

A Model-Based Image Steganography Method Using Watson's Visual Model

Mohammad Fakhredanesh, Reza Safabakhsh, and Mohammad Rahmati

This paper presents a model-based image steganography method based on Watson's visual model. Model-based steganography assumes a model for cover image statistics. This approach, however, has some weaknesses, including perceptual detectability. We propose to use Watson's visual model to improve perceptual undetectability of model-based steganography. The proposed method prevents visually perceptible changes during embedding. First, the maximum acceptable change in each discrete cosine transform coefficient is extracted based on Watson's visual model. Then, a model is fitted to a low-precision histogram of such coefficients and the message bits are encoded to this model. Finally, the encoded message bits are embedded in those coefficients whose maximum possible changes are visually imperceptible. Experimental results show that changes resulting from the proposed method are perceptually undetectable, whereas model-based steganography retains perceptually detectable changes. This perceptual undetectability is achieved while the perceptual quality — based on the structural similarity measure — and the security — based on two steganalysis methods — do not show any significant changes.

Keywords: Data hiding image steganography, model-based steganography human visual system, Watson's visual model, discrete cosine transform.

I. Introduction

Steganography involves hiding a secret message in a cover medium so that only the intended recipient becomes aware of the existence of the hidden message [1]. For higher security, the message is usually coded before being embedded into the cover medium. The result of this coding procedure is a pseudo-random sequence of 0's and 1's, which is then embedded in the cover. Embedding the message bits in an image usually introduces some visual artifacts and changes the image statistics. The image statistics are analyzed by steganalysis methods, whereas visual artifacts may be detected by the human eye.

Adaptive steganography utilizes image features and content information to improve the embedding process [2]. For example, less detectable artifacts may be introduced into the cover image by the embedding method [3]. In some cases these artifacts are detected with difficulty by steganalysis methods due to trivial changes introduced in image statistics. However, the artifacts may be visually perceptible to the human eye. Therefore, perceptual undetectability is very important in steganography. This means that applying a steganography method to an image should not leave any visually perceptible phenomena. To support this idea and improve robustness of watermarking methods, perceptual models, such as Watson's visual model [4], were proposed to describe the human visual system. The methods which use the human visual model in steganography are discussed in the next section.

Different approaches, such as statistics-preserving steganography, model-based steganography, and masking embedding as natural processing methods, exist for image steganography. In model-based steganography, a model is

Manuscript received Feb. 18, 2013; revised Oct. 21, 2013; accepted Nov. 5, 2013.

Mohammad Fakhredanesh (phone: +98 21 64542746, m-fakhredanesh@aut.ac.ir), Reza Safabakhsh (corresponding author, safa@aut.ac.ir), and Mohammad Rahmati (rahmati@aut.ac.ir) are with the Department of Computer Engineering and Information Technology, Amirkabir University of Technology, Tehran, Iran.

assumed for the cover image [5]. The embedding procedure changes image elements (pixel values or transform coefficients) so that they do not violate the assumed cover model. For example, Sallee [6] proposed a generalized Cauchy distribution (GCD) to model the distribution of each discrete cosine transform (DCT) coefficient and assumed that all covers obey this model. The idea of using a model for image statistics in his approach was a novel idea. But, unfortunately, steganalysis methods easily detect stego images generated by this method [7], [8].

While the model-based steganography approach attempts to model image statistics, visual models attempt to model the human visual system. We propose to use both visual and cover models for improving steganography techniques. In the proposed method, Watson's visual model [4] is used to strengthen model-based steganography. In this scheme, Watson's visual model prevents perceptual phenomena during the embedding phase. In addition, we model the DCT coefficients via the GCD.

First, the maximum acceptable changes are extracted based on Watson's model [4]. The DCT AC¹⁾ coefficients whose changes are acceptable based on Watson's visual model are called "proper AC coefficients." Then, the low-precision histogram of proper AC coefficients is computed and the distribution parameters are estimated. The message bits are encoded to obtain new symbols based on the model. Finally, new symbols are embedded in proper AC coefficients such that the estimated distribution is satisfied.

Experimental results show that the proposed steganography algorithm improves the model-based steganography algorithm [6] in terms of visual imperceptibility, image quality, and security.

The rest of the paper is organized as follows. In section II, we give a brief overview of the related work involving model-based steganography and human visual models. Watson's visual model is explained in section III. The details of the proposed steganography method are explained in section IV. In section V, using some experiments, we investigate the imperceptibility and quality of the resulting images and the security of the algorithm. Finally, in section VI we conclude the paper and outline some possible future works.

II. Related Work

The human visual system has been considered in JPEG image compression [9]. Based on Watson's study [4], this compression standard selects higher quantization coefficients

for frequencies that are less sensitive to the human eye. Ahumada and Peterson proposed a luminance-based model to approximate the perceptual sensitivity of DCT coefficients to quantization errors [10]. Peterson and others psychophysically measured the smallest coefficient of each frequency that causes a visible phenomenon [11]. These coefficients provided threshold amplitudes for the DCT basis functions. Watson called this model the image-independent perceptual (IIP) approach and proposed the image-dependent perceptual (IDP) method [4]. The IDP model is an extension of the IIP model, which is optimized for a specific image. It, as described in the next section, provides luminance masking, contrast masking, and error pooling. Podilchuk and Zeng used this IDP model for watermarking still images [12]. The IIP model, which is independent of image content, is not a suitable choice for data embedding. Adaptive data hiding implies using local image content during the embedding process. Kankanhalli and others proposed such a new analysis method of the noise sensitivity for every pixel [13]. Awrangjeb and others used the method of Kankanhalli and others [13] for their data-hiding method in the spatial domain [14], since this method is applicable to this domain.

Cox and others dedicated one chapter of their book to perceptual watermarking [5]. They noted that a perceptual model generally considers sensitivity, masking, and pooling phenomena in a perceptive sense, such as vision. Watson's visual model [4] properly modeled such phenomena in the DCT domain, while Kankanhalli and others [13] limited their model to edge, texture, smooth areas, and brightness sensitivity. Therefore, we chose Watson's model to apply a human visual model in the DCT domain.

Human visual models usually measure the distance between the original image and the watermarked images. This distance can be used as a restriction on changes of coefficients and improves imperceptibility [15]–[19] and robustness of watermarking [20]–[24]. In addition, it is used as a measure for determining suitable regions in an image in which to embed more data [25], [26], a measure of distortion [26], and to simplify the computation of watermark insertion [27].

In some studies, the human visual system has been used with a neural network [28], [29] or a fuzzy system [30] to recover the watermarking signals from the watermarked images.

Watson's visual model [4] is used to model the human visual system in the DCT domain. However, in some investigations it is modeled in other transform domains [20]. Levicky and Foris combined three human visual models in the DCT, discrete wavelet transform (DWT), and spatial frequency domains [31]. Xie and others compared two human visual models: the DCT-based Watson's model and the DWT-based pixel-wise masking model and showed that Watson's model was the more superior.

1) Alternating Current. In DCT, the upper-left-most element in the coefficients matrix of each block is the DC or direct current coefficient and other coefficients are the AC coefficients.

[32]. Pan and others analyzed the influences of luminance, frequency, and texture masking factor to cover distortion for embedding in the DWT domain [33]. Shu and others proposed two watermarking methods for the discrete curvelet and discrete contourlet transform domains [34], [35]. They concluded that these domains have higher capacity than the wavelet domain. Kim and others proposed a watermarking method based on the human visual system in the DWT domain, which is robust against JPEG compression, smoothing, cropping, collusion, and multiple watermarking [36].

III. Watson's Visual Model

A perceptual model usually measures the perceptibility of a hidden message. The human visual system is the basis for visual models such as Watson's model [4], [5]. Watson's model attempts to account for frequency sensitivity, luminance masking, and contrast masking; all of which are in the DCT domain.

1. Frequency Sensitivity

Watson's visual model defines the frequency sensitivity in a matrix t , in which each $t[i, j]: i, j = 1, \dots, 8$, is to be the smallest visible value of the corresponding DCT coefficients in a block without any masking noise. Smaller values of $t[i, j]$ indicate higher sensitivity of the human eye to the corresponding frequency [4], [5]. The resulting matrix is

$$t = \begin{bmatrix} 1.40 & 1.01 & 1.16 & 1.66 & 2.40 & 3.43 & 4.79 & 6.56 \\ 1.01 & 1.45 & 1.32 & 1.52 & 2.00 & 2.71 & 3.67 & 4.93 \\ 1.16 & 1.32 & 2.24 & 2.59 & 2.98 & 3.64 & 4.60 & 5.88 \\ 1.66 & 1.52 & 2.59 & 3.77 & 4.55 & 5.30 & 6.28 & 7.60 \\ 2.40 & 2.00 & 2.98 & 4.55 & 6.15 & 7.46 & 8.71 & 10.17 \\ 3.43 & 2.71 & 3.64 & 5.30 & 7.46 & 9.62 & 11.58 & 13.51 \\ 4.79 & 3.67 & 4.60 & 6.28 & 8.71 & 11.58 & 14.50 & 17.29 \\ 6.56 & 4.93 & 5.88 & 7.60 & 10.17 & 13.51 & 17.29 & 21.15 \end{bmatrix}, \quad (1)$$

where lower frequencies (upper-left corner of the matrix) have smaller t -values. The high sensitivity of the human eye to these frequencies results in detecting perceptible phenomena in the image, despite the triviality of the introduced changes.

2. Luminance Masking

The human eye sensitivity also depends on the average intensity of the block. This means that in brighter blocks, larger changes in DCT coefficients can be imperceptible. Watson's model modifies matrix $t[i, j]$ in each block k according to the block's average luminance or DC term. The luminance

masking threshold $t_L[i, j, k]$ is defined as

$$t_L[i, j, k] = t[i, j](C_o[0, 0, k] / C_{0,0})^{a_T}, \quad (2)$$

where $C_o[0, 0, k]$ is the DC coefficient of the k th block of the cover image, $C_{0,0}$ is the average of DC coefficients in all cover image blocks, and a_T is the constant 0.649 [4], [5].

3. Contrast Masking

The reduction of visibility of a change in a frequency due to its energy is called contrast masking. The contrast masking threshold $s[i, j, k]$ is defined as

$$s[i, j, k] = \max\{t_L[i, j, k], |C_c[i, j, k]|^{w[i, j]} t_L[i, j, k]^{1-w[i, j]}\}, \quad (3)$$

where $w[i, j]$ is a constant between 0 and 1 (The value 0.7 has been used by Watson for all i, j [4], [5]).

4. Pooling

A level of distortion perceptible for half of the experimental trails is called a just-noticeable difference (JND) [5]. Perceptual distance is defined to allow comparison of a cover image C_c and a stego image C_s . The perceptual distance $d[i, j, k]$ between the (i, j) th DCT coefficients of the k th block in both images is defined as

$$d[i, j, k] = \frac{C_s[i, j, k] - C_c[i, j, k]}{s[i, j, k]}, \quad (4)$$

which measures the distance between them as a fraction or multiple of one JND. The perceptual distances are pooled into a single perceptual distance $D_{wat}(C_c, C_s)$, given by

$$D_{wat}(C_c, C_s) = \left(\sum_{i,j,k} |d[i, j, k]|^p \right)^{\frac{1}{p}}, \quad (5)$$

where Watson proposed a value of four for p . The value of p determines type or degree of pooling. For example if $p=1$, the pooling is a linear summation of absolute values, but if $p=2$, the pooling is a quadratic combination of errors [4], [5].

IV. The Proposed Method

In this section, we propose a steganography method that uses both a cover model and Watson's visual model [4] for steganography. The cover modelling assumes a model for the cover image statistics. Watson's visual model perceptually considers differences between the cover and stego images.

1. Model-Based Steganography in JPEG Images

The proposed method uses a model-based approach for

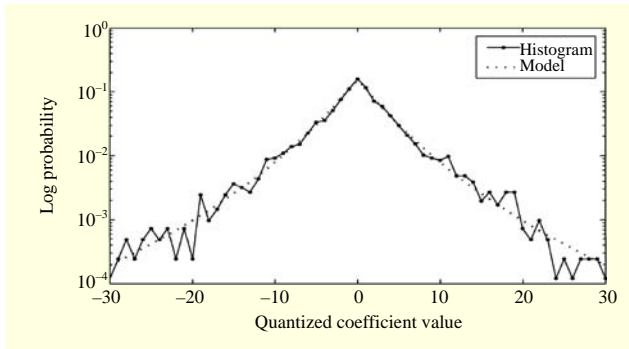


Fig. 1. Histogram of DCT coefficient (2, 2) for goldhill image and model pdf with parameters $s=18.28$ and $p=6.92$ [6].

embedding message bits in the DCT coefficients of JPEG images, based on the method proposed by Sallee [6]. Let x denote an image that is an instance of a random variable X and therefore probability has distribution $P_X(x)$. Then, x can be separated into two distinct parts: x_α , which is not changed, and x_β , which may be changed to x'_β during embedding. For example, if we use naive least significant bit embedding, x_α represents the seven most significant bits and x_β represents the least significant bit of the binary representation of the selected coefficients.

The main idea of this approach is treating x_α and x_β as instances of two dependent random variables X_α and X_β , respectively and estimating the conditional distribution $\hat{P}_{X_\beta|X_\alpha}(X_\beta | X_\alpha = x_\alpha)$ based on the proposed model distribution \hat{P}_X . In this approach, if x'_β is selected to obey this conditional distribution, the resulting $x' = (x_\alpha, x'_\beta)$ will be correctly distributed according to model \hat{P}_X . Sallee [6] proposed to use a parametric model of P_X to estimate $\hat{P}_{X_\beta|X_\alpha}$ used to select x'_β such that it conveys the intended message and is also distributed according to the estimate of $\hat{P}_{X_\beta|X_\alpha}$ [6].

This approach skips zero-value DCT coefficients due to the large number of these coefficients and sensitivity of the image to their changes [37], [38]. In addition, no bit is embedded into the DC coefficients, since changing these coefficients results in more detectable statistical and visual artifacts [5], [37]. The remaining AC coefficients are modeled using a specific form of a GCD as a parametric density function, which is defined as

$$P(u) = \frac{p-1}{2s} \left(1 + \left|\frac{u}{s}\right|\right)^{-p}, \quad (6)$$

where u is the coefficient value and $p > 1$, $s > 0$ are the two parameters of the model. This function is a better fit than other parametric models to the coefficient distributions [6]. Figure 1 depicts the histogram of the (2,2) DCT coefficient for a sample image and the model density function fitted using the maximum-likelihood approach. The corresponding cumulative

density function is

$$D(u) = \begin{cases} \frac{1}{2} \left(1 + \left|\frac{u}{s}\right|\right)^{1-p} & \text{if } u \leq 0, \\ 1 - \frac{1}{2} \left(1 + \left|\frac{u}{s}\right|\right)^{1-p} & \text{if } u > 0. \end{cases} \quad (7)$$

2. Data-Embedding Algorithm

The pseudocode of the data-embedding algorithm is given in Fig. 2. In the first step of the algorithm, after reading a cover image and generating a message, the DCT of the image is computed as done in the JPEG standard. For each coefficient x , x_β consists of the k least significant bits and x_α the $(8 - k)$ most significant bits of the coefficient. Changing coefficient values only changes x_β , which results in the coefficient bin offset. Therefore, the maximum change in the value of coefficients is $step - 1$. This change may visually be perceptible to the human eye. We use Watson's visual model [4] to prevent perceptible changes. Using this model, for each coefficient of each block, if the maximum change ($step - 1$) is visually imperceptible, then the embedding is valid. These coefficients are called "proper AC coefficients," and the rest of the procedure is applied only to them.

The pseudocode for the proposed method is given in Fig. 2, and the method for selecting such coefficients is presented in detail in Fig. 3. In this method, after computing the DCT of the image, the frequency sensitivity matrix t is defined. Then, the luminance masking and contrast masking thresholds are computed for each block. Then for each coefficient, the maximum possible change, which is equal to $(step - 1)$, is

```

1: function WMB(step)
2:   c ← ReadCoverImage()
3:   msg ← GetRandMessage()
4:   ACw ← ACCoefWatsonModel(c, step)
5:   for each AC coefficient i
6:     Xwlhi ← LowPrecHist(ACwi, step)
7:     (pi, si) ← FitGCD(Xwlhi, step)
8:     cdfi ← ComputeCDF(pi, si)
9:     symPi ← SymbolProbability(cdfi)
10:  end for
11:  order ← RandPermutation()
12:  newSym ← ArithEncode(msg, symP, order)
13:  newXw ← NewCoefficients(newSym)
14:  s ← WriteStegoImage(newXw)
15: end function

```

Fig. 2. Pseudocode of proposed algorithm.

```

1: function ACw ← ACCoefWatsonModel(c,step)
2:   dctC ← DCT(c)
3:   t ← DefineT()
4:   for each DCT block k
5:     tlk ← ComputeLuminanceMask(dctC)
6:     slk ← ComputeContrastMask(dctC)
7:     for each AC coefficient i
8:       frqSensk,i ← IsLower(step, ti)
9:       lumMaskk,i ← IsLower(step, tlk,i)
10:      contMaskk,i ← IsLower(step, slk,i)
11:     end for
12:   end for
13:   ACw ← frqSens AND lumMask AND contMask
14: return ACw
15: end function

```

Fig. 3. Pseudocode for selecting proper coefficients.

compared with the corresponding thresholds of frequency sensitivity, luminance masking, and contrast masking through the *IsLower* procedure. If the maximum possible change is lower than or equal to each threshold t_i , $tl_{k,i}$ or $sl_{k,i}$, this procedure returns the true value; otherwise it returns the false value. As shown in line 13 of Fig. 3, all coefficients whose maximum possible changes are lower than or equal to the corresponding thresholds of frequency sensitivity, luminance masking and contrast masking are returned as proper coefficients for embedding.

In the next steps, some procedures are applied to the proper coefficients of each type of AC coefficients. The *LowPrecHist* procedure in Fig. 2 constructs a low-precision histogram for the x_α 's. In this histogram, all coefficient values x with equal x_α but different x_β are counted for each bin. The term "low-precision histogram" is used to indicate the counting of the coefficient values x , which have equal most-significant bits, but unequal or equal least-significant bits in the same bin.

The bin size, which is called the *step*, is defined as the number of consecutive coefficient values that are assigned to a bin; and is equal to 2^k , where k shows the number of least significant bits that comprise x_β . Each coefficient value is represented by a histogram-bin index and a symbol that indicates its offset within the bin from 0 to 2^k-1 . The bin indices for all the coefficients comprise x_α , which will remain unchanged, and the bin offsets comprise x_β , which may be changed during the embedding procedure.

After computing all low-precision histograms, the model parameters p and s , of (6), are fitted to these histograms individually by *FitGCD*. The steganography decoder does not know the least significant portions x_β 's, which may have been changed by the encoder. During embedding, the proper AC coefficients are changed only within these low-precision histogram bins (only the bin offsets are changed) so that the

same estimates of p and s for each coefficient type may be obtained by the decoder.

Once the model parameters p_i and s_i are fitted to the histograms, the procedure *ComputeCDF* computes the cumulative distribution function cdf_i based on the GCD model parameters p_i and s_i . Then, the *SymbolProbability* procedure computes the probability of each possible offset symbol for a coefficient given its bin index, which is then assigned to the variable $symP_i$. These offset symbols, and their respective probabilities, are passed to a non-adaptive arithmetic entropy decoder, along with the secret message. The details of arithmetic encoding can be found in [39]. The order in which coefficients are used for encoding the message by *ArithEncode* is previously determined by the *RandPermutation* procedure. This procedure, which is called permutative straddling [37], computes a pseudo-random permutation to avoid changing coefficients only in a part of the image. This straddling procedure was proposed in the F5 algorithm [37] to shuffle all coefficients using a permutation. This permutation is an order for embedding and is dependent upon a key. The offset symbols returned by the entropy decoder comprise x'_β 's, which are combined with the bin indices (x_α 's) by the *NewCoefficients* procedure. This combination forms the new coefficient value x' , which is called *newXw* in Fig. 2.

We summarize the method in the following steganography algorithm.

The Proposed Algorithm:

Input: cover image I , embedded message m , *step*.

Output: stego image I_{cs} .

Step 1: Compute the DCT of image I to generate the DCT coefficients X_w .

Step 2: Compute proper AC coefficients whose changes are visually imperceptible based on Watson's visual model and the value of *step*. A coefficient change is visually imperceptible if it is less than or equal to the corresponding thresholds of the frequency sensitivity matrix (t), the luminance masking matrix (tl), and the contrast masking matrix (sl).

Step 3: For each AC coefficient i , do steps 3.1 to 3.4.

Step 3.1: Compute the low-precision histogram of proper AC coefficients i .

Step 3.2: Compute p_i and s_i by fitting a GCD model to this low-precision histogram.

Step 3.3: Compute the cumulative distribution function cdf_i based on the GCD model parameters p_i and s_i .

Step 3.4: Compute the symbol probability of cdf_i and generate $symP_i$.

Step 4: Generate *order* by a random permutation.

Step 5: Generate new symbols *newSym* by arithmetic encoding of the message based on *symP* and *order*.

Step 6: Generate new DCT coefficients $newXw$ using new symbols $newSym$.

Step 7: Generate and return stego image I_{cs} by applying inverse DCT on DCT coefficients $newXw$.

V. Experimental Results

In the following discussion, we use the term “model-based” for the steganography method proposed by Sallee [6] and use the “proposed method” for our method. The model-based steganography method embeds message bits so that the DCT changes may or may not be perceptible to the human eye. In this work, we restrict changes of DCT coefficients to be visually imperceptible changes.

The algorithm was implemented based on the model-based algorithm implementation using MATLAB.

We also used a JPEG toolbox developed for manipulating DCT coefficients of the JPEG images. In the following experiments, the bin size or *step* is set equal to 2.

Figure 4 depicts an application of the model-based and proposed methods on the goldhill image. The original cover (goldhill) image is shown in Fig. 4(a). Stego images generated by the model-based and the proposed methods are depicted in Figs. 4(b) and 4(c), respectively. In both cases *step* = 2. Figs. 4(d) and 4(e) depict perceptible changes for the model-based and proposed methods, respectively. The white dots correspond to the DCT coefficients whose changes during the embedding procedure are perceptible to human eyes. They are obtained based on Watson’s visual model of the human visual system in the DCT domain. As Fig. 4 shows, the model-based method results in perceptible changes, but the proposed method does not retain any perceptible results. Perceptually undetectable changes in the proposed method are guaranteed. We used a large data set of cover images called BossBase [40]. This data set contains ten thousand 512 pixels × 512 pixels gray-scale images, which are originally in raw PGM²⁾ format. These images were converted to JPEG format and then the model-based and the proposed method were applied to them. In all experiments *step* = 2. The proposed method did not retain any perceptible changes based on Watson’s visual model while the model-based method retained some visually perceptible changes.

In another quality test, we examined the proposed method based on the structural similarity (SSIM) index. SSIM measures structural similarity, which assesses the perceptual quality of the image. Wang and others [41] defined SSIM between two images x and y as

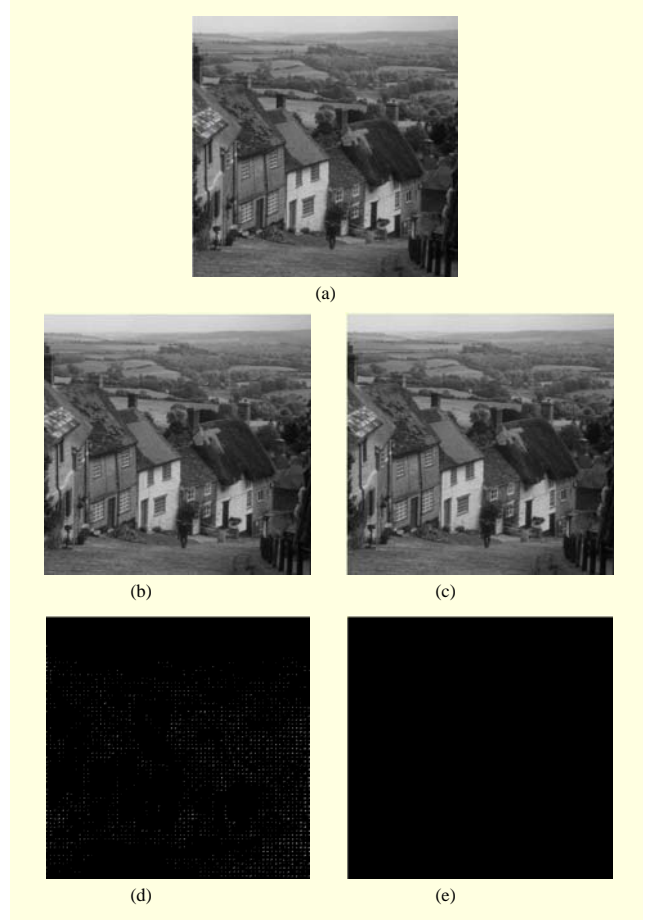


Fig. 4. Perceptible changes in DCT coefficients based on Watson’s visual model: (a) goldhill cover image, (b) stego image generated by model-based steganography, (c) stego image generated by proposed steganography, (d) white dots are perceptible changes in DCT coefficients after model-based steganography, and (e) no white noise (no perceptible change) resulted after the proposed steganography.

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (8)$$

where μ_x and σ_x^2 are, respectively, the mean and variance of image x and are defined as

$$\mu_x = \frac{1}{N} \sum_{i=1}^N x_i, \quad (9)$$

$$\sigma_x^2 = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2. \quad (10)$$

The variables μ_y and σ_y^2 are similarly defined. In addition, σ_{xy} is defined as

$$\sigma_{xy} = \frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)(y_i - \mu_y), \quad (11)$$

²⁾ Portable Gray Map

Table 1. Average, maximum, minimum, and standard deviation of the SSIM vs. payload. Stego images are produced by model-based and proposed methods for different payloads.

Method	Statistics	Payload (bpac)				
		0.01	0.02	0.03	0.04	0.05
Model-based	Average	0.99997742	0.99995518	0.99993287	0.99991101	0.99988911
	Maximum	0.99999957	1	1	1	1
	Minimum	0.99949459	0.99895124	0.99853924	0.99793887	0.99739842
	Standard deviation	0.00002419	0.00004775	0.00007145	0.00009513	0.00011845
Proposed	Average	0.99997607	0.99995249	0.99992900	0.99990575	0.99988268
	Maximum	0.99999946	1	0.99999849	1	1
	Minimum	0.99949405	0.99900100	0.99842562	0.99794140	0.99730463
	Standard deviation	0.00002509	0.00004927	0.00007405	0.00004927	0.00012218

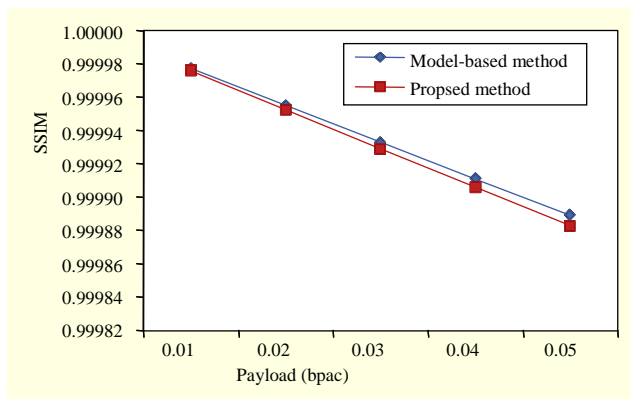


Fig. 5. SSIM vs. payload. Stego images are the result of model-based and proposed method for different payloads.

and C_1 and C_2 are two small constants included to avoid instability. We set both of them equal to 0.001 in our experiments.

For different payloads, the SSIM indexes of all images are computed for each method and their average values are reported in Fig. 5. In our experiments, the payloads are measured based on the “bpac” unit, which stands for “bit per non-zero AC coefficients.” In this measure, the payload is specified as the ratio of the message size to the number of non-zero AC coefficients of all DCT coefficients. This experiment shows that the stego images resulting from the two methods have similar qualities. The maximum difference between the corresponding values in the two methods is less than 10^{-5} of their values. Therefore, considering only the SSIM measure, it can be said that the proposed method has no significant effect on the images compared to the model-based method. Other statistics—as in the maximum, minimum, and standard deviation are also reported for this experiment in Table 1. Similar results for these statistics are reported for both methods.

For example, a comparison of the minimum SSIM’s demonstrates that although our method has higher minimum SSIM values for payloads of 0.02 bpac and 0.04 bpac, it has lower values for payloads of 0.01 bpac, 0.03 bpac, and 0.05 bpac.

Figure 6 demonstrates a sample cover image and its corresponding stego image generated by the model-based and proposed method with a payload of 0.03 bpac. The SSIM values reported for these images show that the stego image generated by the proposed method has slightly better SSIM than the SSIM of the other stego image.

Witnessing better results for the model-based method in some cases is due to the nature of the SSIM. This criterion gives structural similarity and not content similarity. SSIM measures the difference between two images based on the means and variances of the images and is not sensitive to the places where changes appear. On the other hand, Watson’s model considers the difference between the two images, as well as the place where the change appears. If the change appears in a place that is sensitive to the human eye, the change is important; otherwise, it is not.

Although the main contribution of the proposed method is human visual imperceptibility, we tested the security of the proposed method in the other experiments with two steganalysis attacks.

Both steganalysis methods consist of two parts: feature extraction and classification. In the first steganalysis method, the extracted features include 274 merged extended DCT and Markov features (as introduced in [42],) which are calibrated either by difference [42] or by the Cartesian product (as introduced in [43]), thus, yielding a total of 548 features. For classification, we applied the ensemble classifiers described by Kodovský and others [44] to the extracted features. We call this

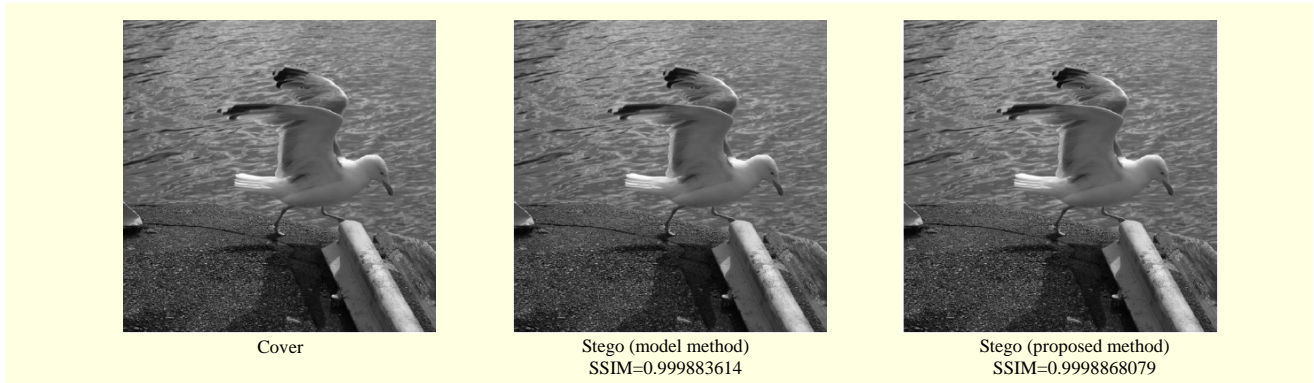


Fig. 6. SSIM for a sample cover image and its corresponding stego image generated by model-based and proposed methods with a payload of 0.03 bpac.

Table 2. Detection rates of steganalysis attack [38]–[40] for different DCT-based steganography methods.

Steganalysis method	DCT-Mkv-548-ensCls		Higher-order statistics		
	Payload (bpac)	Model-based (%)	Proposed (%)	Model-based (%)	Proposed (%)
	0.01	58.42	58.83	49.95	50.07
	0.02	66.58	67.23	49.97	49.95
	0.03	73.71	74.60	49.95	49.98
	0.04	80.08	81.04	50.13	49.91
	0.05	85.04	86.09	50.04	49.99
	0.1	97.20	97.43	50.01	50.03
	0.2	99.76	99.78	50.06	50.11
	0.3	99.95	99.95	50.15	50.15
	0.4	99.98	99.97	50.32	50.35
	0.5	99.98	99.97	50.50	50.42

method the “DCT-Mkv-548-ensCls”. In the second steganalysis method, we tested our method by the higher-order statistics steganalysis method [45], [46]. This method uses a wavelet-like decomposition to build higher-order statistical models of natural images. Then a Fisher linear discriminant analysis is used for the classification.

We applied DCT-Mkv-548-ensCls and higher-order statistics method to the model-based method and our proposed methods. Table 2 compares the detection rates of both methods for different payloads. We define the detection rate of an algorithm as the average of the correct recognition rate of the cover images and the correct recognition rate of the stego images produced by the algorithm. Obviously, lower detection rates are desirable for a steganalyzer. It shows that both methods have similar securities. In the DCT-Mkv-548-ensCls attack, the model-based method has better results for lower payloads, but our method remains more secure than the model-based method when increasing the payloads. At a payload of 0.3 bpac, the

two methods have the same detection rates. For payloads of more than 0.3 bpac, our method is more secure than the model-based method. In the higher-order statistics attack, both methods have about the same results. The results show that neither the model-based nor the proposed method can be detected by the higher-order statistics attack, because their detection rates are near 50% (random detection).

VI. Conclusion and Future Work

Model-based steganography is an important approach for image steganography. Although this approach is based on a novel idea, experimental steganalysis methods easily detect stego images generated with it. Watson’s visual model is a model of the human visual system in the DCT domain. Several studies in data hiding have used the human visual model in which stego images do not show any perceptible changes.

In this paper, we combined these two approaches to improve

model-based steganography. We proposed utilizing Watson's visual model for model-based steganography, which improves the model-based steganography method in terms of visual imperceptibility.

Experimental results show that the proposed method does not retain any perceptible change in the image while the model-based method retains many perceptible changes in the stego images. This is the purpose and the main contribution of the present work. We tested our method with the SSIM perceptual quality measure, which showed that the proposed method does not leave significant effects on images as compared to the model-based method. Furthermore, we tested our method by two steganalysis attacks called DCT-Mkv-548-ensCls and higher-order statistics. The results showed that the two methods have similar securities.

Determination of the optimum changes based on the perceptual and statistical undetectability and extending the proposed method to other transform domains, such as wavelet and contourlet transform constitute the ideas for future works.

References

- [1] L.M. Marvel, C.G. Boncelet, and C.T. Retter, "Spread Spectrum Image Steganography," *IEEE Trans. Image Process.*, vol. 8, no. 8, Aug. 1999, pp. 1075–1083.
- [2] A. Cheddad et al., "Digital Image Steganography: Survey and Analysis of Current Methods," *Signal Process.*, vol. 90, no. 3, Mar. 2010, pp. 727–752.
- [3] J. Fridrich and R. Du, "Secure Steganographic Methods for Palette Images," *Proc. IH*, Dresden, Germany, Sept. 29 – Oct. 1, 1999, pp. 47–60.
- [4] A.B. Watson, "DCT Quantization Matrices Visually Optimized for Individual Images," *Proc. SPIE*, San Jose, CA, USA, vol. 1913, Jan. 31, 1993, pp. 202–216.
- [5] I.J. Cox et al., *Digital Watermarking and Steganography*, 2nd ed., Burlington: Morgan Kaufmann, 2008.
- [6] P. Sallee, "Model-Based Steganography," *Digit. Watermarking*, LNCS, Seoul, Rep. of Korea, vol. 2939, Oct. 20–22, 2004, pp. 154–167.
- [7] J. Fridrich, "Feature-Based Steganalysis for JPEG Images and its Implications for Future Design of Steganographic Schemes," *Proc. IH*, LNCS, Toronto, Canada, vol. 3200, May 23–25, 2005, pp. 67–81.
- [8] R. Bohme and A. Westfeld, "Breaking Cauchy Model-Based JPEG Steganography with First Order Statistics," *Proc. ESORICS*, LNCS, vol. 3193, 2004, pp. 125–140.
- [9] G. Wallace, "The JPEG Still Picture Compression Standard," *IEEE Trans. Consum. Electron.*, vol. 38, no. 1, Feb. 1992, pp. 18–34.
- [10] A.J. Ahumada and H.A. Peterson, "Luminance-Model-Based DCT Quantization for Color Image Compression," *Proc. SPIE, Human Vis., Visual Process. Digital Display III*, San Jose, CA, USA, vol. 1666, Aug. 27, 1992, pp. 365–374.
- [11] H.A. Peterson et al., "Quantization of Color Image Components in the DCT Domain," *Proc. SPIE, Human Vis., Visual Process., Digital Display II*, San Jose, CA, USA vol. 1453, June 1, 1991, pp. 210–222.
- [12] C.I. Podilchuk and W. Zeng, "Perceptual Watermarking of Still Images," *Proc. IEEE Multimedia Signal Process.*, Princeton, NJ, USA, June 23–25, 1997, pp. 363–368.
- [13] M.S. Kankanhalli, Rajmohan, and K.R. Ramakrishnan, "Content Based Watermarking of Images," *Proc. MULTIMEDIA*, Bristol, UK, Sept. 1998, pp. 61–70.
- [14] M. Awrangjeb and M. Kankanhalli, "Reversible Watermarking Using a Perceptual Model," *J. Electron. Imag.*, vol. 14, no. 1, Mar. 30, 2005, pp. 1–8.
- [15] J.F. Delaigle, C. de Vleeschouwer, and B. Macq, "Watermarking Algorithm Based on a Human Visual Model," *Signal Process.*, vol. 66, no. 3, May 28, 1998, pp. 319–335.
- [16] M. Awrangjeb and M.S. Kankanhalli, "Lossless Watermarking Considering the Human Visual System," *Digit. Watermarking*, LNCS, Seoul, Rep. of Korea, vol. 2939, Oct. 20–22, 2003, pp. 581–592.
- [17] G. Zhu and N. Sang, "An Adaptive Quantitative Information Hiding Algorithm Based on DCT Domain of New Visual Model," *Proc. ISISE*, Shanghai, China, vol. 1, Dec. 20–22, 2008, pp. 546–550.
- [18] N. Ahmidi and A.L. Neyestanak, "A Human Visual Model for Steganography," *Proc. CCECE*, Niagara Falls, ON, USA, May 4–7, 2008, pp. 1077–1080.
- [19] R. Hu, F. Chen, and H. Yu, "Incorporating Watson's Perceptual Model into Patchwork Watermarking for Digital Images," *Proc. ICIP*, Hong Kong, China, Sept. 26–29, 2010, pp. 3705–3708.
- [20] T.H. Chen, G. Horng, and S.H. Wang, "A Robust Wavelet-Based Watermarking Scheme Using Quantization and Human Visual System Model," *Pakistan J. Inf. Technol.*, vol. 2, no. 3, 2003, pp. 213–230.
- [21] Y.J. Jung, M. Hahn, and Y.M. Ro, "Spatial Frequency Band Division in Human Visual System Based Watermarking," *Proc. IWDW*, Seoul, Rep. of Korea, vol. 2613, Nov. 21–22, 2003, pp. 224–234.
- [22] J. Porter and P. Rajan, "Image Adaptive Watermarking Techniques Using Models of the Human Visual System," *Proc. SSST*, Cookeville, TN, USA, Mar. 5–7, 2006, pp. 354–357.
- [23] M. Jayalakshmi, S.N. Merchant, and U.B. Desai, "Significant Pixel Watermarking Using Human Visual System Model in Wavelet Domain," *Proc. ICVGIP*, Madurai, India, Dec. 13–16, 2006, pp. 206–215.
- [24] O.H. Kwon, Y.S. Kim, and R.H. Park, "Watermarking for Still Images Using the Human Visual System in the DCT Domain,"

Proc. ISCAS, Orlando, FL, USA, vol. 4, July 1999, pp. 76–79.

- [25] S.W. Jung, L.T. Ha, and S.J. Ko, “A New Histogram Modification Based Reversible Data Hiding Algorithm Considering the Human Visual System,” *IEEE Signal Process. Lett.*, vol. 18, no. 2, Feb. 2011, pp. 95–98.
- [26] X. Zhang and S. Wang, “Steganography Using Multiple-base Notational System and Human Vision Sensitivity,” *IEEE Signal Process. Lett.*, vol. 12, no. 1, Jan. 2005, pp. 67–70.
- [27] Y. Li et al., “An Adaptive Blind Watermarking Algorithm Based on DCT and Modified Watson’s Visual Model,” *Proc. ISECS*, Guangzhou, China, Aug. 3–5, 2008, pp. 904–907.
- [28] D.C. Lou, J.L. Liu, and M.C. Hu, “Adaptive Digital Watermarking Using Neural Network Technique,” *Proc. IEEE, ICCST*, Oct. 14–16, 2003, pp. 325–332.
- [29] Y. Zhang, “Blind Watermark Algorithm Based on HVS, and RBF Neural Network in DWT Domain,” *WSEAS*, Stevens Point, WI, USA, vol. 8, no. 1, Jan. 2009, pp. 174–183.
- [30] S. Oueslati, A. Cherif, and B. Solaiman, “A Fuzzy Watermarking Approach Based on the Human Visual System,” *Int. J. Image Process.*, vol. 4, no. 3, 2010, pp. 218–231.
- [31] D. Levicky and P. Foris, “Human Visual System Models in Digital Image Watermarking,” *Radioeng.*, vol. 13, no. 4, Dec. 2004, pp. 38–43.
- [32] G. Xie, M. Swamy, and M.O. Ahmad, “Perceptual-Shaping Comparison of DWT-Based Pixel-Wise Masking Model with DCT-Based Watson Model,” *Proc. IEEE, ICIP*, Atlanta, GA, USA, Oct. 8–11, 2006, pp. 1381–1384.
- [33] F. Pan et al., “Steganography Based on Minimizing Embedding Impact Function and HVS,” *Proc. ICECC*, Zhejiang, China, Sept. 9–11, 2011, pp. 490–493.
- [34] Z. Shu et al., “Watermarking Algorithm Based on Contourlet Transform and Human Visual Model,” *Proc. ICES*, Sichuan, China, July 29–31, 2008, pp. 348–352.
- [35] Z. Shu et al., “Watermarking Algorithm Based on Curvelet Transform and Human Visual Model,” *Proc. ISECS*, Nanchang, China, vol. 1, May 22–24, 2009, pp. 208–212.
- [36] Y.S. Kim, O.H. Kwon, and R.H. Park, “Wavelet Based Watermarking Method for Digital Images Using the Human Visual System,” *Proc. IEEE, ISCAS*, Orlando, FL, USA, vol. 4, July 1999, pp. 80–83.
- [37] A. Westfeld, “F5-A Steganographic Algorithm,” *Proc. IH*, Pittsburgh, PA, USA, Apr. 25–27, 2001, pp. 289–302.
- [38] N. Provos, “Defending Against Statistical Steganalysis,” *Proc. USENIX Security Symp.*, Washington, WA, USA, Aug. 13–17, 2001, pp. 323–335.
- [39] T. Cover and J. Thomas, *Elements of Information Theory*, NY: Wiley, New York, 1991.
- [40] “BOSSBase (v1.01).” Accessed Jan. 2013. <http://exile.felk.cvut.cz/boss/BOSSFinal/index.php?mode=VIEW&tmpl=materials>
- [41] Z. Wang et al., “Image Quality Assessment: From Error Visibility

to Structural Similarity,” *IEEE Trans. Image Process.*, vol. 13, no. 4, Apr. 2004, pp. 600–612.

- [42] T. Pevny and J. Fridrich, “Merging Markov and DCT Features for Multi-class JPEG Steganalysis,” *Proc. SPIE*, San Jose, CA, USA, vol. 6505, Jan. 28, 2007.
- [43] J. Kodovsky and J. Fridrich, “Calibration Revisited,” *Proc. ACM MM & Sec.*, Princeton, NJ, USA, Sept. 7–8, 2009, pp. 63–74.
- [44] J. Kodovsky, J. Fridrich, and V. Holub, “Ensemble Classifiers for Steganalysis of Digital Media,” *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 2, Apr. 2012, pp. 432–444.
- [45] H. Farid, “Detecting Hidden Messages Using Higher-Order Statistical Models,” *Proc. ICIP*, Rochester, NY, USA, vol. 2, Sept. 22–25, 2002, pp. 905–908.
- [46] Accessed Nov. 2012. www.cs.dartmouth.edu/farid/downloads/research/steg.m



Mohammad Fakhredanesh received his BS and MS degrees in computer science from the Mathematical and Computer Science Department, Amirkabir University of Technology (Tehran Polytechnic), Iran, in 2005 and 2007, respectively. He is currently a PhD candidate at the Computer Engineering Department, Amirkabir University of Technology (Tehran Polytechnic), Iran. His research interests are in the fields of data hiding, image processing, computer vision, and pattern recognition.



Reza Safabakhsh received his BS degree in electrical engineering from the Sharif University of Technology, Tehran, Iran, in 1976 and his MS and PhD degrees in electrical and computer engineering from the University of Tennessee, Knoxville, USA, in 1980 and 1986, respectively. He worked at the Center of Excellence in Information Systems, Nashville, TN, USA, from 1986 to 1988. Since 1988, he has been with the Computer Engineering Department, Amirkabir University of Technology, Tehran, Iran, where he is currently a professor and a director of the Computer Vision Laboratory. His current research interests include computer vision, neural networks, and information hiding. He is a member of the IEEE and several honor societies, including Phi Kappa Phi and Eta Kappa Nu. He was the founder and a member of the Board of Executives of the Computer Society of Iran and was the president of this society for the first four years.



Mohammad Rahmati received his MS in electrical engineering from the University of New Orleans, USA, in 1997 and his PhD degree in electrical and computer engineering from the University of Kentucky, Lexington, KY, USA, in 2003. He is currently an associate professor at the Computer Engineering

Department, Amirkabir University of Technology (Tehran Polytechnic), Iran. His research interests are in the fields of pattern recognition, image processing, bioinformatics, video processing, and data mining. He is the research coordinator of the department and is a member of the IEEE Signal Processing Society.