QIM-BASED HISTOGRAM-PRESERVING AND HIGH CAPACITY STEGANOGRAPHY FOR JPEG IMAGES

Xiaojun Qi	**Daniel Lewis**
Computer Science Department	Computer Science Department
Utah State University	Utah State University
Logan, UT, United States	Logan, UT, United States
xqi@cc.usu.edu	daniell@cc.usu.edu

Abstract - *We propose a secure steganographic approach with a high capacity. We first use small neighborhoods of the original histogram to calculate the minimum remapping from a histogram to the odd-only and even-only histograms. We then employ a quantization index modulation-based odd-even embedding to preserve the histogram of each embedding subchannel. A minimum distortion remapping policy is used to decide the quantization value so the coefficient is increased or decreased according to a constant probability. Our embedding ensures the modification occurs in a small neighborhood and the details of the original histogram are maintained. We extensively compare our approach with two state-of-the-art steganographic approaches, MB1 and MB2, using Fridrich's 23-feature-based blind steganalysis tool with a linear discriminant analysis. The results show that our method achieves better embedding capacity, comparable image quality, and comparable security to MB1 and MB2.*

Keywords: steganography, steganalysis, histogram preserving, quantization index modulation

1 Introduction

Steganography has been mainly used in information security applications. It transmits information by embedding messages into innocuous-looking cover objects, such as digital images, to conceal the existence of communication. As a result, it is the art of invisible communication and the science of data smuggling. The most desirable property of any steganographic approach is to maximize the amount of hidden information (embedding rate) while preserving security against detection by unauthorized parties. The information-theoretic model [1] proves that a perfectly secure steganographic system should ensure the statistics of the cover image and its stego are identical. In this paper, we present a high capacity DCT-based steganographic approach by preserving odd-even histograms. Our work is presented with the passive warden scenario [2] and assumes that there is no channel noise during data transmission. We choose to work in the DCT domain due to an unavoidable JPEG compression "attack". That is, we focus on necessary data modifications for information hiding to ensure the PDFs (Probability Mass Functions) of the quantized DCT coefficients are almost identical to those of the JPEG compressed data without any hidden data information.

We will briefly review several related steganographic methods, which focus on preserving image statistics to increase the undetectability of the stego by steganalysis tools. Eggers et al. [3] and Tzschoppe et al. [4] respectively propose to preserve the histograms of each DCT subchannel by applying the HPDM (Histogram-Preserving Data-Mapping) to calculate a remapping for guiding minimal changes to individual DCT coefficients by encoding one bit per change. Specifically, an entropy encoder [3] is used to match the message distribution with the subchannel parity distribution. This method generally preserves the histograms with the sacrifice of the payload. However, it usually skews in JPEG images with an average quality due to the large number of zeroes introduced in the embedding process. To address this issue, Tzschoppe et al. [4] propose to keep a uniform parity distribution of each subchannel within an embeddable zone, which contains coefficients with absolute value greater than one. But, this embeddable zone further decreases the embedding capacity. Recently, a new paradigm called the MB (Model-Based) steganography [5, 6] was invented to ensure the stego signal conforms to an implicit model. Specifically, the MB1 method [5] models subchannel histograms as parameterized Cauchy or Gaussian distributions and calculates minimal modifications for preserving the distributions. The modifications are then used as a symbol set to encode messages. The MB2 method [6], an extension to MB1, uses the statistical restoration technique for reserving half of the DCT coefficients to maintain more image statistics. To our knowledge, these two model-based

approaches (e.g., MB1 and MB2) are the most robust steganographic systems to date.

Along with the evolution of steganographic methods, many steganalysis methods have been invented to detect the presence of the stego and/or provide information about the length or location of the message. These methods can be generally classified into two categories: specific steganalysis against a method or a class of methods, and universal blind steganalysis. The first category needs to be customized to each embedding paradigm, and the design of proper distinguishing statistics cannot be easily automated. The second category is formed by blind classifiers without knowledge about the applied steganographic system. Here, we review several universal blind steganalysis tools due to their effective power in detecting stegos. Some pioneer work is reported in higher order statistical models, wherein a linear classifier [7] or a SVM (Support vector machine) [8] is trained on 72 features to learn the differences between cover and stego images. These two methods work well on J-Steg, OutGuess, and F5. However, it cannot successfully detect the stegos constructed by HPDM methods. Fridrich's feature-based steganalysis method for JPEG images [9] improves the above methods and is the most effective blind method to date. The Fisher linear discriminant classifier uses 23 calibrated features, which are likely altered during embedding, to learn the differences between cover images and their stegos. It can detect the stegos generated by JP Hide & Seek, OutGuess, F5, HPDM-based methods, MB1, and MB2. SVMs trained on the same features [10] are also used to improve the performance of the steganalysis tool.

In this paper, we propose a novel QIM-based (Quantization Index Modulation [11]) odd-even histogram preserving and high capacity steganographic method for still images. The remainder of the paper is organized as follows. Section 2 presents our proposed method. Section 3 compares our proposed method with the two powerful steganographic methods, MB1 and MB2, by using the 23-feature-based blind steganalysis tool. Section 4 concludes the paper and shows the direction of the future work.

2 Our proposed approach

In this research, we improve the HPDM algorithm to implement a secure steganographic approach with a high capacity. First, we use small neighborhoods of the original histogram rather than large overall measures (i.e., the cumulative probability distribution) to calculate the minimum remapping from a histogram to the odd-only and even-only histograms. This improvement tends to reduce the effect of skewed distributions on the entire histogram in [3, 4]. Second, we implement a QIM-based odd-even embedding to preserve the histogram of each individual DCT subchannel. That is, we force a quantized DCT coefficient to be odd when embedding a 1, and even when embedding a 0. Third, we use a minimum distortion remapping policy to respectively decide the quantizers for the altered quantized even and odd DCT coefficients to preserve the original histogram. This strategy ensures the modification occurs in a small neighborhood and the details of the original histogram are approximated more precisely after embedding. As a result, it creates less modification of the original image and maintains more image statistics. Fourth, we embed the message in appropriate mid-frequency JPEG subchannels to avoid heavily quantized subchannels and the DC subchannel. The detailed components of the proposed method are explained in the following subsections.

2.1 The embedding scheme

The block diagram of the proposed method is shown in Fig. 1. In the following, we will discuss each block in detail.

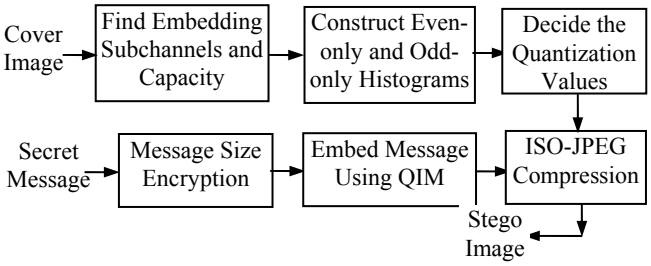


Figure 1: Block diagram of our proposed approach

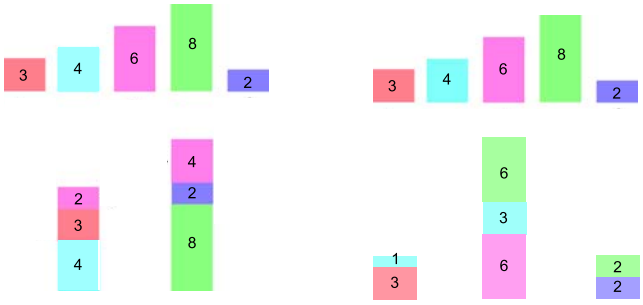
Find Embedding Subchannels and Capacity: Collect the embeddable quantized DCT coefficients (i.e., all nonzero values) at the zig-zag number 2 through 36 in each 8×8 DCT block. The DCT coefficients at a specific zig-zag number from all DCT blocks form one embedding subchannel. In total, 35 embedding subchannels are constructed. These 35 subchannels are empirically chosen to ensure good capacity and robustness since low-frequency stego (noise) is usually more noticeable and high-frequency stego leads to more heavily skewed quantized coefficients in a distribution. The embedding capacity equals the total number of embeddable DCT coefficients in the 35 embedding subchannels.

Construct Even-Only and Odd-Only Histograms: The histograms of positive and negative quantized DCT coefficients are separately calculated for each subchannel. Our improved HPDM method is then applied to calculate a minimum remapping from a histogram to the odd-only and even-only histograms by considering small neighborhoods of the original positive and negative histograms. Fig. 2 illustrates the idea by simultaneously remapping a histogram to the even-only and odd-only histograms. The detailed steps for the localized remapping from a histogram to an even-only histogram are as follows:

1. Copy all the original histogram's even bins to the corresponding even bins in the new histogram.
2. Redistribute the neighboring odd bins of each even bin in the new histogram. Two adjacent odd bins exist when the even bin is away from the protected

zone, where no embedding is allowed. One adjacent odd bin exists when the even bin is next to the border of the protected zone. In the first case, the contents of the odd bin are shared in proportion to the sizes of the adjacent even bins in the original histogram. In the second case, the original odd bin's contents are added to the even-only bin's total. In our system, we specify all zero DCT coefficients as the protected zone for maintaining more statistical properties.

The same strategy can be used in remapping a histogram to the odd-only histogram since the even-only and odd-only histograms are dual.



(a) Original to even histogram (b) Original to odd histogram

Figure 2: Illustration of the localized mapping from a histogram to an even-odd histogram

Decide the Quantization Values: The minimum distortion remapping policy is applied to decide the quantization values for the QIM embedding scheme. Here, we explain the process of using the even-only histogram to decide the quantization values for odd quantized DCT coefficients. The dual process can be easily derived to decide the quantization values for even quantized DCT coefficients using the odd-only histogram. Below is the strategy to apply the minimum distortion remapping policy to decide the quantization values for odd coefficients.

1. Create cumulative versions of the original and the even-only histogram as running totals starting from the lowest-absolute-value bin outside the protected zone.
2. If the endpoints of the original histogram are odd bins, no calculation is necessary since their parity can only be forced to the adjacent even value.
3. Otherwise, each count in the cumulative even-only histogram is used twice to respectively compute the quantization value for its two odd neighbors by using two corresponding bracketed counts (i.e., the odd counts) in the cumulative original histogram. The interpolation technique is used to compute the percentage of the distance between the two bracketing counts to the even-only count. This percentage defines the remap-down rate for the odd quantized DCT coefficient.

Embed Message Using QIM: Five steps are involved in this process:

1. Find the message size in bits, append this message size as the first 16 bits of the message, and encrypt the message using the AES (Advanced Encryption Standard) method [12] with a shared key.
2. Divide the total message length (16 bits + message size) by the image capacity to get the embedding rate.
3. Find the number of message bits to be embedded in each embeddable subchannel by rounding the multiplication of each subchannel's capacity with the embedding rate. If the total number of message bits calculated this way is less than the total message length due to the round-off error, we simply add the leftover message length to each subchannel one bit at a time until the entire message length is used.
4. Split the message into the subchannel as computed in the previous step.
5. Apply the QIM scheme to embed the split message in each subchannel in a permutated order determined by a user-specified key. That is, to embed a 0 or 1 in a given quantized DCT coefficient, force that coefficient to the corresponding parity. Specifically, if embedding 0 in even coefficients or 1 in odd coefficients, the coefficient requires no changes since the parity is the same as the embedded bit. On the other hand, if embedding 0 in odd coefficients or 1 in even coefficients, force the parity up or down using the quantization values determined from the minimum distortion remapping policy.

ISO-JPEG Compression: The modified block DCT coefficients undergo the same zig-zag scan, and differential and run-length coding as in the JPEG compression scheme. The resulting stego image is transmitted to the intended receiver as a JPEG image.

2.2 The extraction scheme

The extraction process can be carried out by reversing the embedding procedure. That is, generate the same random permutations of each embedding subchannel as the encoder by using the same user-provided key. Decrypt the first 16 bits in the message to find the embedded message length and compute the embedded message length in each subchannel. This length together with the shuffler shows exactly which quantized DCT coefficients carry the secret message. The message bits can then be reconstructed by taking remainder of each shuffled, embedded quantized DCT coefficients divided by 2.

3 Experimental results

We performed several controlled experiments on our proposed approach to demonstrate its capacity and security. Furthermore, we compared our approach with two state-of-the-art steganographic approaches, MB1 (without deblocking) and MB2 (with deblocking), using Fridrich's universal blind steganalysis features. In all experiments, we embedded messages whose lengths are proportional to

the number of non-zero DCT coefficients in each image to create the corresponding stego image databases. The testing was done for the following relative embedding rates expressed in bpc (Bits Per non-zero DCT coefficients), bpc = 0.1, 0.2, 0.4, 0.6, and max. If the bpc rate is larger than the maximal bpc rate bpc_{max} determined by the image capacity, we use bpc_{max} as the embedding rate. This configuration ensures the detection is approximately of the same level of difficulty and is also consistent with the experiments performed in Fridrich's 23-feature steganalysis tool. Furthermore, it caters to the real-world scenario where the user needs to know which algorithm best preserves secrecy for a given message and cover image.

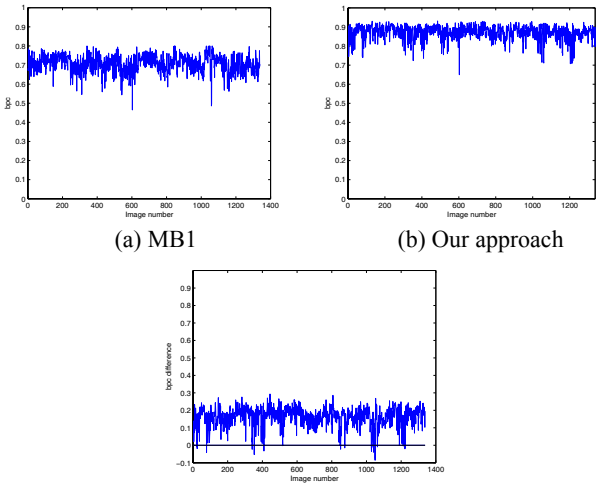
3.1 Data set

We tested our approach using the images in the Uncompressed Color Image Database [13], which contains a set of 1338 uncompressed RGB images with dimensions near 512×480. All images were converted to grayscale and compressed using an 80% JPEG quality. The quality factor of 80 was chosen because it is the default for some applications and is commonly seen on the Internet. Such a choice also ensures fruitful comparisons since prior work in this area has used 80 as the default setting.

3.2 Embedding capacity

High embedding capacity is a desirable feature for any steganographic method. We compute a theoretical capacity for each individual image by dividing the total number of allowable embedding bits by the total number of non-zero quantized DCT coefficients. It is measured as the relative embedding rate expressed in bpc. Fig. 3 shows the capacity of MB1 (a scheme that has a substantially higher embedding efficiency than previous art) and our approach together with the capacity difference between these two approaches on each of the 1338 images. The capacity of MB2 is not shown here since it is half of that of MB1. The figure clearly shows that our capacity is superior to MB1's. As shown in Fig. 3(c), our algorithm can embed messages approximately 25% longer than MB1. For certain images, MB1 has more capacity than our algorithm, but his happens only 21 times, less than 2% of the image set. Consequently, we claim that our proposed approach has a substantially higher embedding rate than previous art.

There are two reasons for the difference in the embedding capacities. First, MB1's image capacity depends heavily on the distributions of coefficients in various subchannels. Specifically, when the symbol set created for embedding is not favorable to embedding in rarer coefficients with high information content, the embedding capacity will be compromised. Our algorithm tends to have better capacity in this situation. Second, our algorithm's capacity uses a more consistent calculation of capacity by fixing the embedding channels at 2 through 36. We compared using this set of subchannels with using subchannels 2 through 21 (as in [3, 4]) and found that adding subchannels above 21 increased the overall capacity dramatically without increasing the detection rate. We do not take advantage of as many subchannels as MB1, even though other JPEG quantization factors are similar to those of the subchannels we do use. A more adaptive approach for a given quality factor could use a subchannel for embedding when the quantization factor passes a certain threshold, and might increase the capacity.



(a) MB1 (b) Our approach

(c) Advantage of our approach over MB1

Figure 3: Capacity of the tested techniques expressed in bpc

3.3 Objective image quality

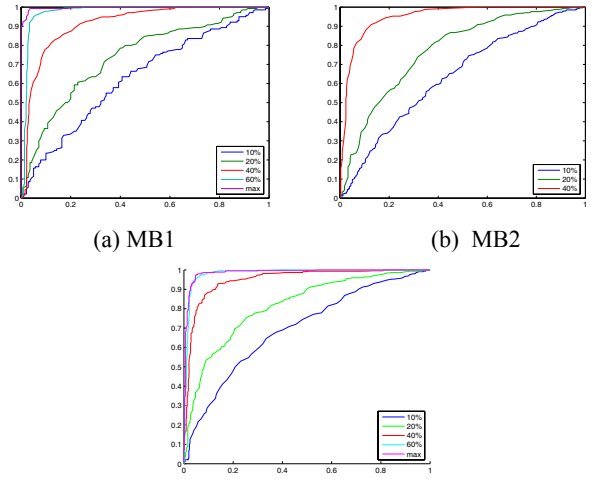
Image quality is another important measure for any steganographic method. We use the SSIM (Structural SIMilarity index) [14] to quantify image quality relative to reference images. The SSIM normalizes local luminance and contrast to highlight structural comparison and treat non-structural information as noise before applying similarity measures. It correlates with human perceptions of image quality better than the PSNR value. That is, the higher value of SSIM, the less distortion on the cover image in terms of the human perceptions on the visual similarity.

To test visual degradation of stego images, we used the cover images as references. At the embedding capacity of 0.6 bpc on all 1338 cover images, MB1 has an average image quality of 0.9789, while our method has an average image quality of 0.9728. At the maximum embedding capacity for MB1 on all 1317 cover images, wherein our method has a higher embedding capacity than MB1, MB1 has an average image quality of 0.9752 and our method has an average image quality of 0.9679. As a result, our method is capable of embedding messages with comparable perceptual distortion on the cover images. However, the alterations are invisible to the human system for both embedding algorithms.

3.4 ROC (Receiver Operating Characteristic) curves

We used ROC curves to analyze each steganographic technique and compare their embedding security. For a fixed relative message length expressed in terms of bpc = 0.1, 0.2, 0.4, and 0.6 for each of the 1338 images, we created a training database of stego images. The Fisher linear discriminant classifier was trained on Fridrich's 23 features derived from 892 cover and its corresponding stego images. The generalized eigenvector obtained from this training was used to compute the ROC for the remaining 446 cover and corresponding stego images. We also plotted the ROC curve for the maximum embedding rate for MB1 by removing 21 total images, for which our method cannot embed MB1's maximum, from the training and testing sets.

Fig. 4 plots the ROC curves of three steganographic techniques, namely, MB1, MB2, and our approach, at several fixed relative message lengths. It clearly shows that the detection accuracy, shown as the area under the ROC curve, increases with increased embedding for all three techniques. Our approach also has comparable detection accuracy for the largest embedding rates when compared with MB1. This is desirable since maximizing the embedding rate while preserving security against detection by unauthorized parties is the most desirable property of any steganographic approach.



(a) MB1 (b) MB2

(c) Our Approach

Figure 4: ROC curves for three methods at constant bpc rates

We also quantitatively evaluated the detection performance using detection reliability ρ defined as:

$$\rho = 2A - 1 \qquad (1)$$

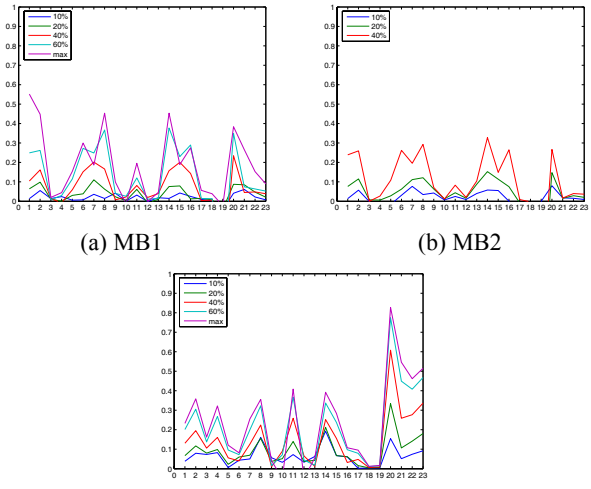
where A is the area under the ROC curve, also known as detection accuracy. This scales detection accuracy so that ρ=1 indicates a perfect detection and ρ=0 indicates a failure detection. When the ROC curve coincides with the diagonal line, the reliability of detection is 0. The detection reliability for MB1, MB2, and our approach at different embedding rates is shown in Table 1. It is clear that our algorithm is the least detectable for higher embedding rates. It is interesting that MB2 is a little bit easier to detect than MB1 on our testing images, which contradicts a result in [9] that MB2 is significantly more secure than MB1. The images used in [9] are drawn from a different database, so different results are to be expected. However, the fact that MB2's security boost has not generalized to our image set is interesting. In the following subsections, we will discuss the detection reliability of individual and combined features for all three embedding algorithms and offer some insights to the above observations.

Table 1: Detection reliability ρ for MB1, MB2, and our approach for different embedding rates (U means unachievable rate, Max is equal to the maximum embedding rates of MB1)

bpc	MB1	MB2	Ours
0.1	0.2602	0.2768	0.3942
0.2	0.4774	0.5440	0.6320
0.4	0.8276	0.8990	0.8996
0.6	0.9500	U	0.9667
Max	0.9934	U	0.9844

3.5 Impacts of individual features

We studied the importance of each of Fridrich's 23 calibrated features for MB1, MB2, and our proposed method by creating classifiers using only one feature at a time and measuring the detection reliability using Eq. (1) for each. Fig. 5 shows classifications that use individual features for different embedding rates. The collection of individual detection accuracy may not capture the performance of the detection algorithm in the 23-dimensional space. This is because that it is possible that none of the individual features themselves have any distinguishing power, but the collection of some features or all features achieves a better or perfect detection. Nevertheless, we use the detection reliability ρ as a measure of each feature's detection power.



(a) MB1 (b) MB2

(c) Our approach

Figure 5: Detection reliability ρ for classifiers using one feature at a time

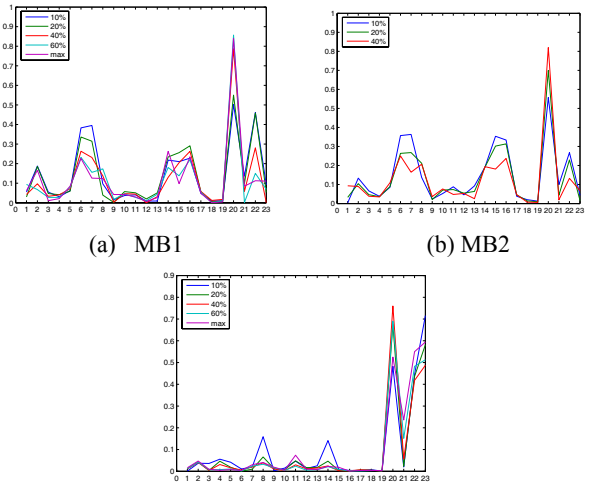
These graphs clearly show that the detection reliability generally increases as the embedding rate increases for

using each individual feature of all steganographic methods as the inputs to the classifier. However, the detection reliability using certain individual features (e.g., features 9, 10, and 12: the dual histograms of -2, -1, and 1) decreases at higher embedding rates in our scheme. In addition, detection reliability curves of different embedding rates follow different patterns, and the detection reliability does not have a drastic increment as the embedding rate increases. These observations may indicate that our approach has less detection reliability when the embedding rates increase.

Another interesting observation is that the impact of individual features on MB2 does not seem to be improved when compared to MB1 with the same level of embedding rate. In particular, the blockiness (features 19 and 20 as shown in the diagram) has little power when used as an individual feature, and the overall detection profiles are very similar. Somewhat surprisingly, all three schemes show detection spikes for the dual histograms for -3 and 3, features 8 and 14, even though the embedding schemes are rather different. Other than these, the profiles show that our algorithm, broadly speaking, does better than MB1 at preserving individual subchannels, but worse at preserving joint co-occurrence statistics (features 20 through 22).

3.6 Impacts of all 23 features

In view of the shortcomings of the study on the impacts of the individual features, we further investigate the detection reliability of using all Fridrich's 23 calibrated features. Fig. 6 shows this classification by using all 23 features on three steganographic approaches (MB1, MB2, and ours) at different embedding rates.



(a) MB1 (b) MB2

(c) Our approach

Figure 6: Individual feature strength for classifiers using all features at once

To study a particular algorithm, we first train the classifier using all 23 features of the cover images and their corresponding stego images. Second, we multiply the weights by the corresponding features and subtract them by the mean for normalization. This gives a rough estimate of how much each feature weighs in the classification relative to the other features by eliminating biases in the scale of the features. Third, we find the percentage change from each cover's features to each stego's features and subtract this percentage change by the mean of the absolute values of these errors for normalization. Finally, we multiply the normalized relative weights by the normalized error rate for each feature to obtain the importance of each feature relative to the others. This combines how much each feature weighs with how much it can vary.

Features with low weight and low variation contribute little to the classification, and therefore receive low scores on this method. Features with low weights and high variation make larger changes to image statistics, but these changes are not as useful to the classifier. Similarly, features with high weights and low variation make small changes to image statistics that are useful to the classifier. Each of these cases receives medium scores on this method. Features with high weights and high variation affect the classification most strongly by creating higher uncertainties about their effect on the classification, and therefore receive high scores on this method. Thus, this method compares features' classification power relative to each other, rather than independently.

For MB1 and MB2, several spikes in the feature graph become less relevant as the embedding rate increases. Features 6, 7, and 22 (dual histograms for -5 and -4, and co-occurrence N_{11}) are good examples. However, feature 20 (co-occurrence N_{00}), a large spike in the feature graph, increases in importance as the embedding rate increases. This confirms our observation that MB2 does not substantially alter the detection reliability because feature 20 dominates the detection as much as in MB1. For our method, most features become less relevant as the embedding rate increases. Especially, the spikes corresponding to features 8, 11, and 14 (the dual histograms for -3, 0, 3) decrease in prominence as the embedding rate is increased. However, features 20 and 22 (co-occurrence N_{00} and co-occurrence N_{11}) increase in importance as the embedding rate increases. Also, our approach has the lowest detection reliability for more features than MB1. This may indicate that our approach is expected to have better performance for a wider class of statistics.

In MB2 and our approach, the non-embedded coefficients can be used for statistical restoration. For MB2, the restriction on these coefficients is that any changes to the coefficients must preserve the statistical model of the subchannel coefficients calculated by the embedder. For our algorithm, the only restriction on these coefficients is that zero coefficients remain zero and non-zero coefficients remain non-zero; this allows the same permutations of embeddable coefficients performed at the encoder to be calculated at the decoder. This looser restriction allows more flexible approaches to statistical restoration that would be impossible to achieve with MB2. If a method to restore joint-co-occurrence statistics

becomes available, it should be easier to implement on top of our embedding scheme.

4 Conclusions

In this paper, we propose a robust steganographic approach to embed a high capacity message in a JPEG cover image. This approach improves both HPDM [3] and modified HPDM [4] by taking local idiosyncrasies of subchannel histograms into account. The contributions of the proposed approach are: 1) Use small neighborhoods of the original histogram to calculate the minimum remapping from a histogram to the odd-only and even-only histograms. 2) Employ a QIM-based odd-even embedding to preserve the histogram of each embedding subchannel. 3) Use a minimum distortion remapping policy to decide the quantization value so the coefficient is increased or decreased according to a constant probability for preserving the original histogram. 4) Ensure the modification occurs in a small neighborhood and the details of the original histogram are maintained.

Our extensive experimental results show that our scheme achieves higher embedding capacity, comparable image quality, and comparable security to MB1 and MB2 at the embedding rate of bpc>0.6. In addition, we have extensively studied the importance of each individual feature and all 23 features on the detection reliability at different embedding rates for three steganographic approaches, i.e., MB1, MB2, and our approach. The experimental results show that our scheme affects individual features in a different manner as both MB1 and MB2. That is, the detection reliability of some individual features is decreased when the embedding rate increases. Furthermore, our approach has the lowest detection reliability for more features than MB1 when all 23 features are considered for classification.

In the future, we will identify the features that our algorithm alters heavily by omitting the key features identified in the 23-feature classifier and redoing classification on the remaining features. Some iterative steps may be needed in locating the features our algorithm affects most, but the co-occurrence statistics are likely candidates. We will further plan on using the remaining capacity to reduce the modifications on the identified important features.

5 References

[1] C. Cachin, "An Information-Theoretic Model for Steganography," *Proc. of the 2nd Int. Workshop on Information Hiding*, LNCS Vol. 1525, pp. 306-318, 1998.

[2] S. Katzenbeisser and F. Petitcolas, *Information Hiding Techniques for Steganography and Digital Watermarking*, Artech House Publishers, 2000.

[3] J. J. Eggers, R. Bäuml, and B. Girod, "A Communications Approach to Image Steganography," *Proc. of SPIE on Electronic Imaging, Security and Watermarking of Multimedia Contents* IV, Vol. 4675, pp. 26-37, 2002.

[4] R. Tzschoppe, R. Bäuml, J. Huber, and A. Kaup, "Steganographic System Based on Higher-Order Statistics," *Proc. of SPIE on Electronic Imaging, Security and Watermarking of Multimedia Contents V*, pp. 156-166, 2003.

[5] P. Sallee, "Model-Based Steganography," *Proc. of Int. Workshop on Digital Watermarking*, pp. 154-167, 2003.

[6] P. Sallee, "Model-Based methods for Steganography and Steganalysis," *Int. J. of Image Graphics*, Vol. 5, No. 1, pp. 167-190, 2005.

[7] H. Farid, "Detecting Hidden Messages Using Higher-Order Statistical Models," *Proc. of IEEE Int. Conf. on Image Processing*, Rochester, NY, 2002.

[8] H. Farid, and L. Siwei, "Detecting Hidden Messages Using Higher-Order Statistical Models and Support Vector Machines," *Proc. of the 5th Int. Workshop on Information Hiding*, LNCS Vol. 2578, pp. 340-354, 2002.

[9] J. Fridrich, "Feature-Based Steganalysis for JPEG Images and its Implications for Future Design of Steganographic Schemes," *Proc. of the 6th Int. Workshop on Information Hiding*, LNCS Vol. 3200, pp. 67-81, 2004.

[10] J. Fridrich and T. Pevny, "Towards Multi-class Blind Steganalyzer for JPEG Images," *Proc. of Int. Workshop on Digital Watermarking*, LNCS Vol. 3710, Springer-Verlag, pp. 39-53, 2005.

[11] B. Chen and G. W. Wornell, "Quantization Index Modulation: A Class of Provably Good Methods for Digital Watermarking and Information Embedding," *IEEE Trans. on Information Theory*, Vol. 47, No. 4, pp. 1423-1443, 2001.

[12] J. Daemen and V. Rijmen, *The Design of Rijndael: AES - The Advanced Encryption Standard*, Springer, NY: 2002.

[13] G. Schaefer and M. Stich, "UCID: An Uncompressed Color Image Database," *Proc. of SPIE on Storage and Retrieval Methods and Applications for Multimedia*, Vol. 5307, pp. 472-480, 2003.

[14] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Trans. on Image Processing*, Vol. 13, No. 4, pp. 600-612, 2004.