ is the number of operations required to match a single pair of blocks, and k is the number of covers. The total amount of operations required for matching the given secret image with a particular cover container is determined as shown in Eq. (21):

$$N_{\text{operations}} = \frac{2N_{\text{pixels}}^{\text{secret}} N_{\text{pixels}}^{\text{cover}} N_{\text{covers}}}{N_b^2} \qquad (21)$$

The number of operations required for local block matching is proportional to the number of cover images, N_{covers}. Therefore, the complexity of MSCI is $O(N_{\text{covers}} N_{\text{operations}})$= $O(N_{\text{covers}} N_{\text{pixels}}^{\text{secret}} N_{\text{pixels}}^{\text{cover}})$. The advantage of MSCI is that the complexity can be reduced by tuning the parameters in the global filtering step, i.e., the number of cover candidates is controlled. Thus, a balance between the algorithm performance and the visual quality and security of the resulting stego-images can be found.

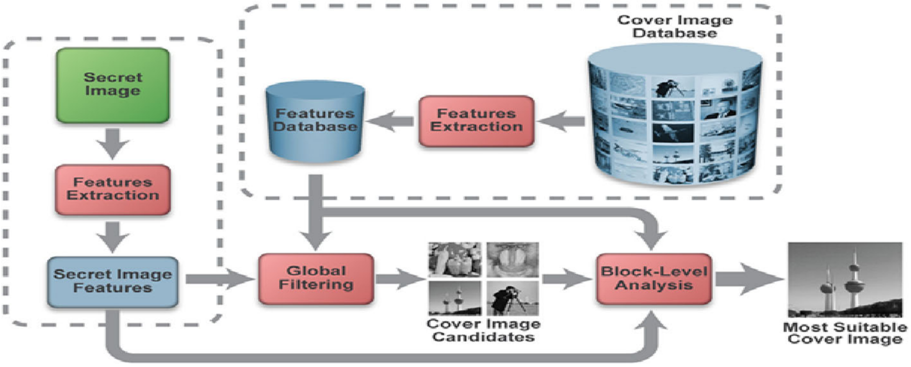

Fig. 7 The pipeline of the most suitable cover image (MSCI) framework

2.3 Related work

Cover image selection was first proposed in [2]. The cover image was divided into a sequence of local blocks of size 4 × 4. Sample mean and variance were calculated for each block. This operation was applied to images decomposed with Gabor filters of different scale and orientation. Thus, each local block was described via a 24-dimensional feature vector (two features for 12 Gabor images). For each block of the secret image, the closest block from a set of cover images was found (based on the Euclidean distance between feature vectors). The cover image with the maximum number of closest blocks was chosen as the most suitable container for a given secret image.

An improvement to this technique was proposed in [3]. The authors used a different approach for feature vector construction. Mean, variance, and skewness were calculated for 4 × 4 sub-blocks within each block. Mean pixel values from the local neighborhood were used to improve visual quality. As a result, the 16-dimensional feature vector (three features for four sub-blocks, and four features from the neighborhood) was constructed for each block. The cover selection procedure was similar to that described above. Using the information from the local neighborhood improved the visual quality of the stego-image.

An interesting analysis of the cover image embedding capacity was conducted in [20]. Here, the authors studied the resistance of different image containers to steganalysis attacks. Specifically, the relation between image complexity and embedding capacity was analyzed. As a result, the safe capacity was determined for each image container. Recent steganalysis algorithms utilize machine-learning approaches for the classification of cover and stego-images. Support vector machines (SVM) are the tool of choice for such problems [22]. However, in [23], it was illustrated that usage ensemble classifiers based on random forest results provided comparable performance with significantly lower training complexity [20].

In contrast to the methods introduced above, there is a group of algorithms based on cover image analysis, which do not consider a certain secret image. In [24], the authors introduced an agent-based system for image feature analysis. Images with the highest contrast and high entropy were chosen as the most suitable candidates. However, no steganalysis tests were conducted in this experiment. Steganography performance dependent on the global characteristics of the cover image was studied in [25]. A specific cover selection technique based on the analysis of the correlation coefficient within the cover image was proposed in [26]. Authors represented the cover image as a Gauss-Markov process and demonstrated that the images with smaller correlation coefficients led to lower detectability. But the embedding rate, unfortunately, was only considered to a maximum of one. The technique was also limited to spatial domain steganography.

A series of experiments on cover selection and steganalysis were performed in [27]. Steganographic security was analyzed for three different scenarios: when the cover image selector had no knowledge, partial knowledge, or full knowledge of the applied steganalyzer type and the classification principle. A set of different image quality metrics were tested together with different steganalysis tools. Proper use of the steganalyzer type decreased detection rate.

The steganalysis technique in [28] used the features based on intra- and inter-block DCT correlations while [29] used the fusion of DCT-based and Markov-based features. In [30], the authors obtained the receiver operating characteristics (ROC) to evaluate the classifier performance [30]. Furthermore, two recent steganalysis algorithms were proposed in [31, 32]. The authors in [31] proposed a new feature set for steganalysis for JPEG images with less complexity, less dimensionality, and better performance compared with other proposed JPEG domain steganalysis features. In [32], a steganalysis feature extraction technique based on 2D Gabor filters is offered for adaptive JPEG steganography. The evaluation of the detection performance of the proposed steganalysis feature can be enhanced effectively compared with other methods.

In [33], the authors proposed a cover selection method that is secure against both the single object steganalysis and pooled steganalysis at the same time. Additionally, a detailed explanation of steganography security in image level and individual level was showed while pointing on the theoretical weakness of existing cover-selection methods in individual level. The authors shortened the difference between the selected cover images and the whole set of possible images using the maximum mean discrepancy (MMD) distance during the cover selection to enhance the security in individual level. Experimental results demonstrated that the security in individual level and image level is assured.

Another interesting recent work describing text localization and web image understanding was presented in [34, 35], respectively. In [34], a novel solution was proposed to fast localize text in complex backgrounds which is considered as a challenge. This solution was based on two effective algorithms; stroke-specific FASTroke keypoint detector and the component similarity clustering algorithm. Performance results showed that this method outperformed the existing solutions and reported best performance as FASTroke generates less than twice the amount of components and at least 10% more characters are recognized. The problem of image understanding was presented and solved in [35] by proposing cross-modality bridging dictionary. Images were considered as probability distribution of semantic groups for the visual appearances. Moreover, the probability distributions were transferred for related categories by proposing knowledge-based semantic propagation. Experimental results showed the effectiveness of this method as it outperformed the state-of-the-art methods on four public datasets.

A new spatial-temporal attention model (STAT) was introduced in [36] in order to solve the problems of error recognition for the videos and missing the description details. STAT model was proposed within an encoder-decoder neural network for video captioning. It focuses on both the spatial and temporal structures in a video, and significant region was selected in the subset of frames instead of the subset only for accurate word prediction. Performance results showed that STAT generated detailed and accurate descriptions of videos. It also accomplished state-of-the-art performance on two well-known benchmarks.

3 Results and discussion

In this section, the visual quality and security of the stego-images produced using the developed framework are analyzed. Visual quality analysis is performed first. The PSNR was adopted as its numerical characteristic. The performance of MSCI was compared with the framework proposed in [3]. In steganalysis experiments, randomly chosen covers were used for both training and comparison purposes.

For the experiments, both algorithms were programmed in MATLAB. With computation and visualization services, it was a very convenient testing environment. The steganalysis techniques were also implemented as MATLAB scripts. A CPU Intel Core i5-7600K at 3.8 GHz with 16 Gb of RAM was used to perform the experiments.

3.1 Visual quality experiments

In this experiment, 1000 images from the BOSS database were used as a base for cover selection. The MSCI algorithm and the cover selection technique proposed in [3] were applied to find the best cover for three secret images of size 64×64 with different characteristics. Figure 8 shows the PSNR value after embedding the secret image into the 20 best cover images obtained using both techniques. The number of the cover in the diagrams corresponds to its number in the queue of the technique. The complex measure in Eq. (19) ensured that two or more cover image candidates were in the same position in the queue (have the same value of the measure). In the case of using the maximum number of similar blocks [2, 3], the situation with several candidates having exactly the same characteristic was quite possible, especially when $N_{\text{covers}} >> N^{\text{secret}}_{\text{blocks}}$.

The experiment proved that stego-images after MSCI demonstrated higher PSNR values than after cover selection based on statistical features [3] (Fig. 8). This is explained by the close relation between the distance measure Eq. (18) and the local MSE Eq. (8) (see Section 2.2.4).

3.2 Steganalysis experiments

To analyze the security of the developed cover selection framework, a series of steganalysis experiments were performed. Steganalysis tools were applied to detect the changes in the images that typically are not visible to HVS.

For the steganalysis experiments, the BOSS database (10000 images of size 512×512) was divided into two halves. The first half was used for training purposes (to choose the secret images to be embedded and to form a training set from the pairs of cover image and corresponding stego-image). Two types of features were used: features based on intra- and inter-block DCT correlations [28] and the fusion of DCT-based and Markov-based features [29]. The other half (5000 images) was used for testing. To evaluate the classifier performance, the receiver operating characteristics (ROC) were obtained [30].

The steganalysis techniques used to test the security of MSCI are described in [28, 29]. In addition, two recent steganalysis algorithms [31, 32] were implemented and

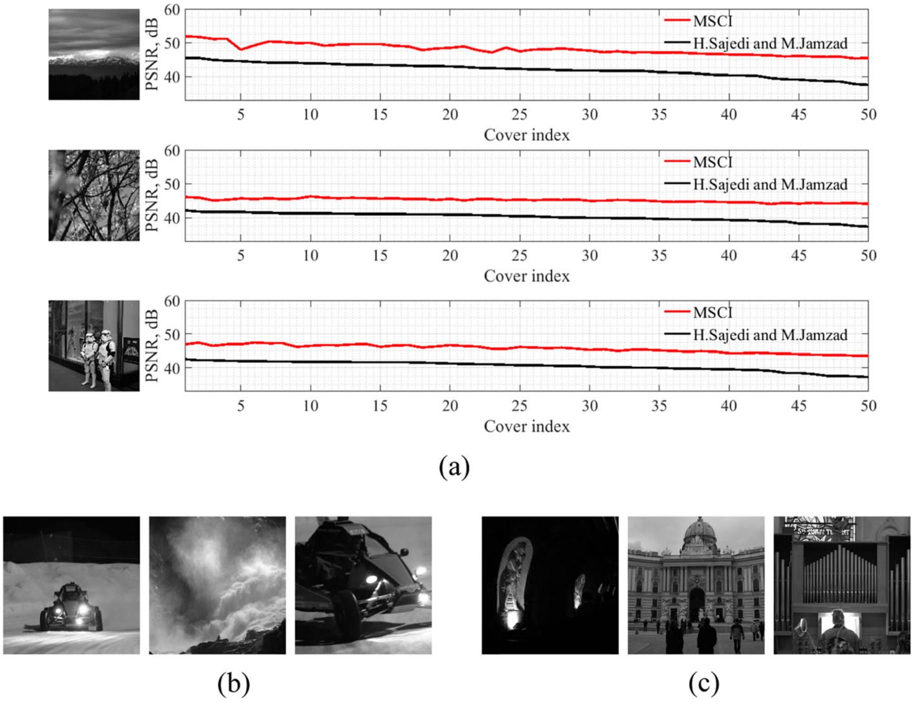

Fig. 8 a Comparing PSNR values of the stego-images obtained with two different cover selection and embedding techniques for three secret images, (**b**) the best covers chosen for the secret images with the MSCI framework, and (**c**) with the algorithm proposed in [3]

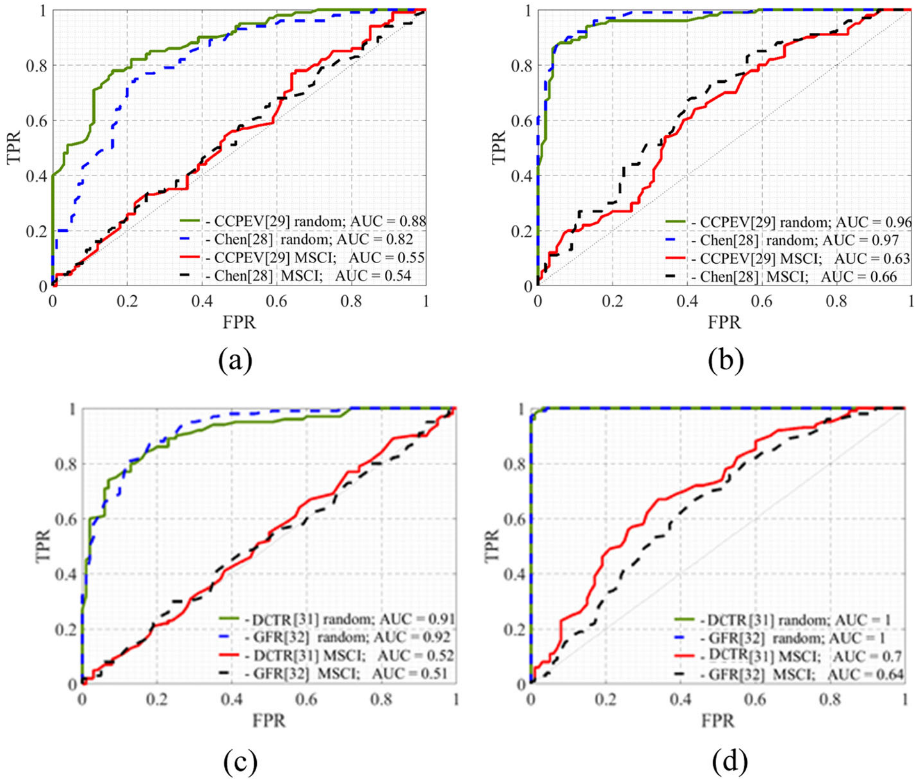

Fig. 9 ROC for four steganalysis techniques: the closer the ROC is to this line, the worse its stegnalyser performance. **a**, **b** CCPEVand Chen, **c**, **d** GFR and DCTR. The cover selection was performed with MSCI and random selection for different secret image sizes: (**a**, **c**): 32 × 32 and (**b**, **d**): 64 × 64

applied to both randomly chosen covers and image containers found by MSCI. Secret images of size 32 × 32 and 64 × 64 were used. The ROCs in Fig. 9 demonstrate the steganalysis results. The straight line between the points (0, 0) and (1, 1) in ROC space corresponds to random classification (the area under the curve (AUC) being equal to 0.5) [30]. Thus, the closer the ROC curve is to this line, the worse its steganalyser performance. In both cases, MSCI noticeably outperformed random cover selection. The ROC curve for the secret image of size 64 × 64 was closer to the point (0, 1) (Fig. 9b, d). The amount of embedded information in this case was larger. This affected the local characteristics of the stego-images; thus, the steganalyzers were more sensitive, and the secret information was detected with higher accuracy.

The ROC curve may be used to find the steganalyser parameters that provide the optimal classification accuracy. This is accomplished by maximizing Youden's index:

$$J = \mathrm{TPR} + \mathrm{TNR}{-}1 = \mathrm{TPR}{-}\mathrm{FPR} \tag{22}$$

while moving along the ROC curve [37]. Here, ***TPR*** is the true positives rate (or sensitivity), ***TNR*** is the true negatives rate (or specificity), and ***FPR*** is the false positive rate. The optimal classification accuracy is shown in Fig. 10 for the same steganalyzers and techniques used for cover selection. The accuracy increases with increasing size of secret image. To demonstrate this, a larger range of secret image size was analyzed. Starting from a size 64 × 64, the detection accuracy for random covers was larger than 0.8 (Fig. 10a), which indicated low security. MSCI always provided lower detection accuracy (Fig. 10b). The same was true for the steganalysis techniques seen in [31, 32] (Fig. 10c and d, respectively). For a secret image of size 32 × 32 and 48 × 48, the proposed method was very stable considering all four steganalysis techniques. This guaranteed secret communication without detection. The detection rate is reduced by 30% on average when using the MSCI as shown in Fig. 10b, d compared with the four steganalysis techniques [28–32] shown in Fig. 10a, c.

MSCI found a cover image that provided minimal visual distortion. As a result, the DCT coefficients and local image characteristics did not change significantly. This made the resulting stego-images more transparent for the steganalysis tools.

4 Conclusions and future work

In this paper, automatic cover image selection was examined and analyzed in detail. The use of secret image

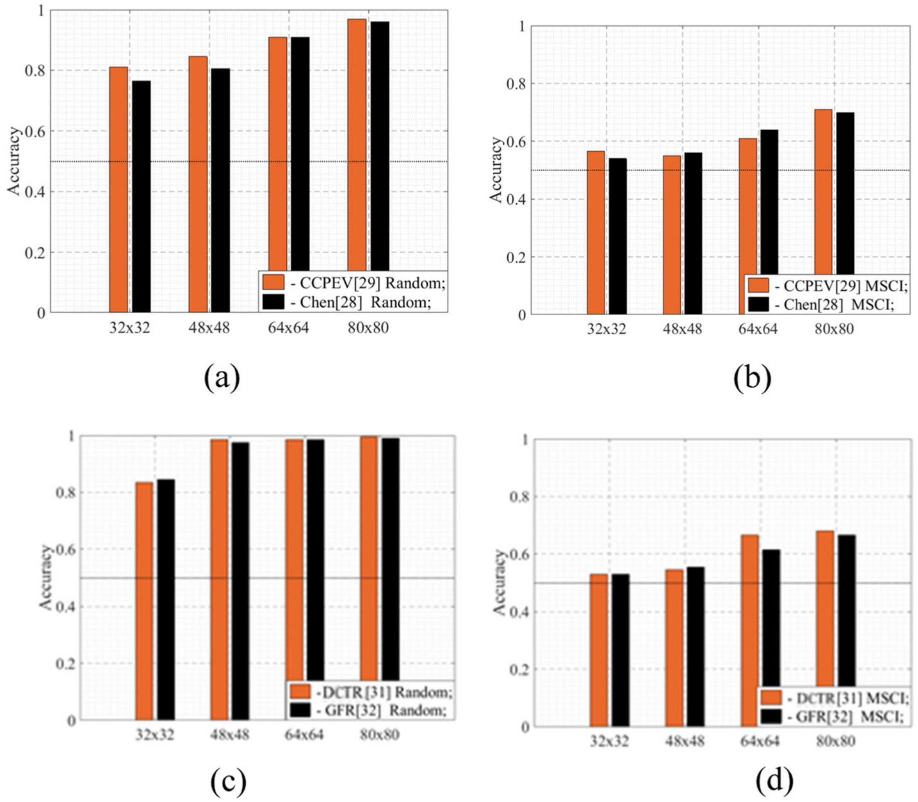

Fig. 10 **a**, **c** Detection accuracy for random cover selection and **b**, **d** MSCI choice; different steganalysis techniques were used: **a**, **b** CCPEV and Chen **c**, **d** GFR and DCTR

characteristics for global filtration significantly improved efficiency. The histogram and the entropy were presented as characteristics. An intensity-based local feature set was also proposed for local block matching. Its direct relation with such metrics as the PSNR and MSE was explained analytically, and its efficiency against the existing local block analysis techniques was demonstrated experimentally. In addition, analysis of the optimal block size and local block geometrical manipulations was introduced as mechanisms for further improvement of stego-image quality.

The experiments confirmed the improvements when using MSCI, both in terms of visual quality and security. The proposed method outperformed the existing methods in terms of visual quality. MSCI could be used in security applications, due to its resistance to modern steganalysis techniques. We believe that our work introduced a new way of looking at improving steganographic process as a whole instead of just focusing on the embedding part.

Further development of the MSCI algorithm is planned. In particular, more global image characteristics will be implemented as criteria for cover filtering suitability. Image distortions can be analyzed in both spatial and DCT domains. MSCI will also be combined with more advanced DCT domain steganography techniques.

Abbreviations

AUC: Area under the curve; DCT: Discrete cosine transform; HVS: Human visual system; LSB: Least-significant bit; MSCI: Most suitable cover image; MSE: Mean square error; PSNR: Peak signal-to-noise ratio; SVM: Support vector machines

Acknowledgements

Not applicable

Authors' contributions

SA proposed the methodology of the research, analyzed and interpreted the results and was a major contributor in writing the manuscript.
SAA did the literature review, was a major contributor in the experimental results and in writing the manuscript. MA revised the work and analysed the results. All authors read and approved the final manuscript.

Funding

The authors declare that they have no sources of funding.

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests.

Received: 15 May 2019 Accepted: 19 November 2019
[redacted]

References

1. S. Katzenbeisser, F.A. Petitcolas, *Information hiding techniques for steganography and digital watermarking* (Artech House, Inc., Norwood, USA, 2000)
2. Z. Kermani, M. Jamzad (2005) A robust steganography algorithm based on texture similarity using Gabor filter. Fifth IEEE International Symposium on Signal Processing and Information Technology, pp. 578–582.
3. H. Sajedi, M. Jamzad (2008) Cover selection steganography method based on similarity of image blocks. IEEE 8th International Conference on Computer and Information Technology Workshops, Sydney, Australia, pp. 379–384.
4. M. Mahajan, N. Kaur, Adaptive steganography: a survey of recent statistical aware steganography techniques. International Journal on Computer Network and Information Security **10**, 76–92 (2012)
5. K. Wu, W. Chung-Ming, Steganography using reversible texture synthesis. IEEE Transactions on Image Processing **24**(1), 130–139 (2015)
6. L. Guo, J. Ni, Y.Q. Shi, Uniform embedding for efficient JPEG steganography. IEEE Transactions on Information Forensics and Security **9**(5), 814–825 (2014)
7. L. Guo, J. Ni, W. Su, C. Tang, Y. Shi, Using statistical image model for JPEG steganography: Uniform embedding revisited. IEEE Transactions on Information Forensics and Security **10**(12), 2669–2680 (2015)
8. W. Luo, F. Huang, J. Huang, Edge adaptive image steganography based on LSB matching revisited. IEEE Transactions on Information Forensics and Security **5**(2), 201–214 (2010)
9. S. Islam, M. Modi, P. Gupta, Edge-based image steganography. EURASIP Journal on Information Security **8**, 1–14 (2014)
10. D.C. Lou, N.I. Wu, C.M. Wang, Z.H. Lin, C. S (2010) Tsai. A novel adaptive steganography based on local complexity and human vision sensitivity. Journal of Systems and Software 83(7): 1236-1248.
11. V. Sedighi, J. Fridrich (2016) Effect of saturated pixels on security of steganographic schemes for digital images. IEEE International Conference Image Processing (ICIP), pp. 2747-2751.
12. A. El Sayed, A. Elleithy, P. Thunga, Z. Wu (2015) Highly secure image steganography algorithm using curvelet transform and DCT encryption. 2015 Long Island Systems, Applications and technology (LISAT), pp. 15295289
13. Y. Zhang, X. Luo, C. Yang, D. Ye, F. Liu, *A JPEG-compression resistant adaptive steganography based on relative relationship between DCT Coefficients* (International Conference on Availability, Reliability and Security (ARES), 2015), pp. 461–466
14. T. Denemark, J. Fridrich (2017) Steganography with two JPEGS of the same scene. The 42nd IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1-5.
15. G. Liu, W. Liu, Y. Dai, S. Lian, Adaptive steganography based on block complexity and matrix embedding. Multimedia Systems **20**(2), 227–238 (2014)
16. S.-R. Tsui, C.-T. Huang, W.-J. Wang, A new adaptive steganographic method based on gradient adjacent prediction and side-match vector quantization. Journal of Information Hiding and Multimedia Signal Processing **4**(4), 215–224 (2013)
17. M. Jamzad, F. Yaghmaee, Achieving higher stability in watermarking according to image complexity. Sci. Iran. J. **13**(4), 404–412 (2006)
18. G. Bradsky, A. Kaehler, *Learning OpenCV* (O'Really Media Inc., Sebastopol, CA, 2008)
19. Y. K. Seong, Y.-H. Choi, T.-S. Choi (2004) Scene-based watermarking method for copy protection using image complexity and motion vector amplitude. In Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing, (ICASSP '04), 3: 409-412.
20. H. Sajedi, M. Jamzad, Secure steganography based on embedding capacity. International Journal of Information Security **8**(6), 433–445 (2009)
21. P. Bas, T. Filler, T. Pevny (2011) Break our steganographic system – the ins and outs of organizing BOSS. Proceedings of the 13th international conference on Information Hiding, LNCS, Springer Berlin Heidelberg, 6958: 59–70.
22. C. C. Chang, C. J. Lin. LIBSVM: A library for support vector machines. https://www.csie.ntu.edu.tw/~cjlin/libsvm/ (2001). .
23. J. Kodovsky, J. Fridrich, V. Holub, Ensemble classifiers for steganalysis of digital media. IEEE Transactions on Information Forensics and Security **7**(2), 432–444 (2012)
24. S.B. Sadkhan, A.M. Al-Barky, N.N. Muhammad, An agent based image steganography using information theoretic parameters. MASAUM Journal of Computing **1**(2), 258–264 (2009)
25. R. Yang, Z. Zheng, J. Wei, Cover selection for image steganography based on image characteristics. Journal of Optoelectronics Laser **25**(4), 764–768 (2014)
26. Y. Sun, F. Liu, Selecting cover for image steganography by correlation coefficient. Second International Workshop on Education Technology and Computer Science **2**, 159–162 (2010)
27. M. Kharrazi, H. Sencar, N. Memon (2006) Cover selection for steganographic embedding. IEEE International Conference on Image Processing, pp. 117–120.

28. C. Chen, Y. Shi (2008) JPEG image steganalysis utilizing both intrablock and interblock correlations. IEEE International Symposium on Circuits and Systems (ISCAS 2008), pp. 3029–3032.
29. T. Pevny, J. Fridrich, Merging Markov and DCT features for multi-class JPEG steganalysis. In proc. of SPIE. San Jose, CA **6505**, 1–13 (2007)
30. T. Fawcett, An introduction to ROC analysis. Pattern Recogn. Lett **27**(8), 861–874 (2006)
31. V. Holub, J. Fridrich, Low Complexity Features for JPEG Steganalysis Using Undecimated DCT. IEEE Transactions on Information forensics and security **10**(2), 219–228 (2015)
32. X. Song, F. Liu, C. Yang, X. Luo and Y. Zhang (2015) Steganalysis of Adaptive JPEG Steganography Using 2D Gabor Filters. Proceedings of the 3rd ACM Workshop on Information Hiding and Multimedia Security, ACM, pp. 15-23.
33. Z. Wang, X. Zhang, Secure cover selection for steganography. IEEE Access **7**, 57857–57867 (2019)
34. C. Yan, H. Xie, J. Chen, Z. Zha, X. Hao, Y. Zhang, Q. Dai, A fast Uyghur text detector for complex background images. IEEE Transactions on Multimedia **20**(12), 3389–3398 (2018)
35. C. Yan, L. Liang, C. Zhang, B. Liu, Y. Zhang, Q. Dai, Cross-modality bridging and knowledge transferring for image understanding. IEEE Transactions on Multimedia, 1–10 (2019)
36. C. Yan, T. Yunbin, X. Wang, Y. Zhang, X. Hao, Y. Zhang, Q. Dai, STAT: Spatial-temporal attention mechanism for video captioning. IEEE Transactions on Multimedia, 1–13 (2019)
37. E.F. Schisterman, N.J. Perkins, A. Liu, H. Bondell, Optimal cut-point and its corresponding Youden index to discriminate individuals using pooled blood samples. Epidemiology **16**, 73–87 (2005)

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.