

A Method of Universal Steganalysis Using Independence with Image Format

QiuyuZhang*, QichangShang, Ruihong Dong, YanYan, HangzhouZuo

AbstractA new universal steganalysis method based on statistical characteristic of multi-domain features extraction is proposed in this paper, mainly aimed at detecting hidden information in images with multiple common formats. Features are firstly extracted in contourlet domain and extended to spatial domain afterwards, by calculating correlation between DCT (discrete cosine transform) coefficients using joint probability density and calculating distribution of coefficients in image using co-occurrence matrix. The experimental results shows that the proposed method has better detection effects on images with different steganography carriers, and achieves higher independence of image formats and better average detection effects compared with typical universal steganalysis algorithms at present.

KeywordsUniversal steganalysis Contourlet Joint probability density Multi-domain feature Independence of image formats

1 Introduction

With the rapid development of computer networks and digital media technology, a large number of steganography methods emerge and people can easily obtain and use a variety of steganography tools. However, the abuse of these tools is becoming a serious threat to network security. Therefore, steganalysis is gaining more and more attention. Steganalysis methods are mainly divided into two categories: specific steganalysis [1-4] and universal steganalysis. Universal steganalysis algorithms highlight the prospects for future development from a practical point of view [5].

The theory of universal steganalysis was first proposed by Avicibas[6] etc. which is based on IQM (Image Quality Metrics); then Farid[7] and Shi[8] proposed different steganography methods based on single characteristic. Lie[9] pointed out that the single feature generally does not have the ability to distinguish

*QiuyuZhang (✉), QichangShang (✉)

School of Computer and Communication, Lanzhou University of Technology, Lanzhou, 730050, China

e-mail: zhangqylz@163.com; shangqichang@163.com

the hidden signal and the original signal effectively, and proposed blind detection methods combined with features of DCT domain and spatial domain. Meanwhile experiments were carried out on three kinds of LSB hiding schemes in spatial domain, three kinds of hiding schemes in DCT-domain and one hiding scheme in wavelet domain. In order to improve the rate of blind detection for hidden information in images, Li Zhuo[10]etc. proposed an universal steganalysis algorithm for JPEG images by extracting statistical features in multi-domain. In recent years, there has also been some progress for multi-carrier image detection. Luo[11] etc. made some improvements on Lie's works about feature extraction in spatial domain and combined it with DWT domain features. This method put forward a new thought of blind detection method for three domains and achieved detection of hidden information in images with both BMP and JPEG format using BP neural network as a classifier. Jessica Fridrich[12] proposed the Pairs Analysis method that can detect hidden information in different image formats.

Proposed method for universal image steganalysis in this paper is an improvement on Ref. [10,11] and combined with contourlet domain when extracting features, achieving the independence of image format. Experimental results show that the proposed universal steganalysis algorithm have a higher average detection rate than Ref. [11].

2 Feature Extractions

One of the key problems of steganalysis is feature extraction. As the same way, universal steganalysis is mainly to extract features to train classifiers. These features are sensitive to the embedded secret messages other than the content of the image. The feature extraction process of this paper is shown in Fig.2.1.

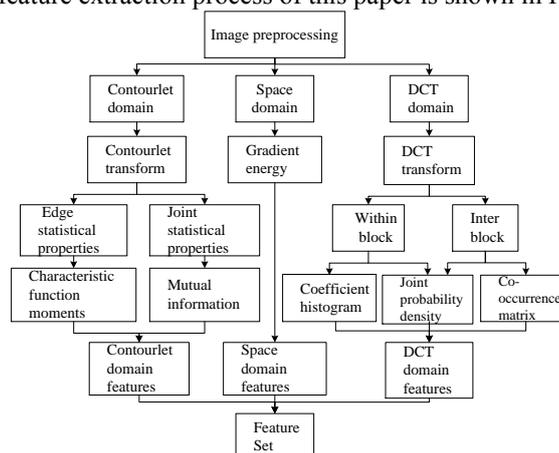


Fig.2.1 Feature extraction process

2.1 Based on the Statistical Feature of Wavelet Transform Contourlet

Contourlet transform is a kind of inseparable and real two-dimensional transformation, which is designed for natural image geometry. It has inbuilt advantage over wavelet transform for its flexibility, multi-direction and high efficiency, and potential. Contourlet transform is a multi-scale computational framework of discrete image[13].

For each contourlet coefficient X , the sub-band in the same direction has eight neighboring coefficients X_N , one parent coefficient X_P in same location within sub-band of last level and many cousin coefficients X_C in all directions with the same scale.

2.1.1 Edge of the Statistical Properties

After contourlet transform, Fig.2.2 shows the marginal distribution $p(X)$ of X and the joint probability distribution $p(X,S)$ of state $S=(S1,S2)$ and contour coefficients X , where $p(X|S1)$ and $p(X|S2)$ is Gaussian distribution with a mean value 0 and a variance value σ_1 and σ_2 where S_1 as outline and another S_2 as background[14].

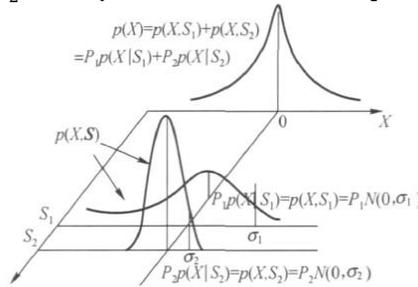


Fig.2.2 Edge of statistical properties of contourlet sub-band coefficients

If we denote Fourier transform of $p(X)$ by $F(x)$ and marginal distribution of X by $p(X)$, then the equation will be shown as formula (2.1):

$$p(X) = \int_{-\infty}^{\infty} F(x) e^{-i2\pi x} dx \quad (2.1)$$

Results are calculated by performing an n-order derivation function and using an absolute value where $X=0$.

$$\left| \frac{d^n}{dx^n} p(X) \right|_{X=0} \leq 2(2\pi)^n \int_0^{\infty} x^n |F(x)| dx \quad (2.2)$$

According to the characteristic of Fourier transform, equality holds in formula (2.2) at other points. After embedding secret messages, Marginal distribution of contourlet coefficients will be lower and this change can be used to detect secret messages. Obviously, n-order moment of characteristic function is the upper bound of n-order derivation of marginal probability distribution, which means moment of characteristic function is more effective than statistical properties of

the edge contourlet coefficients when capturing the changes of statistical properties of the edge, caused by embedding secret messages.

2.1.2 Combined Statistical Properties

Statistical properties of edge only describe statistical characteristics rather than dependence of contourlet coefficients. To measure dependence of contourlet coefficients quantitatively, the joint probability distribution $p(x, y)$ is needed to calculate mutual information between the coefficients X and Y in formula (2.3) represent two random variables of contour coefficients.

$$I(X, Y) = \iint p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx dy \quad (2.3)$$

In most cases, there is a relationship as $I(X, X_N) > I(X, X_C) > I(X, X_P)$ which is determined by their quantity and dependence with X . But in fact result is on contrary after comparing the average mutual information as X and X_P hold the most close relationship of parents, for which reason X_P becomes best predictor of X . Cross-scale joint distribution at same location of Contourlet coefficients X and X_P is calculated using Bayes' formula as:

$$p(X, X_P) = p(X_P)p(X|X_P) \quad (2.4)$$

Mutual information is calculated as formula (2.4) and describes the dependence by calculating X and X_P . Decrement of the mutual information between X and X_P after embedding secret messages can be used to detect the existence of secret messages.

2.2 DCT Domain Feature Extraction

2.2.1 Within the Block Correlation

If we denote the size of test image by $M \times N$, F will represent the two-dimensional matrix formed by the absolute value of DCT coefficients. The matrix is divided into 8×8 blocks without overlapping, which are noted by $F_i (i \in [1, \lfloor M/8 \rfloor \times \lfloor N/8 \rfloor])$. Difference between neighboring coefficients in horizontal, vertical, diagonal and anti-diagonal direction in each block is calculated as follows, generating four matrixes of difference values.

$$\begin{aligned} F_h(u, v) &= F(u, v) - F(u, v + 1) \\ F_v(u, v) &= F(u, v) - F(u + 1, v) \\ F_d(u, v) &= F(u, v) - F(u + 1, v + 1) \\ F_m(u, v) &= F(u, v + 1) - F(u + 1, v) \end{aligned} \quad (2.5)$$

Including: $u \in [1, 7]$; $v \in [1, 7]$

According to the formula (2.5) can be calculated by the coefficients of adjacent joint probability density matrix. Calculated as follows:

$$\begin{aligned}
M_h(i, j) &= \frac{1}{(M-1) \times (N-2)} \sum_{u=1}^{M-1} \sum_{v=1}^{N-2} \varphi(F_h(u, v) = i, F_h(u, v+1) = j) \\
M_v(i, j) &= \frac{1}{(M-2) \times (N-1)} \sum_{u=1}^{M-2} \sum_{v=1}^{N-1} \varphi(F_v(u, v) = i, F_v(u+1, v) = j) \quad (2.6) \\
M_d(i, j) &= \frac{1}{(M-2) \times (N-2)} \sum_{u=1}^{M-2} \sum_{v=1}^{N-2} \varphi(F_d(u, v) = i, F_d(u+1, v+1) = j) \\
M_m(i, j) &= \frac{1}{(M-2) \times (N-2)} \sum_{u=1}^{M-2} \sum_{v=1}^{N-2} \varphi(F_h(u, v+1) = i, F_h(u+1, v) = j)
\end{aligned}$$

General form of $\varphi(F_h(u, v) = i, F_h(u, v+1) = j)$ is denoted by $\varphi(x, y)$. $\varphi(x, y)$ will be equal to 1 while $x=y$; otherwise it will be zero. Feature vector is formed by mean value of all joint probability density matrixes.

Single DCT coefficients are regularly distributed and can be affected by embedding secret messages. It will be declared whether images are embedded through distribution of DCT coefficients among 64 DCT blocks. Owing to over 96.59% of AC coefficients range from -7 to 7, for which part we take into consideration in the proposed method. Calculated as follow:

$$H_d(i, j) = \sum_{r=1}^{M/8} \sum_{l=1}^{N/8} \varphi(d, d_{r,l}(i, j)) \quad (2.7)$$

In formula (2.7) $d_{r,l}(i, j)$ represents coefficient of i -th row and j -th column in DCT block (r, l) . Values of i and j can't be equal to 0 at same time, otherwise the first value of DCT coefficients $d_{r,l}(0,0)$ would be much larger than 7 as a DC coefficient. So do not statsthis value.

2.2.2 Correlation between the Blocks

Statistical properties are similar within the same block and calculated through the joint probability density matrix of neighboring coefficients within blocks. All AC coefficients with same frequency in 8×8 sized DCT coefficient blocks from F are translated to 63 new two-dimensional matrix (DC coefficient at point of (0,0) is removed). Feature vector is formed by mean value of all joint probability density matrixes as follow:

$$M = \frac{1}{63} \left[\sum_{i=1}^{63} (M_{ih} + M_{iv} + M_{id} + M_{im}) \right] \quad (2.8)$$

Embedded messages will affect continuity between neighboring coefficients of carrier images and thereby distribution of coefficients in the image. Co-occurrence matrix of coefficients at same location of neighboring DCT blocks can entirely de-

scribe these changes, for which reason it's used to determine whether images are embedded.

$$H(d_1, d_2) = \sum_{r=1}^{\frac{M}{8}-1} \sum_{l=1}^{N/8} \sum_{i,j=0}^7 \varphi(d_1, d_{r,l}(i,j)) \varphi(d_2, d_{r+1,l}(i,j)) + \sum_{l=1}^{\frac{N}{8}-1} \sum_{r=1}^{M/8} \sum_{i,j=0}^7 \varphi(d_1, d_{r,l}(i,j)) \varphi(d_2, d_{r,l+1}(i,j)) \quad (2.9)$$

Formula (2.9) is used to count distribution of coefficient pair numbered $(d1, d2)$ in 63 positions of each DCT block. Absolute value of this coefficient pair is set to $(0,0), (0,1), (1,1), (1,2), (2,3)$. Of which i and j can't be equal to 0 at same time.

2.3 Feature Extraction in Space Domain

Gradient as a space feature reflects the changing rate of signals and has significant changes when images are embedded. Extraction process is as follows:

$$\begin{aligned} G_V &= \frac{1}{N_H \cdot N_W} \sum_x \sum_y (I(x, y) - I(x, y - 1))^2 + \frac{1}{N_H \cdot N_W} \sum_x \sum_y (I(x, y + 1) - I(x, y))^2 \\ G_H &= \frac{1}{N_H \cdot N_W} \sum_x \sum_y (I(x, y) - I(x - 1, y))^2 + \frac{1}{N_H \cdot N_W} \sum_x \sum_y (I(x + 1, y) - I(x, y))^2 \\ G_D &= \frac{1}{N_H \cdot N_W} \sum_x \sum_y (I(x, y) - I(x - 1, y - 1))^2 + \frac{1}{N_H \cdot N_W} \sum_x \sum_y (I(x + 1, y + 1) - I(x, y))^2 \\ G_M &= \frac{1}{N_H \cdot N_W} \sum_x \sum_y (I(x, y) - I(x - 1, y + 1))^2 + \frac{1}{N_H \cdot N_W} \sum_x \sum_y (I(x + 1, y - 1) - I(x, y))^2 \end{aligned} \quad (2.10)$$

Where N_H and N_W represent the height and width of image.

According to the formula (2.10) can be calculated the total gradient energy, denoted by F as feature of spatial domain.

$$F = G_V + G_H + G_D + G_M \quad (2.11)$$

The proposed method of gradient is an essential analysis for images by smoothness of image space around process of steganography. Gradient energy of main and sub diagonal is taken into consideration to improve the accuracy of blind detection the algorithm.

3 Classifier Selections

The basic idea of Support Vector Machines [15] as a classifier is to separate two sample points correctly as much as possible using separating hyper-plane, while keeping them farthest from classification surface. This paper uses the LIBSVM as a classifier and a RBF (radial basis function) as the kernel function. The parameter g in kernel function and factor C mainly determine classification performance and therefore should be optimized. Here parameter C and g are based on theory of cross-validation ideological grid searching.

4 Experimental Results and Analysis

Experimental images are from standard image database of NRCS, UCID and USC-SIPI. 1004 of JPEG images are selected from NRCS database and 24 from standard image database; 1014 of BMP images are from NRCS database and 24 from standard image database; 1338 GIF images are from images of TIFF format in UCID database and converted to GIF format with a MATLAB disturbance script. Image contents are wide and various, including sceneries (trees, mountains, rivers, flowers, etc), artificial facilities (buildings, bridges, streets, cars, etc), personal portraits, animal close-up and so on. In order to calculate and test easily, all the images are cut into 512×512 pixels around center. Experiments contain two parts: individual and comprehensive experiments of training and testing. The former one refers to training with part of original images and steganography images, and detects steganography images with different steganography methods for the carrier of specific format; the latter refers to training with some original images and steganography images generated by different steganography methods with three kinds of carrier images in different formats, tested with other corresponding steganography images.

There are much steganography algorithms for carrier images in different formats, but only several typical ones are chosen to test for limitation in paper. For images in JPEG format, methods of JSteg, Jphide and Outguess are selected; for images in BMP format, methods of PMK, Stools and LSBare selected; for images in GIF format, methods of EzStego and F5 are selected. 2/3 steganography images and 2/3 carrier images are select randomly for training SVM and the remaining 1/3 ones are used as testing set.

Table 4.1 shows the experimental results of individual training and testing. The data in Table 4.1 indicates test results for carrier images in different formats. Due to the general property, the detection accuracy is lower than some universal steganalysis methods designed for single format, yet achieving effective detection.

Table 4.1 Separated detection result of steganalysis performance

Carrier format	Embedding method	Detection accuracy (%)					
		Embedding rate (50%)			Embedding rate (100%)		
		TPR	TNR	ACC	TPR	TNR	ACC
JPEG	JSteg	87.135	82.749	84.942	95.029	88.304	91.667
JPEG	Jphide	84.503	82.164	83.333	92.982	89.766	91.374
JPEG	Outguess	73.977	78.655	76.316	79.532	85.672	82.602
BMP	PMK	79.191	72.543	75.867	82.659	77.168	79.913
BMP	Stools	84.393	79.769	82.080	90.751	82.370	86.561
BMP	LSB	75.723	71.387	73.555	84.104	76.012	80.058
GIF	EzStego	68.386	62.780	65.583	74.215	70.628	72.422
GIF	F5	81.390	76.009	78.700	85.847	79.372	82.623

For comprehensive training and testing, this paper makes a comparison with Ref. [11] where it is able to detect two kinds of image formats. Experimental results are shown in Table 4.2.

Table 4.2 Comprehensive test results

Carrier format	Data embeddin- grate (%)	Detection accuracy (%)	
		Ref.[11]	This paper
JPEG	0.25	53.704	69.396
JPEG	0.5	60.331	73.977
JPEG	0.75	68.226	76.706
JPEG	1	76.316	82.066
BMP	0.25	65.510	72.736
BMP	0.5	73.121	76.012
BMP	0.75	79.190	81.888
BMP	1	82.466	86.127
GIF	0.25	—	63.789
GIF	0.5	—	67.040
GIF	0.75	—	72.422
GIF	1	—	77.915
Mean		63.239	75.006

For Jpeg format each one thirds of images are embedded using steganography methods of JSteg, Jphide and Outguess respectively and under embedding rate of 25%, 50%, 75% and 100%; Under the same embedding rate, for Bmp format each one thirds of images are embedded using steganography methods of PMK, Stools and LSB respectively; for GIF format each one seconds of images are embedded using steganography methods of EzStego and F5 respectively. Average results are illustrated in Fig.4.1 using a ROC curve.

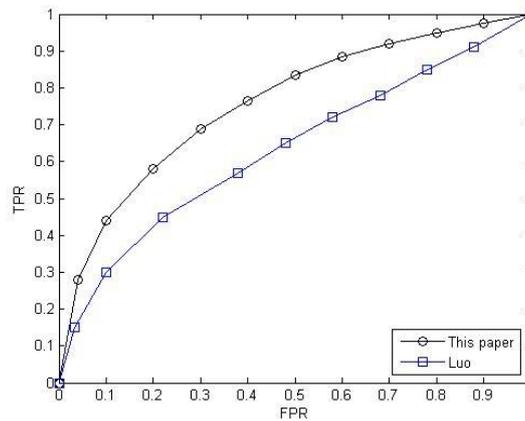


Fig.4.1 ROC curve of comprehensive test

Table 4.1 illustrates when detecting steganography of JSteg and Jphide on Jpeg images, the resulting rate upper than 90% is close to methods designed for single

steganography; the resulting rate stable around 82% when detecting three kinds of stego methods of BMP images certify the stability of the proposed method; the resulting rate of 77% when detecting GIF images certify the versatility of proposed method; Table 4.2 illustrates that for different steganalysis methods, the proposed method achieve about 12% higher effects than Ref. [11] under a small embedding rate using an image calibration function to sensitively reflect changes in statistical characteristics.

It can be seen in results of comprehensive tests that the proposed method can effectively detect images in formats of JPEG, BMP and GIF. When compared with Ref. [11], all of the testing results achieve better than the effect, especially in detection of GIF images. Reasons are that in Ref. [11] proposed method an image is decomposed into three dimensions by wavelet packet transform, and moments of the higher-order histogram features are extracted from these sub-bands. A contourlet transform function hold the same features of multi-resolution and time-frequency local characteristics with a wavelet transform function, as well as features of a directional selectivity and anisotropy which describe images better in nature; For most of steganalysis methods based on DCT domain, the proposed method describe the changes by correlation between the blocks of DCT coefficients when embedding secret messages; On the space domain, Ref. [11] didn't do any feature extraction.

It can be seen that proposed method in this paper captures the characteristics of the image from different angles and is more sensitive and general compared with other feature extraction, achieving easy classification and better results.

5 Conclusions

This paper proposes a universal steganalysis method for image in different formats. Statistical features are extracted from multi-domain based on changes caused by process of steganography. Consistency of contourlet domain features is first taken into account when extracting features from images in all formats. As embedding methods at present are mainly based on DCT domain, we use a co-occurrence matrix and correlation of DCT coefficients to describe changes of statistical features when embedding secret messages, combined with characteristics of spatial domain extracted from all carrier images. A support vector machine is used as a classifier in order to achieve the correct classification. Although the proposed steganalysis method has lower detection rate than ones designed for a single format, the initial realization of detection for different carrier formats is achieved.

Acknowledgments This work is supported by the National Natural Science Foundation of China (No. 61363078), the Natural Science Foundation of Gansu Province of China (No. 1212RJZA006). The authors would like to thank the anonymous reviewers for their helpful comments and suggestions.

6 References

1. Cogramne Rémi,Zitzmann Cathel, Retraint Florent (2012) Statistical detection of LSB matching in the presence of nuisance parameters.*IEEE Statistical Signal Processing Workshop*,Ann Arbor,Michigan,912-915.doi:10.1109/SSP.2012.6319857.
2. Ker, A. D. (2007) Steganalysis of embedding in two least-significant bits. *IEEE Transactions on Information Forensics and Security*, 2(1): 46-54.doi:10.1109/TIFS.2006.890519.
3. Pan Xiao-zhong,Zhao Shi-wei (2010) Anew steganalysis method for BPCS. *Computer Application and System Modeling*, Taiyuan,7:757-759.doi:10.1109/ICCASM.2010.5620237.
4. Zhang Xin-peng,Wang Shuo-zhong (2005)StatisticalAnalysisAgainstSpatialBPCSSSteganography.*Journal of Computer-Aided Design*,17(7):1626-1629(in Chinese).
5. Wang Shuo-zhong,Zhang Xin-peng,Zhang Wei-ming (2009) Recent Advances in Image-Based Steganalysis Research.*Journal of Computers*,2(7):1247-1263(in Chinese).
6. Avcibas,I., Memon, N.Sankur,B.(2003) Steganalysis using image quality metrics. *IEEE Transactions on Image Processing*, 12(2), 221-229.doi:10.1109/TIP.2002.807363.
7. Farid,H. (2000) Detecting steganographic messages in digital images.Technical Report, TR2001-412 USA: Dartmouth College.
8. Shi Y Q,Xuan G R,Yang C Y (2005) Effective Steganalysis Based on Statistical Moments of Wavelet Characteristic Function.*International Conference on Information Technology:Coding and Computing*, Las Vegas, Nevada,768-773.doi:10.1109/ITCC.2005.138.
9. Lie W N,Lin G S (2005) A Feature-Based Classification Technique for Blind Image Steganalysis.*Transactions on Multimedia*,7(6):1007-1020.doi:10.1109/TMM.2005.858377
10. Li Zhuo,Chen Jian (2011) Blind JPEG steganalysis based on multi-domain features.*Journal of Zhejiang University(Engineering Science)*,(9):1528-1538(in Chinese).
- 11.Luo Xiang-yang,Liu Fen-lin,Chen Jian-ming,Zhang Yi-ning (2008) Image universal steganalysis based on wavelet packet transform. *2008 IEEE 10th Workshop on Multimedia Signal Processing*, Cairns, Qld,780-784.doi:10.1109/MMSP.2008.4665180.
- 12.Fridrich Jessica,Goljan Miroslav,Soukal David (2003)Higher-order statistical steganalysis of palette images. *Proc. SPIE 5020Security and Watermarking of Multimedia Contents V*,Santa Clara, California, 178-190. doi:10.1117/12.473140.
- 13.Sajedi,Hedieh (2010) Contourlet-based steganalysis method. *Journal of Signal Processing Systems*,61(3):367-373.
14. Song Xiao-yang (2009) Contourlet Transform and HiddenMarkovModel in ContourletDomain. *Journal of Image andGraphics*, 14 (9):1721-1731(in Chinese).
- 15.CorinnaCortes,VladimirVapnik.(1995) Support-vector networks.*Machine Learning*, 20(3): 273-29. doi:10.1007/BF00994018