Statistical Attack Resilient Data Hiding

Palak K. Amin, Ning Liu, and K. P. Subbalakshmi
(Corresponding author: K. P. Subbalakshmi)

ECE Department, Stevens Institute of Technology
Burchard 208, Hoboken, NJ 07030, USA. (Email: ksubbala@stevens.edu)

(Received Nov. 20, 2005; revised and accepted Dec. 31, 2005 & Feb. 15, 2006)

Abstract

In this paper, we propose a discrete cosine transform (DCT) based spread spectrum data-hiding algorithm that is resilient to statistical attacks. Unlike other spread spectrum based data-hiding algorithms, the proposed algorithm does not introduce a low-pass filtering effect in the histogram of the stego image. The distance between the center of gravity (CoG) as defined by [14] of the unmarked host and the stego images was reduced by 74% in the proposed algorithm. The proposed algorithm is also resilient against the Chi-Square attack and does not compromise on robustness or capacity to achieve this goal. When compared to the generic block based DCT data-hiding scheme, the proposed algorithm provides a 41% reduction in the relative entropy between the host and stego images. In other words, the proposed algorithm is 41% more secure than generic DCT based data-hiding algorithms when measured in terms of relative entropy. The proposed algorithm also provides statistical resilience against a steganalysis attack specifically designed for block DCT data-hiding algorithms [29]. The proposed algorithm is robust against a variety of image manipulating attacks such as noise addition, filtering, blurring, sharpening, JPEG compression etc. In the cases of dislocating attacks such as blurring and despeckling, the bit error ratio (BER) was 0.1045 and 0.0435 respectively, thereby yielding retrieval rates of over 89% and 95% respectively. In the case of low quality JPEG attack (Q-30) the retrieval rate was 92%. In the case of noise addition attacks, the retrieval rates were more than 92%.

Keywords: Data-hiding, information security, multimedia security, steganography

1 Introduction

Data hiding is the science of hiding one data type within another. Depending upon the purpose of data-hiding, it can be classified into broad categories: steganography, where the main application is covert communication; and watermarking where the application is often related to the integrity of the host data itself. Some applications of watermarking include copyright protection, copy protection, authentication, etc. Data-hiding techniques can also be classified into two categories: perceptible and imperceptible. For obvious reasons, imperceptible techniques are considered more suitable for all of the applications above. So, for the remainder of this paper, we will only consider imperceptible data-hiding. Based on the processing domain, steganography can be classified into two categories: spatial domain based [20, 21] and transform domain based [22, 26, 28, 34]. A very popular class of transform based data hiding techniques is spread spectrum data hiding [22, 26, 28, 34]. In the sections that follow we propose a spread spectrum based data hiding technique that provides resilience against statistical attacks.

Just as data hiding is classified into steganography and watermarking, based on the application, the attacks on the system also can be classified into two categories based on whether the system is watermarking or steganographic. In the former case, the attacks usually include geometric manipulations of the image (rotation, scaling, translation, cropping etc. [12, 19, 23]), compression, noise addition etc. The aim of these attacks is to destroy the synchronization between the watermark extractor and the embedder. Attacks on steganographic systems are referred to as steganalysis attacks [11, 16, 32]. Most popular of these are: visual attacks [32], histogram analysis [11], dual statistics methods [25], JPEG compatibility steganalysis [10, 16, 30], universal blind steganalysis [11], unique fingerprinting [7], etc.

Most data-hiding schemes embed data sequentially or in some pseudo-random fashion. If the host image contains connected areas or uniform colors, then a simple visual inspection of a suspected stego image can reveal the existence of the sequentially embedded hidden data. However, it may be harder to distinguish noisy images or highly textured images from stego images using this technique. Hence, Westfeld et al. [32] introduced a new statistical attack that can be applied to any data-hiding technique in which a fixed set of pairs of values (PoVs) are flipped in order to embed message bits. Example Povs can be pixel values, quantized image coefficients, palette indices etc. Prior to embedding, these values are unevenly distributed, but after embedding their occurrences in each

pair tends to be equal. Hence a Chi-Square test on such a data-hiding scheme reveals the existence of the hidden data [11, 24].

Other statistical detection methods that start with sample counts, neglect the placement of pixels in the stego image [25]. It is obvious that by using the spatial correlations between parts of the stego image, we could detect the hidden data more reliably and accurately. However, to uncover and possibly quantify the weak relationship between the host image and some pseudo random data is not simple. Another means of detecting steganography is steganalysis via JPEG compression compatibility [10, 16, 30]. If a host is originally stored in the JPEG format [31], it would be possible to recover the JPEG quantization table even after data-hiding. In most cases, the stego-image becomes incompatible with the JPEG decoder when embedding is done in the compressed domain. Hence, the presence of hidden data is confirmed.

All of the steganalysis algorithms discussed so far are tuned to a specific embedding algorithm. Universal blind detection is a meta-detection method, which attempts to do away with the above restriction. The idea is to find an appropriate set of sensitive statistical qualities that possess the characteristics required to allow the steganalyzer to detect hidden data. Farid [7] proposed a set of sensitive higher-order features derived using the wavelet decomposition of the stego-image. These feature vectors are then divided into two linear subspaces, using the Fisher linear discrimination analysis [8]. This analysis gives the distance between the image under test and its low-pass filtered versions, which can allow the steganalyzer to detect hidden data.

In this paper we propose a novel spread spectrum data-hiding algorithm, in the two-dimensional discrete cosine transform (2D-DCT) domain that complies with the following requirements:

1) The algorithm must reduce the chances of statistical detection.

2) The algorithm must provide robustness against a variety of image manipulation attacks.

3) The stego image must not have any distortion artifacts.

4) The algorithm must not sacrifice embedding capacity in order to achieve the above requirements.

The rest of the paper is organized as follows. In Section 2 we describe the working of the proposed algorithm and provide details of an example encoder-decoder pair. In Section 3, we comment on the visual and statistical distortion of the proposed algorithm and discuss the performance of several stegnalyzers on the proposed algorithm. In Section 4, we provide experimental results for robustness tests. In Section 5, we end with observations and concluding remarks.

2 The Proposed Data-Hiding Algorithm

Transform based data-hiding techniques have been well studied [1, 2, 6, 9, 15, 21, 35]. Unlike spatial techniques, these provide the user with an option to embed the hidden data in certain regions of the image that are less sensitive to minor distortions and at the same time are important enough to survive general lossy compression. In this paper, we use the block-based two-dimensional discrete cosine transform (2D-DCT) for embedding the data. The DCT coefficients of digital images have interesting properties that could be exploited to obtain good hidden data imperceptibility and fidelity. Most important property of the DCT is energy compaction. It has been shown that the DCT converges to the Karhunen-Loeve transform (KLT) for certain images that could be statistically modelled as first-order Markov processes [3]. The DCT is an invertible transform, and its inverse is given by the equation below:

$$A_{mn} = \sum_{p=0}^{M-1} \sum_{q=0}^{N-1} \alpha_p \alpha_q B_{pq} cos(\frac{\pi(2m+1)p}{2M}) cos(\frac{\pi(2n+1)q}{2N}),$$

where, $0 < m < M-1$ and $0 \leq n \leq N-1$. Here α_p and α_q are defined as,

$$\alpha_p = \begin{cases} 1/\sqrt{(M)} & \text{if } p = 0 \\ \sqrt{(2/M)} & \text{if } 1 \leq p \leq M-1 \end{cases}$$

$$\alpha_q = \begin{cases} 1/\sqrt{(N)} & \text{if } q = 0 \\ \sqrt{(2/N)} & \text{if } 1 \leq q \leq N-1. \end{cases}$$

The inverse DCT equation can be interpreted as meaning that any M-by-N matrix A can be written as a sum of MN functions of the form,

$$\alpha_p \alpha_q B_{pq} \cos(\frac{\pi(2m+1)p}{2M}) \cos(\frac{\pi(2n+1)q}{2N}), \qquad (1)$$

where, $0 < m < M-1$ and $0 \leq n \leq N-1$.

It has been showed that the samples of the DCT coefficients at the same frequency, in different blocks, are statistically independent [4]. The statistical characterization of the DCT coefficients of the original image is valuable information to the decoder; however, since the original image is unknown to the decoder, this information has to be estimated from the stego image. Given the small alterations generated within the host image due to the data-hiding process and considering that the hidden data can be modelled statistically [18], very good estimates of these distribution parameters can be obtained.

2.1 The Encoder

In the proposed algorithm, 8x8 pixel blocks of the image are first transformed using the DCT and the mid-frequency regions of the DCT coefficient blocks (see Figure 1) are used to embed the hidden data. In this figure,

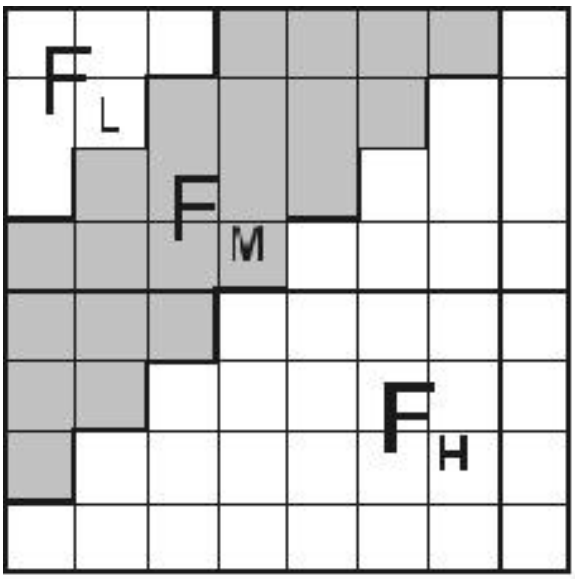


Figure 1: Mid-frequency DCT coefficients that are being used to embed the hidden data

F_L, F_M and F_H represent the low, medium and high frequency regions respectively. At the encoder end, we use a random number generator using a secret key and generate n pseudo random noise (PN) sequences, $\{W_m\}_{m=1}^{n}$, where n is the length of the hidden data sequence and W_m is the m^{th} PN sequence. Hence we have one PN-sequence corresponding to each hidden bit. If the m^{th} bit of the hidden data sequence is '0', we modulate the mid-band DCT coefficients of the m^{th} block with W_m using Equation (2) and similarly for the case where the hidden bit is a '1', we use Equation (3).

$$\tilde{I}_m(u,v) = \begin{cases} I_m(u,v) + kW_m, & \text{if } u,v \in F_M \\ I_m(u,v), & \text{otherwise} \end{cases} \quad (2)$$

$$\tilde{I}_m(u,v) = \begin{cases} I_m(u,v) - kW_m, & \text{if } u,v \in F_M \\ I_m(u,v), & \text{otherwise.} \end{cases} \quad (3)$$

Here, k is the gain factor used to specify the strength of the embedded data. Typically the values k can assume range between 0.0 and 1.0, where 1.0 corresponds to 100% data hiding strength. Increasing k hence reduces the chance of detection errors at the expense of additional image degradation [16]. W_m is the appropriate pseudo random noise sequence, based on the hidden bit, m. $I_m(u,v)$ represents the m^{th} 8x8 DCT block of the host image and $\tilde{I}_m(u,v)$ represents the corresponding marked DCT block. Each block, $\tilde{I}_m(u,v)$, is then inverse transformed to give the final stego image $\tilde{I}(x,y)$. We note that the visual distortion induced in the stego image due to embedding is very low. Figure 2 shows the original image and the hidden data on the left hand side and the stego image on the right hand side.

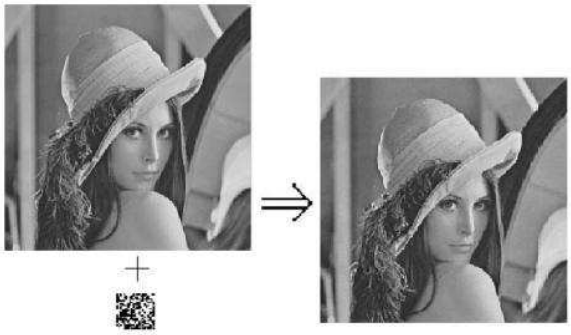


Figure 2: The original image and the hidden data on the left hand side and the stego image on the right hand side

2.2 The Decoder

Since the proposed detection is blind, the decoder has no access to the host image. However, considering the minor distortion generated by the data-hiding process, and the fact that the hidden data can be modelled statistically, a good estimate of the distribution parameters can be obtained from the stego image itself. Let $\beta \triangleq b_1, b_2, ..., b_M$ be the message vectors that correspond to the M possible hidden messages. Let $w_l = P_b l, l \in 1, 2, ..., M$ be the l^{th} PN sequence, whose individual elements may be denoted by $(wl_1, wl_2, ..., wl_n)$. Assuming that all the codewords $b_1, b_2, ..., b_m$ have the same a priori probability, the decoder must minimize the error probability conditioned to a secret key and extract the message vector b_l that satisfies,

$$\ln(\frac{f_z(z \mid b_l)}{f_z(z \mid b_m)}) = \ln(\frac{f_x(z - w_l)}{f_x(z - w_m)}) > 0, \forall m \neq l. \quad (4)$$

With the assumption that the elements of x are independent, and using Equations (1) and (2), we can equate Equation (5) as follows.

$$\sum_{i=1}^{L}(\frac{\mid Z_i - w_{m,i} \mid^{ci} - \mid Z_i - w_{l,i} \mid^{ci}}{\sigma_i^{ci}}) > 0, \forall m \neq l. \quad (5)$$

Assuming that the message vectors b_l satisfy $i \in -1, 1, \forall l \in 1, 2, ..., M, i \in 1, 2, ..., N$, Equation (6) provides sufficient statistics for the hidden message recovery procedure.

$$r_i \triangleq \sum_{k \in S_i}(\frac{\mid Z_k + \alpha_k S_k \mid^{ck} - \mid Z_k - \alpha_k S_k \mid^{ck}}{\sigma_k^{ck}}). \quad (6)$$

If the stego image has not suffered any alteration, then $\forall k \in s, z_k = x_k + b_i \alpha_k s_k$. After using this result in Equation (4) and treating the s_i and the sets S_i as the only random elements of the system, without the loss of generality we can define the first and second order moments.

$$r_i = \sum_{k \in S_i}(\frac{\mid x_k + 2\alpha_k S_k \mid^{ck} - \mid x_k \mid^{ck}}{\sigma_k^{ck}}).$$

In the definition of the sufficient statistics r_i we note that they can be expressed as a sum of statistically independent terms. Hence for a finite N, we can apply the central limit theorem [17, 18] and approximate the distribution of r via a Gaussian *pdf*. Hence we define the signal to noise ratio (SNR) as follows,

$$SNR \triangleq (\frac{E^2[r_i]}{\sigma^2(r_i)}).$$

Lastly under the assumption that the hidden message vectors form a binary constellation with $M = 2^N$ points, the theory of Gaussian approximation suggests that using a bit-by-bit hard decoder the probability of a bit in error is.

$$P_b = Q(\sqrt{SNR}).$$

In the proposed decoder algorithm, the stego image $\tilde{I}(x, y)$ is broken down into 8x8 blocks, and a 2D-DCT transformation is performed. Then the correlation between the mid-band DCT coefficients $\tilde{I}_m(u, v)$ and the appropriate m^{th} PN-sequence W_m is calculated, where W_m is normalized to zero mean. The detection process measures the correlation between the sequences generated using the mid-band DCT coefficients from the stego image and the PN-sequences, which were generated by the same key that was used at the encoder. These correlation values are stored in an array of size n, with each correlation value corresponding to one hidden bit. If the m^{th} correlation value exceeds the average of all of the n correlation values, then we extract a '0′ bit. Else we extract a '1′ bit as seen in Equation (7).

$$H_m = \begin{cases} 0, & \text{if } correlation(m) > average(corr) \\ 1, & \text{otherwise.} \end{cases} \quad (7)$$

3 Steganalyzing the Proposed Algorithm

To discuss the visual and statistical distortion caused by embedding, we use the standard secret key stegosystem [27]. Alice and Bob are the two users of the stegosystem. An adversary, Eve, is also present within the system. Eve has access to the stego image, the embedding algorithm, and a perfect read only access to the channel. Figure 3 shows the model of the stegosystem in detail.

- In the first pass (switch position 0), Alice transmits only the covertext C to Bob over the public channel. No hidden messages are transmitted to Bob in first pass. Eve observes C, but she is unsure of C being the unmarked covertext or marked stego text.
- In the second pass (switch position 1), Alice is active and transmits stegotext S to Bob, generated by using the embedding function F on the covertext C. The stegotext is transmitted to Bob via the public channel. Hence both Bob and Eve observe S. The secret key K however is transmitted to Bob via the secure channel. Hence Eve has no knowledge of K.

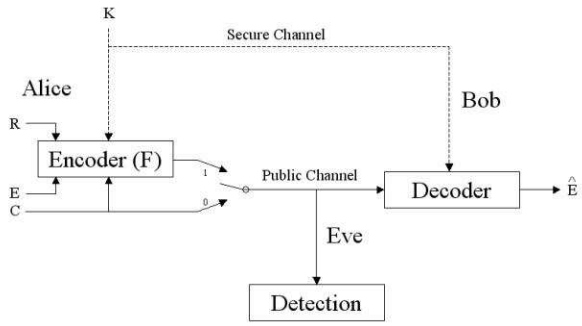


Figure 3: General framework for secret key stego system that generates, embeds and securely transmits the hidden data

The proposed algorithm works along the lines of a universal stegosystem, where no information about the covertext is required at the decoder. In most practical scenarios it is never desirable for an adversary to gain access to the unmarked host. Alice generates the stegotext S based on a secret key K and the covertext C and transmits the stegotext via the public channel and the secret key via the secure channel. In this case Eve has access to the distribution of the stegotext only (P_S). In such a scenario Eve does not have enough information to apply hypothesis testing and stochastic detection algorithms to check for the presence of hidden data. To compare the performance of the proposed algorithm, for all of the following tests, we implemented and tested a generic block based data-hiding scheme working in the DCT domain [25]. The average peak signal to noise ratio (PSNR) for the stego images created using the proposed algorithm was 38.35 dB.

3.1 The Westfeld Attack

For a blind stegosystem, Westfeld [33] proposed a technique that detects the sudden and consistent changes in the color palette of the stego image. Westfeld proposed that the theoretical comparison of the expected frequency distribution in stego-images and the host image will reveal the presence of a hidden signal. For this purpose Westfeld used the Chi-Square test. We tested the proposed and the comparison algorithms using the Chi-Square test available at [13].

A Chi-Square test takes into consideration the correlations between neighboring parts of an image to compute the probability that a message is embedded for a certain assumed hidden message length. We noticed that for a large message length of 1024 bits, the Chi-Square test failed detection, so we performed additional tests to see if there would be a difference for smaller message lengths. We chose to embed a smaller hidden message of size 40 bits. Figure 4 plots the probabilities of detection for assumed hidden message lengths ranging from 0 bits to 40 bits (for the images used in this set of tests, the actual length of hidden data is 40 bits). We note from the plots

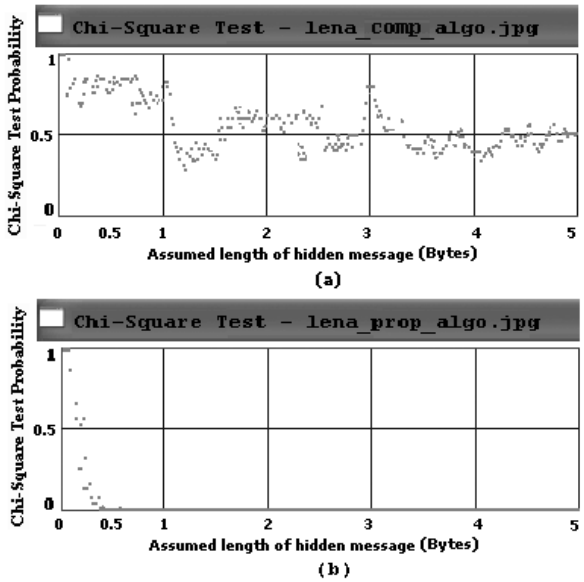


Figure 4: The results of the Chi-Square test for a 40 bit long hidden data in the stego image generated using (a) the comparison algorithm and (b) the proposed algorithm. The X-axis represents the assumed length of the hidden message for the Chi-Square test and the Y-axis denotes the estimated probability (by the Chi-Square test) that a message of that assumed length was hidden.

that as the test assumes a longer hidden message, the probability of detection reduces. This observation is logical because it is more probable for the test to find a high correlation between smaller neighboring parts of the image rather than finding a high correlation between larger neighboring parts of the image. For the proposed algorithm (see Figure 4(b)), the probability that a hidden message of length 0.5 bytes was embedded (4 bits) is less than 0.05 (5%) and the same for the comparison algorithm is 0.75 (75%) (see Figure 4(a)). As a matter of fact, we note that for the proposed algorithm using the Chi-Square test, the probability that a 6-bit or larger hidden message is present is estimated to be 0(0%). Whereas in fact there is a 5 byte (40 bit) hidden message present, which means that the proposed algorithm is not detectable using the Chi-Square attack. However, for the comparison algorithm we note that Chi-Square attack detects the 40-bit hidden data with a probability of 0.5 (50%).

3.2 Statistical Steganalysis

In this sub-section we analyze the effects of statistical attacks and comment on the performance of the proposed algorithm in comparison with the generic spread spectrum based algorithm. Figures 5(a), (b), and (c) show the histograms of the unmarked host image, the stego image for the Hernandez et al. algorithm, and the stego image for the proposed algorithm respectively. The Hernandez et al. algorithm [15] distorts the pdf of the stego image. It introduces two smaller peaks in the pdf as seen in Figure 5(c). However, the pdf of the stego image generated by the proposed algorithm does not have the additional peaks. (See Figure 5(b).)

3.3 The Low-Pass Filtering Effect

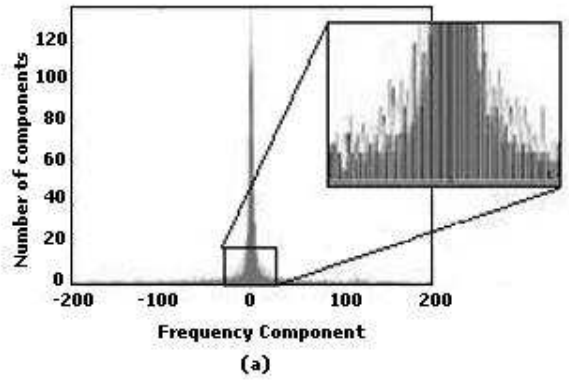
In [14], Harmsen et al. show that noise addition to a host in order to hide stego data corresponds to low pass filtering of the histogram of the host. Hence they suggest that if a noticeable number of high frequency components of the host image have low power, the hidden data could be detected. Their method starts with calculating a histogram, h, of the stego image. Then, h is transformed using the forward Fourier transform to obtain $F(h)$. Lastly the center of gravity (CoG) of the magnitude of $F(h)$, is calculated and used as the distinguishing statistic to differentiate between the stego and cover image. According to Harmsen, it is possible to keep a fixed threshold value (T) that determines whether or not one of the two copies obtained by Eve contains hidden data. Setting a small value for T ensures powerful detection, but also leads to increase in false positives. Similarly, a higher value of T decreases the false positives but the detection power deters as a result. We note that this detection is based on the adversary's knowledge of the cover data. In order to test this steganalyzer, we need to assume that the adversary has obtained a copy of the original host image. The CoG calculated by the procedure outlined by Harmsen et al., for the unmarked host, the stego image generated using the comparison algorithm and the stego image generated using the proposed algorithm was 37.11, 32.94 and 36.01 respectively. Hence we see that the distance between the CoG of the unmarked host and the stego images was reduced by 74% in the proposed algorithm.

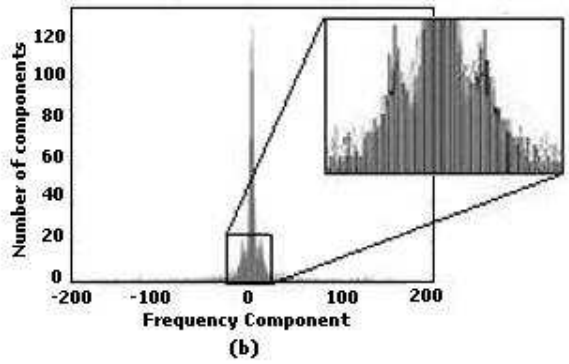
3.4 The Kullback Leibler Distance

Another way to determine the effect of embedding is to determine the Kullback-Leibler distance [5] between the pdf of the marked and unmarked images. This measure can be quantified using the relative entropy between the probability mass functions (PMF) of the host image and stego image. The relative entropy $D(p||q)$ between two PMFs p and q can be defined as

$$D(p \parallel q) = \sum_{x \in X_i} p(x)(\log(\frac{p(x)}{q(x)})),$$

where, $0 \log(0 \mid 0) = 0$ and $p \log(p \mid 0) = \infty$. We note that the KLD is not symmetric. Hence, $D(p \parallel q)$ is different from $D(q \parallel p)$. The KLD between the host image and the stego images generated by the proposed algorithm and the comparison algorithm ($D(p \parallel q)$) were calculated to be 0.0295 and 0.0505 respectively and the KLD between the stego images and the host image ($D(q \parallel p)$) were 0.0347 and 0.0656 respectively. The proposed algorithm provides a 41% (47%) reduction in the KLD, $D(p \parallel q)$ ($D(q \parallel p)$),





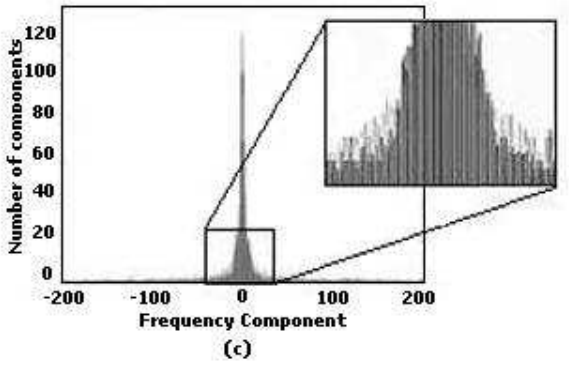


Figure 5: Histograms of (a)Unmarked host image, (b)The stego image for the comparison algorithm, and (c)The stego image for the proposed algorithm

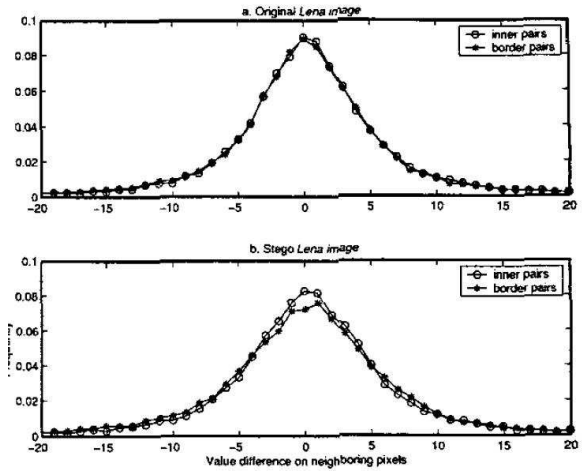


Figure 6: Value difference graphs of the histograms of inner and border DCT coefficient pairs for (a)Original Lena image and (b) Stego image [29]

between the distributions of the host and stego images as compared to the generic algorithm.

3.5 Block-DCT Image Steganalysis

Wang et al. [29] propose a block-DCT image steganalysis method that could be used to determine the presence of hidden data within an image. According to their method, the block structure of DCT embedding would lead to pairs of neighboring pixels within an 8x8 block to have different statistics from those across two blocks. In Figure 6 [29] we see the value difference graph that is reported by Wang et al. after subjecting a block-DCT embedding scheme to their proposed steganalysis algorithm.

We tested stego images generated by our proposed algorithm using the steganalysis method outlined by Wang et al. In Figure 7 we show a value difference graph generated for the difference of inner and border pairs for a stego image generated by the data-hiding algorithm proposed in this paper. As seen in Figure 7, we note that the value difference between the inner and border pairs as expected by Wang et al.'s steganalysis algorithm is not present for the case of the proposed algorithm.

4 Robustness Tests

In practice, there is a very good chance for a stego image to be altered (intentionally and unintentionally) while being transmitted through the channel. These alterations may be a result of intentional attacks such as filtering, blurring, requantization etc. or unintentional distortions such as lossy compression, channel noise addition etc. Distortions and attacks introduce an additional transformation between the stego image and the altered version available to the decoder. As a consequence, the performance of the decoder may be affected. Given a stego

Table 1: BER comparison for popular robustness attacks

Attack	BER(Proposed Algo)	BER(Comparison Algo)
Sharpening	0.0244	0.0343
High Pass Filter	0.0253	0.0298
Despeckle	0.1045	0.1180
Blurring	0.0435	0.0541
JPEG(Q-80)	0.0295	0.0334
JPEG(Q-50)	0.0320	0.0460
JPEG(Q-30)	0.0725	0.0811
5% Uniform Noise addition	0.0403	0.0550
5% Gaussian Noise addition	0.0757	0.0793

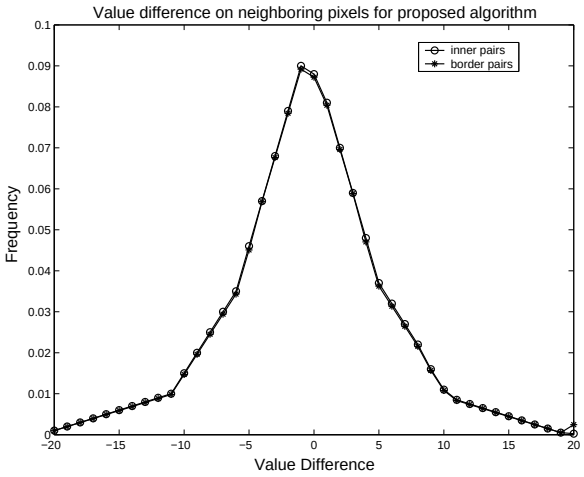


Figure 7: Value difference graph of the histograms of inner and border DCT coefficient pairs for a Stego image generated by the proposed algorithm

system, the role of the attacker is to alter the image in such a way that visually the image is not considerably degraded, however, the frequency distribution is transformed in such a manner that the decoder fails to extract the hidden data. To test the robustness of the proposed algorithm, the stego images were subjected to various image manipulating operations and compression attacks. In Table 1, we present the bit error rate (BER) after subjecting the stego images to popular image manipulation attacks.

As can be seen from Table 1, a simple sharpening or filtering attack does not affect the robustness of the hidden data by a considerable amount. The BER in these cases were 0.0244 and 0.0253 respectively i.e. more than 97% of the embedded data was recovered without any errors. Even in the case of dislocating attacks such as blurring and despeckle, the BER was 0.1045 and 0.0435 respectively. In the case of JPEG compression attacks, the BER was very low, since the data-hiding was performed in the DCT domain. For medium to low quality compression (Q 50-80), the BER was 0.0320-0.0295 i.e. more than 96% of the hidden data was recovered without any errors. Even for very low quality compression (Q-30) more than 92% of the hidden data was recovered without any errors. In the case of noise addition the retrieval rates were over 92% for additive uniform and white Gaussian noise.

5 Conclusion

We proposed a statistically secure data-hiding algorithm in the DCT domain. This algorithm provides statistical security and robustness against several attacks and avoids detection by techniques such as Chi-Square attack or the low-pass filtering attack. When compared to a generic data-hiding scheme based in the DCT domain, the proposed algorithm provides more than 41% reduction in the relative entropy between the PMFs of the host the stego images. The proposed algorithm was shown to be robust against a multitude of image manipulating attacks. In the case of JPEG compression attacks, even very low quality compression (Q-30) resulted in BER of 0.0725 i.e. more than 92% of the hidden data was recovered without any errors. In the case of robustness tests, the proposed algorithm out performed the generic DCT algorithm for all attacks. While providing significant robustness and capacity, this algorithm also induces low distortion in the host image with an average PSNR of more than 38 dB for all the images that were included in the test dataset.

Acknowledgment

This work was funded by National Science Foundation (NSF) award number NSF-DAS 0242417.

References

[1] P. Amin and K. Subbalakshmi, "Rotation and cropping resilient data hiding with Zernike moments," in *International Conference on Image Processing (ICIP'04)*, vol. 4, pp. 2175-2178, Singapore, Oct. 2004.

[2] P. Amin, N. Liu, and K. Subbalakshmi, "Statistically secure digital image data hiding," in *IEEE Multimedia Signal Processing (MMSP05)*, pp. 497-500, China, Oct. 2005.

[3] R. R. Clarke, "Relation between Karhunen-Loeve and cosine transforms," in *Proceedings of Instrumentation and Electrical Engineering*, vol. 128, pp. 359-360, 1981.

[4] R. Clarke, *Transform Coding of Images*, San Diego, CA, USA: Academic Press Professional, Inc., 1986.

[5] T. Cover and J. Thomas, *Elements of Information Theory*, New York, NY: Wiley, 1991.

[6] I. Cox, J. Kilian, T. Leighton, and T. Shamoon, "Secure spread spectrum watermarking for multimedia," *IEEE Transactions on Image Processing*, vol. 6, pp. 1673-1687, 1997.

[7] H. Farid, "Detecting steganographic message in digital images", Technical Report TR2001-412, Dartmouth College, Hanover, 2001.

[8] R. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*, vol. 7, pp. 179-188, 1936.

[9] V. Fotopoulos and A. N. Skodras, "A subband DCT approach to image watermarking," in *Tenth European Signal Processing Conference*, Tampere, Finland, Sept. 2000.

[10] J. Fridrich, M. Goljan, and R. Du, "Steganalysis based on JPEG compatibility," in *SPIE Multimedia Systems and Applications IV*, vol. 4518, pp. 275-280, 2002.

[11] J. Fridrich and M. Goljan, "Practical steganalysis of digital images-state of the art," *SPIE Photonics West*, vol. 4675, pp. 01-13, 2002.

[12] W. Fung and A. Kunisa, "Rotation, scaling, and translation-invariant multi-bit watermarking based on log-polar mapping and discrete fourier transform," in *IEEE International Conference on Multimedia and Expo (ICME)*, pp. 4, Japan, 2005.

[13] S. Guillermito, "A few tools to discover hidden data," http://www.guillermito2.net/stegano/tools/index.html

[14] J. Harmsen and W. Pearlman, "Steganalysis of additive noise modelable information hiding" in *Proceedings of SPIE Electronic Imaging*, vol. 5020, pp. 131-142, Santa Clara CA, 2003.

[15] J. Hernandez, M. Amado, and F. Perez-Gonzalez, "DCT-domain watermarking techniques for still images: Detector performance analysis and a new structure," *IEEE Transactions on Image Processing*, vol. 9, no. 1, pp. 55-68, 2000.

[16] N. Johnson and S. Jajodia, "Steganalysis of images created using current steganography software," in *Second Information Hiding Workshop*, LNCS 1525, pp. 273-289, 1998.

[17] D. Kim and S. Park, "A robust video watermarking method," in *IEEE International Conference on Multimedia and Expo (ICME)*, vol. 2, pp. 763, 2000.

[18] E. Lam and J. Goodman, "A mathematical analysis of the DCT coefficient distributions for images," *IEEE Transactions on Image Processing*, vol. 9, pp. 1661-1666, 2000.

[19] C. Lin, M. Wu, J. Bloom, I. Cox, M. Miller, and Y. Lui, "Rotation, scale, and translation resilient watermarking for images," *IEEE Transactions on Image Processing*, vol. 10, pp. 767-782, 2001.

[20] N. Liu and K. Subbalakshimi, "Vector quantization based scheme for data hiding for images," in *SPIE International Conference on Electronic Images*, pp. 548-559, 2004.

[21] N. Liu, P. Amin, and K. Subbalakshmi, "Secure quantizer based data embedding," in *IEEE Multimedia Signal Processing (MMSP05)*, pp. 509-512, China, Oct. 2005.

[22] L. Marvel, C. G. Boncelet, and C.T.Retter, "Spread spectrum image steganography," *IEEE Transactions on Image Processing*, vol. 8, pp. 1075-1083, 1999.

[23] S. Pholsomboon and S. Vongpradhip, "Rotation, scale, and translation resilient digital watermarking based on complex exponential function," in *2004 IEEE Region 10 Conference TENCON04*, pp. 307-310, Thailand, 2004.

[24] N. Provos, "Defending against statistical steganalysis," in *10th USENIX security symposium*, pp. 323-335, Washington DC, 2001.

[25] N. Provos and P. Honeyman, "Detecting steganography content on the Internet," CITI Technical Report 01-11, University of Michigan, pp. 1-14, 2001.

[26] K. Satish, T. Jayakar, C. Tobin, K. Madhavi, and K. Murali, "Chaos based spread spectrum image steganography," *IEEE Transactions on Consumer Electronics*, vol. 50, pp. 587-590, 2004.

[27] B. P. tzmann, "Information hiding terminology," in *First International Workshop on Information Hiding*, pp. 3233-3236, Dec. 2004.

[28] J. Vila-Forcen, O. Koval, S. Voloshynovskiy, and T. Pun, "Asymmetric spread spectrum data-hiding for laplacian host data," *IEEE International Conference on Image Processing (ICIP05)*, vol. 1, pp. 217-220, 2005.

[29] Y. Wang and P. Moulin, "Steganalysis of block-DCT image steganography," in *2003 IEEE Workshop on Statistical Signal Processing*, pp. 339-342, IL, USA, Sept. 2003.

[30] H. Wang and S. Wang, "Cyber warfare: Steganography vs. steganalysis," *Communications of the ACM*, vol. 47, pp. 76-82, 2004.

[31] T. Welch, "A technique for high performance data compression," *IEEE Computer*, vol. 17, pp. 8-19, 1984.

[32] A. Westfeld and A. Pfitzmann, "Attacks on steganographic systems," in *Proceedings of the Third International Workshop on Information Hiding*, LNCS 1768, pp. 61-75, 2000.

[33] A. Westfeld, "Detecting low embedding rates," in *5th International Workshop on Information Hiding (IH02)*, pp. 324, Netherlands, 2002.

[34] G. Xuan, C. Yang, Y. Zhen, Y. Shi, and N. Zhicheng, "Reversible data hiding based on wavelet spread spectrum," in *IEEE 6th Workshop on Multimedia Signal Processing*, pp. 115-124, Denver CO, 2004.

[35] D. Zheng, J. Zhao, and A. Saddik, "RST invariant digital image watermarking based on log-polar mapping and phase correlation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 1, pp. 753-765, 2003.

Palak Amin received the B.E. and the M.E. degree both in Computer Engineering from the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ in 2003. He is currently working towards the Ph.D. degree in Computer Engineering at Stevens Institute of Technology, Hoboken, NJ. He has been with the MedSW-West Lab, Siemens Medical Solutions at Iselin, NJ for 2001-2002. His research interests include multimedia security - digital image/video watermarking, statistical security, distributed source channel coding (DSCC), and multiple description coding (MDC).

Ning Liu received the B.E in Electrical Engineering from the Sichuan University, China in 1995, and the M.E in Signal Processing Engineering from the Tongji University, China in 2001. Since Fall 2002, He has been a Ph.D. student in the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ, where he works in the MSyNC lab under the guidance of Prof. Subbalakshmi. His research interests include quantizer based steganography and stego-games, digital image/video watermarking, joint source channel coding.

K. P. Subbalakshmi received the B.Sc degree in Physics from the University of Madras, India in 1990, the M.E in Electrical Communication Engineering from the Indian Institute of Science in 1994 and the PhD degree from the School of Engineering Science, Simon Fraser University, Canada in 2000. Since Fall 2000, she has been an Assistant Professor in the Department of Electrical and Computer Engineering, Stevens Institute of Technology, Hoboken, NJ, where she co-directs the MSyNC: Multimedia Systems Networking and Communications Laboratory. Dr. Subbalakshmi leads research projects in information security, joint source-channel, multiple description coding and multimedia networking funded by the National Science Foundation, Air Force Research Laboratory, New Jersey Center for Wireless Telecommunications, the Stevens Center for Wireless Network Security and industry partners. She has been an active participant in several international conference program committees and organizations.