CMC, vol.56, no.2, pp.261-272, 2018

A Novel Universal Steganalysis Algorithm Based on the IQM and the SRM

Yu Yang[1, 2, *], Yuwei Chen[1, 2], Yuling Chen[2] and Wei Bi[3, 4]

Abstract: The state-of-the-art universal steganalysis method, spatial rich model (SRM), and the steganalysis method using image quality metrics (IQM) are both based on image residuals, while they use 34671 and 10 features respectively. This paper proposes a novel steganalysis scheme that combines their advantages in two ways. First, filters used in the IQM are designed according to the models of the SRM owning to their strong abilities for detecting the content adaptive steganographic methods. In addition, a total variant (TV) filter is also used due to its good performance of preserving image edge properties during filtering. Second, due to each type of these filters having own advantages, the multiple filters are used simultaneously and the features extracted from their outputs are combined together. The whole steganalysis procedure is removing steganographic noise using those filters, then measuring the distances between images and their filtered version with the image quality metrics, and last feeding these metrics as features to build a steganalyzer using either an ensemble classifier or a support vector machine. The scheme can work in two modes, the single filter mode using 9 features, and the multi-filter mode using 639 features. We compared the performance of the proposed method, the SRM and the maxSRMd2. The maxSRMd2 is the improved version of the SRM. The simulated results show that the proposed method that worked in the multi-filter mode was about 10% more accurate than the SRM and maxSRMd2 when the data were globally normalized, and had similar performance with the SRM and maxSRMd2 when the data were locally normalized.

Keywords: Image steganalysis, IQM, SRM, total variation, universal image steganalysis.

1 Introduction

Methods that can discover steganographic images are called steganalysis algorithms [Ker (2007a, 2007b); Li, Zeng and Yang (2008); Peony (2007); Ker and Lubenko (2009); Xia, Wang, Sun et al. (2016, 2014)]. The steganalysis methods that are not limited to detect the particular information hiding tools are universal steganalysis algorithms, or blind steganalysis algorithms [Tan and Li (2014); Chen and Shi (2008); Broda, Levicky, Banoci et al. (2014)]. Here, the term universal means that the methods are effective for

[1] Information Security Center, Beijing University of Posts and Telecommunications, Beijing, 100876, China.

[2] Guizhou Provincial Key Laboratory of Public Big Data, Guizhou University, Guiyang, Guizhou, 550025, China.

[3] SeeleTech Corporation, San Francisco, 94107, USA.

[4] Zsbatech Corporation, Beijing, 100088, China.

[*] Corresponding Author: Yu Yang. Email: yangyu@bupt.edu.cn.

CMC. doi:10.3970/cmc.2018.02736

many different steganography algorithms and the term blind emphasizes that the analysis is carried out without the knowledge of the used hiding methods.

Considering the rapid speed at which the variant or new hiding methods have spawned, the universal steganalysis methods are very important to identify potential threats from messages covered by huge amount of images on the internet. A universal steganalysis method should detect as many steganography methods and should have as low dimension of a feature vector as it can. However, current blind steganalysis methods are not yet good enough in these regards.

The state-of-the-art blind steganalysis technique is the spatial rich model (SRM) [Fridrich and Kodovsky (2012)] designed by Fridrich. This method evolved from the subtractive pixel adjacency matrix (SPAM) algorithm [Pevný, Bas and Fridrich (2010)], which focused on relationships between neighboring pixels. The SPAM method utilizes a high-order Markov chain and a transition probability matrix to discover distortions due to steganographic embedding. The SRM replaces these simple multi-directional differences with various types of residuals obtained by linear and nonlinear filters. The SRM can accurately detect most challenging steganographic algorithms such as least significant bit matching (LSBM), edge-adaptive (EA) [Luo, Huang and Huang (2010)], and highly undetectable steganography (HUGO) [Pevný, Filler and Bas (2010)]. However, the SRM uses 34671 features, which means that its procedures of feature extraction and machine learning are time-consuming. Although there are many improved algorithms based on the SRM, the essential problem remains. For example, depending on embedding possibilities of pixels, Tang et al. [Tang, Li, Luo et al. (2015)] assigned different weights to these pixels during feature extraction. Yu et al. [Yu, Li, Cheng et al. (2016)] proposed another feature called contrast feature, which consists of the residual angle and norm. These methods improve SRM’s detection accuracy for the low embedding rate adaptive steganography methods, such as WOW [Holub and Fridrich (2012)] and UNIWARD [Holub, Fridrich and Denemark (2014)], but the feature dimensionalities of these methods are still high.

Based on two considerations, another kind of prevailing steganalysis methods [Ye, Ni and Yi (2017); Xu, Wu, Shi et al. (2016); Zeng, Tan and Huang (2018)] uses deep learning approaches. The first is that the abilities of feature auto-extraction of the deep learning networks can easy the design of the steganalysis feature. The second is that the deep learning architectures are similar to the mechanism of the SRM so they have the potential to replace the SRM with better performance. However the deep learning based steganalysis method are still being explored, and the complexity of these methods remain high.

The steganalysis method using image quality metrics [Avcbas, Memon and Sankur (2003)], called IQM, also focuses on residuals and uses much fewer features. However, the IQM-based approaches [Xu, Wang and Liu (2007); Geetha, Sindhu and Kamaraj (2008, 2007)] are only valid for some steganographic tools because the filter is only applicable to some types of steganographic noise. This paper presents a novel universal blind algorithm based on the IQM and the SRM, which uses at most 639 features to achieve similar or better detection performance than the SRM using 34671 features. Novel zero-watermarking scheme based on DWT-DCT [Yang, Lei, Liu et al. (2016)] and discrete cosine transform (DCT) is discussed. An overview of general theory of security

[Lei, Yang, Niu et al. (2017)] which is devoted to constructing a comprehensive model of network security is discussed.

This paper is organized as follows. Section 2 describes the proposed method and details the design of the filters, distance measures, classifying method, and etc. Section 3 presents the results of simulation and analyzes the performance of the proposed method. Finally, a brief conclusion is given in Section 4.

2 The proposed algorithm

The proposed method is based on the IQM, and the main idea of which is that the distribution of distances between a natural image and its filtered version is different from the distribution of distances between a steganographic image and its denoised version. As shown in Fig. 1, in order to detect whether an image $\boldsymbol{X} \in \mathcal{R}^M \times \mathcal{R}^N$ contains a secret message, the method first filters $\boldsymbol{X}$ using the function $h(\boldsymbol{X})$, then uses a set of functions $m_k(\boldsymbol{X}, \widehat{\boldsymbol{X}})$ to evaluate the distances d_k between the image $\boldsymbol{X}$ and its filtered version $\widehat{\boldsymbol{X}}$, and finally provides a vector $\boldsymbol{d}$ composed of d_k to the classifier c. Similar to most common universal steganalysis methods, the classifier is trained using machine learning methods. The main points of this approach are the design of filters, distance measures, and classifiers. We will discuss them one by one in the following.

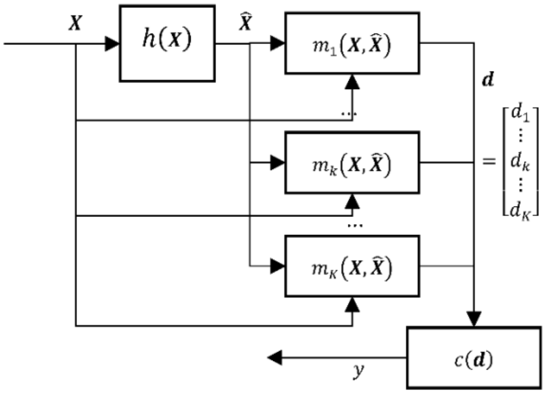
$\boldsymbol{X}$
$h(\boldsymbol{X})$
$\widehat{\boldsymbol{X}}$
$m_1(\boldsymbol{X}, \widehat{\boldsymbol{X}})$
$m_k(\boldsymbol{X}, \widehat{\boldsymbol{X}})$
$m_K(\boldsymbol{X}, \widehat{\boldsymbol{X}})$
$\boldsymbol{d}$
$= \begin{bmatrix} d_1 \\ \vdots \\ d_k \\ \vdots \\ d_K \end{bmatrix}$
$c(\boldsymbol{d})$
y

Figure 1: The mechanism of the proposed method. An estimated image $\widehat{\boldsymbol{X}}$ is got by filtering the image $\boldsymbol{X}$ to be detected. The image quality metrics $\boldsymbol{d_k}$ between them are used as the feature vector $\boldsymbol{d}$ that feed to the classifier $\boldsymbol{c}(.)$

2.1 The filters

The choice of filters is important to the performance of the method. A good filter preserves natural image characteristics as much as possible while removing steganographic noise, which helps the algorithm achieve higher accuracy. In addition to the basic filters, such as the Gaussian and Wiener filter, two more types of filters are used. One of them is a total variation (TV) [Wahid and Lee (2017)] filter called the TV filter. The total variation denoising technique is based on the fact that the noise image has a higher TV value. Therefore, image denoising can be achieved by minimizing the TV value of the image. The TV filter is chosen because it is not only valid for removing noise, but also good at preserving image edge properties.

The other type is SRM-based filter. The SRM [Fridrich and Kodovsky (2012)] uses the high pass filters to get residuals. Since these residuals are very useful for capturing discontinuities introduced by steganography methods, the first- and second-order filters are also used in this proposed approach. For example, according to the SRM second-order residual, which is defined as $R(i,j) = \big(\boldsymbol{X}(i,j+1) + \boldsymbol{X}(i,j-1)\big) - 2\boldsymbol{X}(i,j)$, this method designs and uses the filter defined as $h\big(\boldsymbol{X}(i,j)\big) = \big(\boldsymbol{X}(i,j+1) + \boldsymbol{X}(i,j-1)\big)/2$. The nonlinear filters that taking the minimum (or maximum) of the linear filters' outputs, such as $h\big(\boldsymbol{X}(i,j)\big) = min\Big(h_1\big(\boldsymbol{X}(i,j)\big), h_2\big(\boldsymbol{X}(i,j)\big)\Big) = min\big(\big(\boldsymbol{X}(i,j+1) + \boldsymbol{X}(i,j-1)\big)/2\,,$ $\big(\boldsymbol{X}(i+1,j) + \boldsymbol{X}(i-1,j)\big)/2\big)$, are utilized too. Some filters have four direction forms that they start from the initial forms and rotate counterclockwise in step of 90 degrees until back to the beginning position. For example, the four direction forms of $h\big(\boldsymbol{X}(i,j)\big) = \boldsymbol{X}(i+1,j)areh\big(\boldsymbol{X}(i,j)\big) = \boldsymbol{X}(i,j+1)$, $h\big(\boldsymbol{X}(i,j)\big) = \boldsymbol{X}(i-1,j)$, $h\big(\boldsymbol{X}(i,j)\big) = \boldsymbol{X}(i,j-1)$, and $h\big(\boldsymbol{X}(i,j)\big) = \boldsymbol{X}(i+1,j)$. As a result, this approach uses 71 SRM-based filters in total.

2.2 The distance measures

Embedding messages in the carrier images increases their degree of discontinuity, which means the distances between a steganographic image and its filtered one are greater than the distances between a natural and its filtered version. Let $\widehat{\boldsymbol{X}} = h(\boldsymbol{X})$ be the filtered image of an image X, the distances are evaluated using the image quality metrics [Tang, Li, Luo et al. (2016)] $d_k = m_k\big(\boldsymbol{X}, \widehat{\boldsymbol{X}}\big)$ listed in Tab. 1. These quality metrics are selected according to their accuracies, consistencies and monotonicity.

Table 1: Image quality measures

$M_1 = \left\{\frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left\lvert X(i,j) - \hat{X}(i,j)\right\rvert\right\}$,
$M_2 = \left\{\frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left\lvert X(i,j) - \hat{X}(i,j)\right\rvert^2\right\}^{1/2}$,
$M_3 = \frac{1}{MN}\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(1 - \frac{2min\left(X(i,j),\hat{X}(i,j)\right)}{X(i,j)+\hat{X}(i,j)}\right)$,
$M_4 = 1 - \frac{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(X(i,j)-\hat{X}(i,j)\right)^2}{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(X(i,j)\right)^2},k$
$M_5 = \frac{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(X(i,j)\hat{X}(i,j)\right)}{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(X(i,j)\right)^2}$,
$M_6 = \frac{1}{MN}\sum_{p=0}^{M-1}\sum_{q=0}^{N-1}\left(\lvert\Gamma(p,q)\rvert - \left\lvert\hat{\Gamma}(p,q)\right\rvert\right)^2$, $\Gamma(p,q)$ is the discrete Fourier transform (DCT) of $X(i,j)$.
$M_7 = \underset{l=0\cdots-1}{\text{median}}\,\mathrm{J}_{\varphi}^{l}$, $M_8 = \underset{l=0\cdots L-1}{\text{median}}\,J^l$, $\mathrm{J}_M^l = \left(\sum_{p=0}^{M_b-1}\sum_{q=0}^{N_b-1}\left(\lvert\Gamma^l(p,q)\rvert - \left\lvert\hat{\Gamma}^l(p,q)\right\rvert\right)^2\right)^{1/2}$, $\mathrm{J}_{\varphi}^l =$ $\left(\sum_{p=0}^{M_b-1}\sum_{q=0}^{N_b-1}(\lvert\varphi^l(p,q)\rvert - \lvert\varphi^l(p,q)\rvert)^2\right)^{1/2}$, $J^l = \lambda\mathrm{J}_{\mathrm{M}}^l + (1-\lambda)\mathrm{J}_{\varphi}^l$, where $\Gamma^l(p,q)$is the DCT of the *l*-th 32-by-32 image block $X^l(i,j)$, $\lvert\Gamma^l(p,q)\rvert$ and

$\varphi^{l}(p,q)$are the phase and magnitude spectra.

$$M_9 = \frac{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(X_f(i,j) - \hat{X}_f(i,j)\right)^2}{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}\left(X_f(i,j)\right)^2}$$

Where $D(p,q)$ denotes the discrete cosine transform (DCT) of the signal $X(i,j)$, $D_f(p,q) = D(p,q)h(\rho), \rho = (p^2 + q^2)^{0.5}$ is its band-pass filtered version, and $\mathrm{H}(\rho) = \begin{cases} 0.05e^{\rho^{0.554}} & \rho < 7 \\ e^{-9|\log_{10}\rho - \log_{10}\rho 9|^{2.3}} & \rho \geq 7 \end{cases}$, $X_f(i,j)$ is the inverse DCT of $D_f(p,q)$.

2.3 The Classifying method

The support vector machine (SVM) is an effective classification method for small data sets. The main idea is to transform the original input data set into a high-dimensional feature space by using a kernel function and then to optimize the classification in this new feature space under the assumption of linear separability. Because of the small number of features that the proposed algorithm uses, a SVM classifier is a suitable classifier.

The optimal parameters of the SVM is obtained by grid searching, which is time-consuming when the searching range is large. Therefore, the method also examines the ensemble classifier. We use the same ensemble classifier as the SRM does, which is a random forest. It consists of many fisher linear discriminants (FLDs) that work as base learners, and each of them is trained on a randomly chosen different-dimensional subspace of the feature space. The final decision of the classifier is made using majority voting according to the decisions of all base learners. Both of the two classifiers can work well and should be chosen according to the application environment.

2.4 The working modes

The proposed method works in two modes, a single filter mode and multi-filter mode. Among all of the aforementioned filters, the single filter mode selects the TV filter because of its ability of removing steganographic noise and preserving the image edge properties at the same time. As using only one filter, this mode uses a 9-dimensinal feature vector. In contrast to the single filter mode, the multi-filter mode uses all these filters because they each have their own advantages. This mode combines all the measures corresponding to each filter. As there are 71 filters, each of which corresponding to 9 measures, therefore the multi-filter mode uses a 639-dimensional feature vector. Both modes can use either the ensemble classifier or the SVM classifier.

3 Simulation results and analysis

In order to study the performance of all the candidate filters, four typical space domain steganography methods and one transform domain steganography method were used. They were LSB, EA [Luo, Huang and Huang (2010)], syndrome-trellis codes (STC) [Filler, Judas and Fridrich (2011)], HUGO [Pevný, Filler and Bas (2010)], and QIM [Chen and Wornell (2001)]. EA may be the first steganography method that introduce content adaptive design idea. HUGO improves EA and fully utilizes the knowledge about the steganalysis methods. STC goes further and becomes as the foundation of the current most secure steganography methods, which are content adaptive methods including WOW [Holub and Fridrich (2012)], S-UNIWARD [Holub, Fridrich and Denemark (2014)] and etc. Therefore, we chose HUGO, WOW and S-UNIWARD to verify the performance of the proposed method.

The carrier images were chosen from the BOSSbase version 1.01 database [Filler, Pevný and Bas (2010)]. For each embedding rate and steganography method, 500 natural images and their corresponding steganographic images were prepared, which yielded a total of 1000 samples per steganography method per payload. These samples were divided into two equal groups, one for training and the other one for testing.

We used a SVM and ensemble classifier. The SVM classifier used a RBF kernel, and the RBF scaling factor γ and the box constraint C were optimized by grid search. The searching range for γ was e^{-10} to e^{10} with a step of $\mathrm{e}^{0.1}$, and for C, it was e^{0} to e^{10} with a step of $\mathrm{e}^{0.1}$.

The steganographic images were generated with a series of different embedding rates. The embedding rate is defined as l_m/l_s in units of bits per pixel (bpp), here, l_m is the length of hidden message, and l_s is the total number of the cover image pixels. In addition, the images produced by different steganography methods have different transparencies even at the same embedding rates, and therefore the pitch signal-to-noise ratio (PSNR) was also used to analyze the capability of our method. We also used the accuracy rate (AR) for evaluation. AR is defined as $(n_c + \mathrm{n}_s)/n$, where n is the total number of the samples, and, n_c and n_s are the number of the natural and steganographic images, respectively, correctly classified.

3.1 Performance analysis of the filters

Different filters are sensitive to different types of steganographic noise. To compare their performance, we used one filter at a time to construct detectors in turn. We studied the accuracy rates of the detectors using the Gaussian, Wiener, TV, first-order, and second-order filter, respectively. In order to further examine the performance of these filters, we analyzed the simulation result from two perspectives, that is, the relationship between transparency and detection accuracy, and the relationship between a steganography method and detection accuracy.

We sorted the detection accuracy in the ascending order of the PSNRs. Fig. 2 clearly shows the relationship between transparency and detection accuracy. The results of the detectors using the Gaussian, Wiener and TV filters are shown in Fig. 2(a), and those of the detectors using the first-order, second-order, and TV filters are shown in Fig. 2(b). In general, a small PSNR means a high embedding ratio, so the smaller the PSNR, the higher the detection accuracy. In particular, the detector using the TV filter had high detection accuracy even at low embedding rates. This good detection performance mainly benefited from the fact that TV filters introduce very little secondary noise when removing steganographic noise. The detectors that use the first- or second-order filters were less accurate than that using the TV filter but were more accurate than those that use the Gaussian or the Wiener filter.

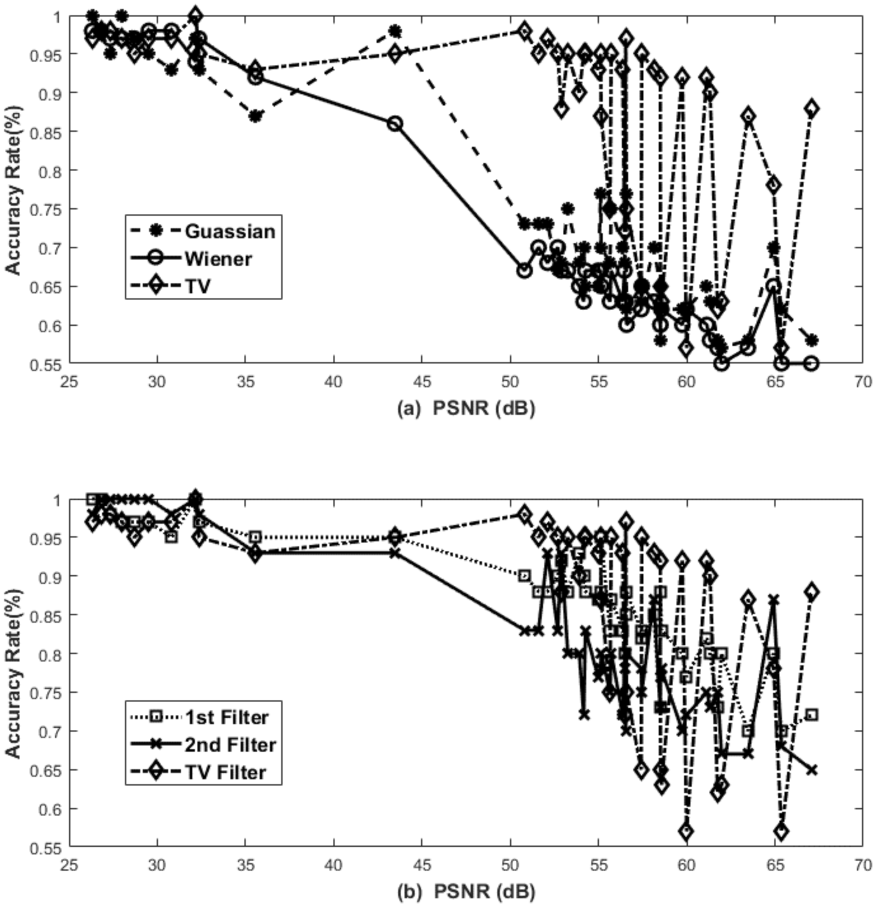

Figure 2: The relationship between the transparency and accuracy. The higher the peak signal-to-noise ratio between the nature and steganographic images were, the less the effect on the nature images were casted by the steganography methods, and therefore the more difficult the detection was. Among all detectors, the detector using the TV filter achieved the best detection accuracy

To examine the relationship between a steganography algorithm and detection accuracy, we compared the detection performance of these detectors, as shown in Fig. 3. The steganography methods using special techniques such as the matrix coding generated images with higher PSNRs than those generated by the other methods, and therefore were more difficult to detect. For the LSB, EA, HUGO, and STC algorithms, the first three most accurate detectors were the ones using the TV, first-, and second-order filters.

In conclusion, the detector using a TV filter was capable to discover all these steganography methods, and using multiple types of filters could improve the performance.

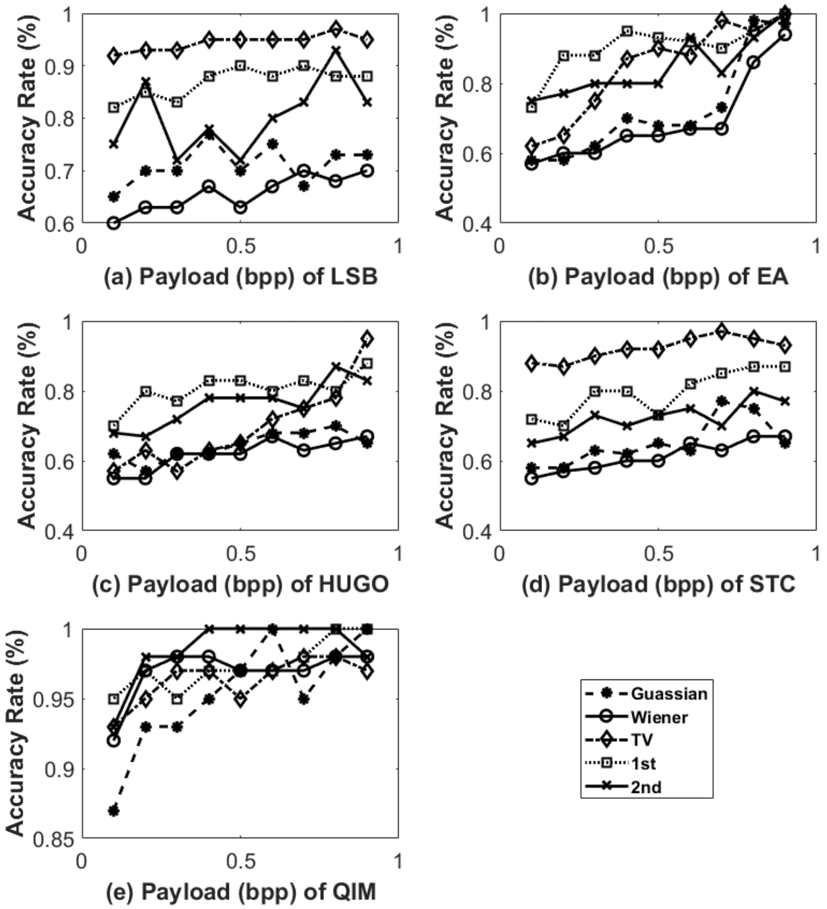

Figure 3: The relationships between the accuracy and steganography methods. The detection accuracy was related with the embedding payload and methods. The higher the embedding payload was, the easier the detection was. The most challenging steganography method was the HUGO. On most conditions, the first three most accurate methods separately used the TV, 1st- and 2nd-filter

3.2 Performance comparison of the steganalysis methods

We compared the proposed method with the SRM and maxSRMd2. The single filter mode and multi-filter mode of this method use 9 and 639 features respectively. For convenience, in Tab. 2, we denoted the single filter mode, the multi-filter mode, the SRM and the maxSRMd2 as A1, A2, A3, and A4 respectively. Here, the SVM classifier was used for the single filter mode, and the ensemble classifier [Kodovsky, Fridrich and Holub (2012)] was used in the multi-filter mode. In addition, the performance data for the SRM and maxSRMd2 is quoted from the literature [Tang, Li, Luo et al. (2015)].

Table 2: The accuracy comparison

bpp	WOW				S-UNIWARD				HUGO			
	A1	A2	A3	A4	A1	A2	A3	A4	A1	A2	A3	A4
0.1	0.50	0.79	0.59	0.69	0.92	0.94	0.59	0.63	0.54	0.90	0.64	0.69
0.2	0.50	0.90	0.68	0.76	0.95	0.95	0.68	0.71	0.56	0.94	0.72	0.76
0.3	0.54	0.94	0.74	0.80	0.68	0.96	0.74	0.75	0.94	0.95	0.78	0.80
0.4	0.85	0.95	0.79	0.83	0.93	0.96	0.79	0.80	0.95	0.97	0.82	0.84

According to the content adaptive steganalysis techniques, the maxSRMd2 assigns higher weights to those complex texture areas. As a result, the maxSRMd2 was more accurate than the SRM especially at the low embedding rate. Owning to the combining the advantages of the SRM and the IQM, for the HUGO, WOW, and S-UNIWARD algorithms, the multi-filter mode was more accurate than the SRM and maxSRMd2. As only one filter is used, the single filter mode poorly performed at 0.3 bpp or lower embedding rate.

At the embedding rate of 0.1 bpp, the multi-filter mode detected all the steganography methods with the accuracy of at least 79%. The single filter mode could accurately detect the HUGO algorithm with a capacity above 0.3 bpp, and the WOW algorithm with a capacity above 0.4 bpp. The result shows the single filter mode is suitable for computing resource-limited environments, while the multi-filter mode is suitable for applications that require higher detection accuracy.

3.3 Effects analysis of data normalization

Data normalization is necessary and important. In the case of a large variation in the dynamic ranges of the different features, normalization makes each feature work instead of being masked due to its smaller dynamic range. Second, normalization can improve the convergence rate of the gradient descent algorithm. When normalizing, we adjusted a variable $x \in [a, b]$ into the new range $[c, d]$ using the linear function $\mathrm{g}(x) = \mathrm{c} + \frac{\mathrm{b-a}}{c-d}(x - \mathrm{a})$, where the new dynamic range $[c, d]$ was $[0, 1]$.

According the range of data normalization, three methods were used. The first method normalized the each feature among the whole data set and then randomly segmented the normalized features into the training and testing set. The first method is actually a global normalization method, so the result obtained by this method is denoted as GN in Tab. 3. The second method normalized the each feature among the training set, and recorded their maximum and minimum values. These values were then used to normalize the features among the testing set. The third method normalized the each feature among the training set and testing set respectively. These two methods are both local normalization methods, and the second method has a more stable and superior performance. Therefore, only the result under the second method is listed in Tab. 3 and denoted as LN.

It is shown in Tab. 3 that normalizing features could improve the accuracy a lot and global normalization performs better than local normalization. However, the second normalization method still performs better than the SRM, and it is easier to implement in actual operation.

Table 3: The effect of the normalization methods

bpp	WOW			S-UNIWARD			HUGO		
	GN	LN	SRM	GN	LN	SRM	GN	LN	SRM
0.1	0.79	0.83	0.59	0.94	0.82	0.59	0.90	0.82	0.64
0.2	0.90	0.77	0.68	0.95	0.83	0.68	0.94	0.76	0.72
0.3	0.94	0.78	0.74	0.96	0.79	0.74	0.95	0.76	0.78
0.4	0.95	0.80	0.79	0.96	0.80	0.79	0.97	0.70	0.82

4 Conclusion

This paper presented a novel steganalysis method based on the IQM and the SRM. Compared with the SRM, the IQM uses the distances between the image and its filtered version as the features, which are actually various weighted multi-domain accumulative sum of the residuals and are more conducive to reflecting the steganographic noise.

Because of the fact that the total variation technique introduces less secondary noise when removing steganographic noise, the single filter mode only uses the TV filter. This mode therefore uses only 9 features but achieved similar performance to the SRM. This characteristic helps to extend the application of steganalysis to the computing resource-limited environments.

Inspired by the SRM, the multi-filter mode uses multiple types of filters including the first- and second-order filters used by the SRM, and combines respective outputting measure as the feature vector. Although the number of the features thus increases to 639, a higher detection accuracy is achieved, resulting in an average accuracy of higher than 79% even at embedding rates as low as 0.1 bpp.

The data normalization method heavily influents the performance of the proposed method especially for the content adaptive steganography method at low embedding rate. Future study will focus on improving data normalization method and optimizing the proposed method with content adaptive steganalysis techniques.

Acknowledgement: This work is supported by the National Key R&D Program of China (2017YFB0802703), the Open Foundation of Guizhou Provincial Key Laboratory of Public Big Data (2017BDKFJJ017) and CCF-Venustech Hongyan Research Initiative (2017-004).

References

Avcibaş, I.; Memon, N.; Sankur, B. (2003): Steganalysis using image quality metrics. *IEEE Transactions on Image Processing*, vol. 12, no. 2, pp. 221-229.

Broda, M.; Levicky, D.; Banoci, V.; Bugar, G. (2014): Universal imagesteganalytic method based on binary similarity measures. *Radioelektronika*, pp. 1-4.

Chen, B.; Wornell, G. W. (2001): Quantization index modulation: A class of provably good methods for digital watermarking and information embedding. *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1423-1443.

Chen, C.; Shi, Y. Q. (2008): JPEG image steganalysis utilizing both intrablock and interblock correlations. *IEEE International Symposium on Circuits and Systems*, pp. 3029-3032.

Filler, T.; Judas, J.; Fridrich, J. (2011): Minimizing additive distortion in steganography using syndrome-trellis codes. *IEEE Transactions on Information Forensics & Security*, vol. 6, no. 3, pp. 920-935.

Filler, T.; Pevný, T.; Bas, P. (2010): BOSS (Break Our Steganography System) 2010. http://agents.fel.cvut.cz/stegodata/.

Fridrich, J.; Kodovsky, J. (2012): Rich models for steganalysis of digital images. *IEEE Transactions on Information Forensics & Security*, vol. 7, no. 3, pp. 868-882.

Geetha, S.; Sindhu, S. S. S.; Kamaraj, N. (2008): Stego-Breaker: Defeating the steganographic systems through Genetic-X-Means approach using image quality metrics. *16th International Conference on Advanced Computing and Communications*, pp. 382-391.

Geetha, S.; Sindhu, S. S. S.; Kamaraj, N. (2007): Evolving GA classifier for breaking the steganographic utilities: Stools, Steganos and Jsteg. *International Conference on Computational Intelligence and Multimedia Applications*, pp. 230-234.

Holub, V.; Fridrich, J.; Denemark, T. (2014): Universal distortion function for steganography in an arbitrary domain. *EURASIP Journal on Information Security*, vol. 2014, no. 1, pp. 1-3.

Holub, V.; Fridrich, J. (2012): Designing steganographic distortion using directional filters. *IEEE International Workshop on Information Forensics and Security*, pp. 234-239.

Ker, A. D. (2007a): A fusion of maximum likelihood and structural steganalysis. *IH'07 Proceedings of the 9th International Conference on Information Hiding*, pp. 204-219.

Ker, A. D. (2007b): Steganalysis of embedding in two least-significant bits. *IEEE Transactions on Information Forensics & Security*, vol. 2, no. 1, pp. 46-54.

Ker, A. D.; Lubenko, I. (2009): Feature reduction and payload location with WAM steganalysis. *Media Forensics and Security I, Part of the IS&T-SPIE Electronic Imaging Symposium*, pp. 72540-72552.

Kodovsky, J.; Fridrich, J.; Holub, V. (2012): Ensemble classifiers for steganalysis of digital media. *IEEE Transactions on Information Forensics & Security*, vol. 7, no. 2, pp. 432-444.

Lei, M.; Yang, Y. X.; Niu, X. X.; Yang, Y.; Hao, J. (2017): An overview of general theory of security. *China Communications*, vol. 14, no. 7, pp. 1-10.

Li, X.; Zeng, T.; Yang, B. (2008): Detecting LSB matching by applying calibration technique for difference image. *10th ACM Workshop on Multimedia and Security*, pp. 133-138.

Lin, J. Q.; Zhong, S. P. (2009): JPEG image steganalysis method based on binary similarity measures. *International Conference on Machine Learning and Cybernetics*, pp. 2238-2243.

Luo, W.; Huang, F.; Huang, J. (2010): Edge adaptive image steganography based on LSB matching revisited. *IEEE Transactions on Information Forensics & Security*, vol. 5, no. 2, pp. 201-214.

Pevný, T.; Bas, P.; Fridrich, J. (2010): Steganalysis by subtractive pixel adjacency matrix. *IEEE Transactions on Information Forensics & Security*, vol. 5, no. 2, pp. 215-224.

Pevny, T.; Fridrich, J. (2007): Merging Markov and DCT features for multi-class JPEG steganalysis. *Security, Steganography, and Watermarking of Multimedia Contents IX International Society for Optics and Photonics*, vol. 6505, pp. 31-314.

Pevný, T.; Filler, T.; Bas, P. (2010): Using high-dimensional image models to perform highly undetectable steganography. *IH'10 the 12th International Conference on Information Hiding*, pp. 161-177.

Tan, S.; Li, B. (2014): Stacked convolutional auto-encoders for steganalysis of digital images. *Signal and Information Processing Association Annual Summit and Conference*, pp. 1-4.

Tang, W.; Li, H.; Luo, W.; Huang, J. (2016): Adaptive steganalysis based on embedding probabilities of pixels. *IEEE Transactions on Information Forensics & Security*, vol. 11, no. 4, pp. 734-745.

Wahid, A.; Lee, H. J. (2017): Image denoising method based on directional total variation filtering. *International Conference on Information and Communication Technology Convergence*, pp. 798-802.

Xia, Z.; Wang, X.; Sun, X.; Liu, Q.; Xiong, N. (2016): Steganalysis of LSB matching using differences between nonadjacent pixels. *Multimedia Tools & Applications*, vol. 75, no. 4, pp. 1947-1962.

Xia, Z.; Wang, X.; Sun, X.; Wang, B. (2014): Steganalysis of least significant bit matching using multi-order differences. *Security & Communication Networks*, vol. 7, no. 8, pp. 1283-1291.

Xu, B.; Wang, J.; Liu, X.; Zhang, Z. (2007): Passive steganalysis using image quality metrics and multi-class support vector machine. *Third International Conference on Natural Computation*, pp. 215-220.

Xu, G.; Wu, H. Z.; Shi, Y. Q. (2016): Structural design of convolutional neural networks for steganalysis. *IEEE Signal Processing Letters*, vol. 23, no. 5, pp. 708-712.

Yang, Y.; Lei, M.; Liu, X.; Qu, Z.; Wang, C. (2016): Novel zero-watermarking scheme based on DWT-DCT. *China Communications*, vol. 13, no. 7, pp. 122-126.

Ye, J.; Ni, J.; Yi, Y. (2017): Deep learning hierarchical representations for image steganalysis. *IEEE Transactions on Information Forensics & Security*, vol. 12, no. 11, pp. 2545-2557.

Yu, J.; Li, F.; Cheng, H.; Zhang, X. (2016): Spatial steganalysis using contrast of residuals. *IEEE Signal Processing Letters*, vol. 23, no. 7, pp. 989-992.

Zeng, J.; Tan, S.; Li, B.; Huang, J. (2018): Large-scale JPEG image steganalysis using hybrid deep-learning framework. *IEEE Transactions on Information Forensics & Security*, vol. 13, no. 5, pp. 1200-1214.