CrossMark

Unsupervised Optimization for Universal Spatial Image Steganalysis

Anita Christaline Johnvictor[1] · Ramesh Rangaswamy[2] ·
Gomathy Chidambaram[1] · Vaishali Durgamahanthi[1]

Abstract Blind universal steganalysis has been the choice of Steganalysers owing to it's capability to detect stego images without any prior information about the embedding method. Universal steganalysis is a two class optimization problem and the detecting efficiency depends on the feature set chosen from the stego and clean images. Though extracting all possible features of an image may lead to more efficiency the classification suffers due to large dimension of feature set. To overcome the problem of dimensionality appropriate feature reduction techniques need to be employed. This paper presents a blind universal image steganalysis technique that extracts the noise models of adjacent pixels of an image. The exact model construction involves the formation of four dimensional co-occurrence matrices of the quantised and truncated noise residues. From the 106 sub models 34,671 features have been extracted and further reduced by a novel unsupervised optimization technique to identify the most appropriate features for classification. The classifiers implemented include Support Vector Machines (SVM), Multi Layer Perceptron (MLP) and three fusion classifiers based on Bayes, Decision Template and Dempster Schafer fusion schemes. It has been identified that MLP performs better than SVM but is not superior to fusion classifiers. Comparing all the classifiers, Decision Template based fusion method gives the best classification accuracy (99.25%). Thus the proposed unsupervised optimization method combined with Decision Template fusion classification scheme provides the best classification of stego and clear images as compared to the existing research work.

✉ Anita Christaline Johnvictor
  anita.j@vdp.srmuniv.ac.in

  Ramesh Rangaswamy
  ramesh@saveetha.ac.in

  Gomathy Chidambaram
  hod.ece@vdp.srmuniv.ac.in

  Vaishali Durgamahanthi
  vaishali.b@vdp.srmuniv.ac.in

[1] Department of ECE, SRM University, Chennai, India

[2] Saveetha Engineering College, Chennai, India

Springer

1 Introduction

Steganography is the art of concealed communication in digital media. Image steganography has gained moment in the recent past as the very existence of the secret data is undetectable in an image. The strength of the steganographic technique lies in the property of statistical undetectability [1] of the secret information. Steganalysis is the identification of an image with secret information and is a two class optimization problem with the outcome as either stego image or cover image. Identifying the image features that were altered due to steganography is a crucial task as different steganographic methods alter different parameters of the image. Modern blind image steganalysis intends to extract as many features as possible to get a rich set of features and classify them. Though a rich data set is expected to give more accuracy in identifying stego images, they suffer from dimensionality problem. Hence feature set optimization to discover the pixel changes has become the prime requirement. Recent steganographic methods preserve the pixel dependencies of the cover image [2] and hence image steganalysis methods use increased number of pixel dependencies as feature vectors. In recent past, spatial domain steganalysis of images using SPAM features (686 features) have been used by Penvy et al. [2] to attack LSB matching steganographic schemes, cross domain features (1234 features) have been employed to attack the YASS [3]. Holub et al. state that while the scope of steganalysis with low dimensional feature set seems ineffective [3] the need for a large feature set has become inevitable.

This research work intends to capture all possible statistical dependencies among pixels in spatial domain as noise models. The low amplitude noise components have small dynamic range and hence increase the SNR between the stego image and the cover image. Statistically represented as joint probability distribution or co-occurrence matrices, these noise residues are modeled as multiple sub models of the images resulting in a large feature set. To overcome the problem of space and time complexity due to this large data set and to identify the specific pixel changes in stego images, an unsupervised clustering based optimization technique has been proposed. The reduced feature set is classified with individual classifiers (SVM and MLP) and a fusion classifier system with Bayes, Dempster Schafer and Decision template fusion schemes. The proposed method has been tested for Universal distortion method of steganography in spatial domain given by Holub et al. [4]. The next section of this paper discusses about the steganographic scheme in universal distortion method, followed by the feature extraction. Later the unsupervised optimization process is explained followed by the classification scheme along with the experimental results.

2 Steganographic Scheme

The lack of accurate model leads to different methods of steganographic embedding in images. Steganography in images pertain to JPEG and spatial domains. Spatial domain image steganography demands the embedding cost to be high in smooth areas and low in

noisy areas. Recent algorithms define the embedding distortion as normalized weighted values of higher order statistical differences of pixels [5–7]. Wavelet Obtained weights (WOW) intend to assess the image content in terms of aggregated directional residuals of high pass filters [5]. This wavelet based steganographic system has high resistance against rich model based steganalysis [6]. A modification of this WOW as the sum of relative changes between stego and cover images is presented as SUNIWARD (Spatial Universal Wavelet Relative Distortion) [4]. This steganographic scheme is used in this research to create the stego images. SUNIWARD intends to embed information in noisy or textured regions by quantifying the output with directional filter banks to resist steganalysis based on rich models. The additive approximation of SUNIWARD is considered in this paper as it uses stabilized filter banks. The choice of the directional filters and the stabilization constant determines the best setting for the SUNIWARD. The directional filter bank is constructed as a set of three linear shift invariant filters with kernels K1, K2, K3 which evaluate the directional residual (smoothness) of an image (I) in three directions, viz. horizontal, diagonal and vertical. The residual has $n_1 \times n_2$ elements as it is the mirror padded convolution of the kernel with the image I. The kernels are obtained from the one dimensional wavelet decomposition filters (low pass—h and high pass—g) [4] as below,

$$K_1 = h \cdot g^T; \quad k_2 = g \cdot h^T; \quad K_3 = g \cdot g^T \tag{1}$$

The choice of the wavelet based directional filters provides good de-correlation and energy compactification when their residuals coincide with the undecimated (1st level) wavelet decomposition of an image into the two dimensional LH, HH and HL sub bands. The steganographic distortion function used is the sum of relative changes of all wavelet coefficients given by,

$$D(I,J) = \sum \sum \sum (|Wxy(I) - Wxy(J)|)/(\propto + |Wxy(I)|) \tag{2}$$

where '$\propto$' is the stabilization constant and is always non negative. The distortion $D(I,J)$ is small when the wavelet coefficient of cover image (I) is large compared to that of the stego image (J). This occurs in the edges and textured regions. Apart from stabilizing the numerical computation, $\propto$ also affects the content adaptivity of the embedding algorithm. A small value of $\propto$ leads to high sensitivity and undesirable embedding change probabilities. A large value of $\propto$ gives smooth embedding change probabilities, enabling SUNIWARD to embed into cover image. These situations prompt the choice of a moderate value for $\propto$. For an embedding or distortion function, the cost involved in changing an element of cover image from I_{xy} to J_{xy} is $C_{xy}(I, J_{xy}) \triangleq D(I, I_{xy} J_{xy})$. The distortion function depends on Syndrome Trellis Code which is a standard tool in steganography. When $I = J$, $C_{xy} = 0$. The additive approximation of Eq. (2) based on this distortion is,

$$D_a(I,J) = \sum \sum C_{xy}(I, J_{xy}) [I_{xy} \neq J_{xy}] \tag{3}$$

Here the value of $[I_{xy} \neq J_{xy}]$ is equal to 1 when $[I_{xy} \neq J_{xy}]$ is true and is equal to 0 when $[I_{xy} \neq J_{xy}]$ is false. For high frequency components, the embedding cost in SUNIWARD is higher as it avoids embedding in zeros.

3 Proposed Steganalyzer

3.1 Feature Extraction

Feature model selection is a crucial task for steganalysis as the classification accuracy depends on the intrinsic nature of the features that undergo changes during steganography. The choice of the image features should be in spatial domain as the steganographic scheme used is in spatial domain (SUNIWARD). As steganography changes only few pixels in cover images, it would suffice to model these changes, in the form of noise components. Also, framing a single model and enlarging it leads to many under populated bins [8]. This necessitates the need for an appropriate model built from many sub models. This research work uses the sub models of the noise components of pixels. As supported by literature, modeling the noise residual seems to be advantageous as the dynamic range of the pixel values reduce giving way for effective statistical evaluations [9–12]. The noise models proposed by Kodovsky et al. [11] form the basis for extracting the feature set in this research work. These are the spam features by Du [13] in horizontal and vertical directions. The residual noise components are calculated from high pass filters for specific neighborhood and specific residual order.

$$P_{ij} = \tilde{Q}_{ij}(\rho_{ij}) - rQ_{ij} \qquad (4)$$

where P_{ij} is the residual, ρ_{ij} is the neighborhood of Q_{ij}, $\tilde{Q}_{ij}(\cdot)$ is a predictor of rQ_{ij} on the neighborhood and $\{Q_{ij} + \rho_{ij}\}$ is the support of the residual. The chosen residuals are further truncated after quantization as in Eq. (5) for better dynamic range of the residuals,

$$P_{ij} = T(P_{ij}/s) \qquad (5)$$

where T is the truncation function and s is the quantization step size. Quantization enhances the embedding changes in the spatial discontinuities to be more obvious in the noise residuals. The feature model construction is based on the co-occurrence matrices of neighboring noise residuals which depends on the choice of the truncation constant T, spatial distribution of the noise residuals and the order of the co-occurrence matrices. Since the pixel correlation falls off at a faster rate in diagonal direction, only correlation between horizontal and vertical pixels are considered in this work. The maximum co-occurrence order is chosen as 4 and the value of T is kept small. The loss of information due to truncation with small T is compensated by considering several sub-models with varying s. For a co-occurrence matrix of dimension 4, indices being $d=(d_1, d_2, d_3, d_4)$ $\acute{T}_4$, $\acute{T}_4$ being $\{-\acute{T}, ..., \acute{T}\}^4$, the number of elements in the co-occurrence array is $(2\acute{T}+1)^4 = 625$. Considering the horizontal co-occurrence matrix, the dth element in the array is given as,

$$A_d^{(h)} = \mathbf{1}/(N)\big|\{(P_{ij}, P_{i,j+1}, P_{i,j+2}, P_{i,j+3})\big|P_{i,j+k-1} = d_k; k = 1, 2, 3, 4\} \qquad (6)$$

here N is the normalization constant assuring $A_d^{(h)}=1$. The analogy of defining the vertical co-occurrence is similar to that of horizontal co-occurrences. For a fixed value of T and fixed order of co-occurrence, finding the sub-models depends on the selection of the predictor $\tilde{Q}_{ij}(\cdot)$ and the quantization step size s. In order to capture the noise differences between nearby as well as distant pixels, the residual order is taken from 1 to 3 for a chosen neighbourhood. The direction of the neighbourhood is in horizontal and vertical directions for spam features. For a first order spam feature the noise residue is given by $N_{xy} = I_{x,y+1} - I_{xy}$ and for a second order spam feature the noise residue is given by $N_{xy} = I_{x,y-1} + I_{x,y+1} -$

$2I_{xy}$ where I_{xy} is the central pixel for which the residual is being calculated. The extracted residuals can be grouped into six categories—the immediate neighbours (1st), a linear quadratic model (2nd and 3rd), the spatial discontinuities as 3×3 edge (4th), circularly symmetrical kernel as 5×5 edge (5th) and square kernel (6th).

Better populated models are created by reducing the number of sub models based on the symmetries of the residuals. Some residuals are non directional i.e. the residual value does not change even when the image is rotated 90°. If the residual value changes, they are directional. All non directional residuals are symmetrical as their horizontal and vertical co-occurrence matrices can be added to form a single matrix. Non symmetrical (directional) residuals produce two co-occurrence matrices one for horizontal direction and one for vertical direction. Each residual is associated with a specific symmetry index (μ) representing the different residuals that can be obtained by rotational mirroring of the image before computing residuals. To reduce the dimensionality and increase the statistical robustness, the non symmetrical residuals of same symmetry index, μ are added. Spam features are also considered for square kernels [6]. These square and edge kernels provide better noise estimates in textures where spatial discontinuities are more. The spam features calculated thus are reduced (duplicates removed) by sequential symmetrisation [6] to have only the unique features. The extraction of the minmax residuals involves more than two linear filters. Each filter corresponds to horizontal or vertical directions and the final residual is based on the minimum or maximum of the filter outputs.

The number of co-occurrence matrices for each of the six categories are calculated as— 22 co-occurrence matrices for the 1st category, 12 co-occurrence matrices for the 2nd, 22 co-occurrences for the 3rd, 2 co-occurrences for the category square and 10 each for edges (5×5 and 3×3). Based on these computations 78 co-occurrence matrices each with 625 elements are used to build the required feature sub models. Since symmetrisation is the key for increasing the population of the co-occurrence bins, it requires care in implementation. Both sign symmetry and direction symmetry are used depending on the type of residue. The rules adopted in symmetrising the spam features are,

$$\bar{A}_d = A_d + A_{-d} \tag{7}$$

$$\bar{\bar{A}}d = \bar{A}_d + \bar{A}_{dn} \tag{8}$$

where $d = (d_1, d_2, d_3, d_4) \in \acute{\Gamma}_4$, $dn = (d_4, d_3, d_2, d_1)$ and $-d = (-d_1, -d_2, -d_3, -d_4)$. Symmetrisation based on this removes the duplicates and gives only 169 elements from 625 elements. The minmax residuals possess directional symmetry and satisfy the relation $min(X) = -max(X)$ for a finite set $X \subseteq R$. The rules for symmetrising minmax residuals are,

$$\bar{A}_d = A_d^{min} + A_{-d}^{max} \tag{9}$$

$$\bar{\bar{A}}d = \bar{A}_d + \bar{A}_{dn} \tag{10}$$

where A_d^{min} is the min co-occurrence and A_{-d}^{max} is the max co-occurrence for a specific symmetry index μ. This reduces the dimensionality from 625×2 to 325. Thus symmetrisation reduces the sub models from 78 to 45. These 45 sub models are due to 12 models from 1st category, 12 from 3rd category, 7 from 2nd category, 2 from square, 6 each from the edges (5×5 and 3×3). Finally, the sub models due to spam residuals are 12 each with 169 elements and the sub models due to minmax are 33 each with 325 elements. So the total dimension of the extracted feature set due to spam and minmax features is 34,671 as in Table 1. The extracted high dimensional feature set has well populated bins.

Table 1 The different categories of spatial features

Category	Number of sub models	SPAM features	MinMax features
1st	22	2×338	20×325
2nd	18	3×338	15×325
3rd	33	3×338	30×325
3×3	15	3×338	12×325
5×5	18	6×338	12×325
Total	106 models	5746	28,925
Total no. of features		34,671 features	

For better results, low value of truncation coefficient is considered and the relation between the residue order r and the quantization step s has to be $s \in [r, 2r]$. This large feature set has been derived for the best values of s and r.

3.2 Unsupervised Optimization

In the above extracted features, though under populated bins have been eliminated, the final dimensionality of the feature set is large. To avoid this curse of dimensionality in terms of computational time and classification accuracy, optimization of the feature set becomes essential [6]. Many statistical feature reduction techniques like Principle Component Analysis (PCA), Fisher Linear Discriminant (FLD) [14] and Projections of residuals [15] have been used in past. With the intuition of obtaining the best possible features for steganalysis a novel unsupervised optimization technique has been proposed to optimize the feature set for better classification. Unsupervised algorithms help in finding structures in data which describe the data in a compact manner [16]. These methods depend only on the input data to find the structures in data and hence are called as unsupervised.

In a high dimensional space, validation of the optimization can be examined in terms of distance based cluster plots [17]. When all data points are in same physical units, Euclidean distance can be used to cluster data. Clustering can be done with Vector Quantization based on codebook vectors with cluster centers [13]. A subset of codebook vectors is associated with each cluster. Grouping of image features (cluster points) are based on the Euclidean distance between each input feature vector and the center cluster according to the nearest neighbor function.

The algorithm works for 4 clusters and each feature model acts as the input space, the optimized output feature vector has dimension 848 (reduced from 34,671). The measure used for minimum distance clustering is the pair wise Euclidean distance measure. The clusters obtained are exclusive with no overlapping. This method adapts the cluster centers in the feature space and tries to minimize the mean distance between a data point and this location. The measuring metric for Euclidean distance

$$D_e^2 = (x_1 - x_2)(x_1 - x_2) \tag{11}$$

This optimization gives a feature reduction ratio of 93.35% while maintaining the useful and valid information in the data. For each image, a feature vector of dimension 848 is

obtained by concatenating the optimized features from each noise sub model. Distribution of image data in cluster space before and after optimization is shown in Figs. 1 and 2.

3.3 Classification Scheme

The optimized feature set is used to classify the image as a clear or stego image. Classification plays a crucial role in steganalysis. One of the most powerful classifier used by many steganalysts is the Support Vector Machines (SVM) [11, 18–21]. SVM maps the input vector onto a high dimensional feature space by non linear mapping [18]. Training in SVM involves determining the decision boundary and minimizing the aggregate distance (cost function) between the hyper plane and the support vectors. Configuring the SVM depends on the chosen kernel, kernel parameters and the margin parameter value. The cost function to be minimized is of the form,

$$F(x) = (\mu/2)\|w\|^2 + \mathbf{1}/\mathbf{L}\sum|1 - y_1 f_o(x_1)|_+ \tag{12}$$

The hyper parameter μ is the weight change parameter which is used to maximize the margin and minimize the margin errors. $|1 - y_l f_o(x_l)|_+ = max(0,x)$. The weight value is given by

$$w = (\mathbf{1}/\mu)\sum Y_l A_l \theta(x_l) \tag{13}$$

the parameter A_l is seldom non zero and hence give only few training samples. These samples are called as support vectors and give a sparse w matrix. Multi Layer Perceptron (MLP) is another classifier that gives good classification accuracy for images. MLP works
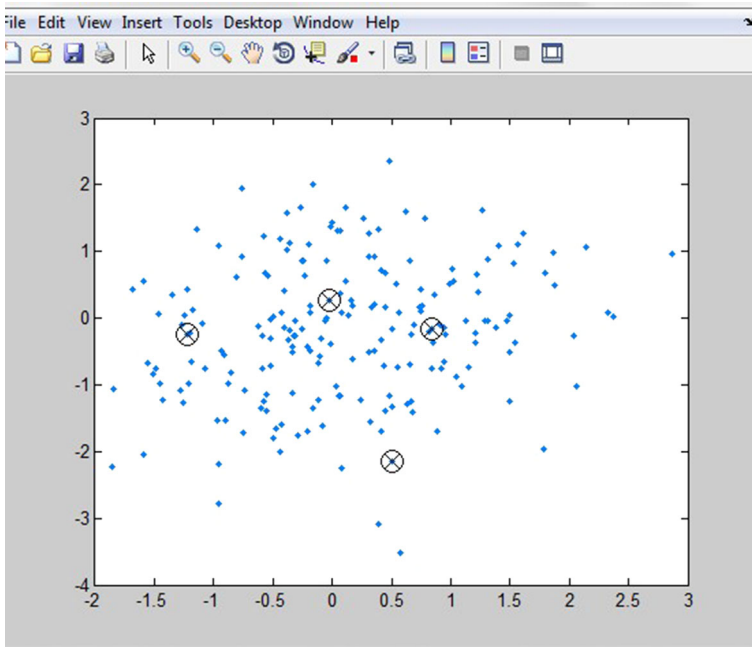

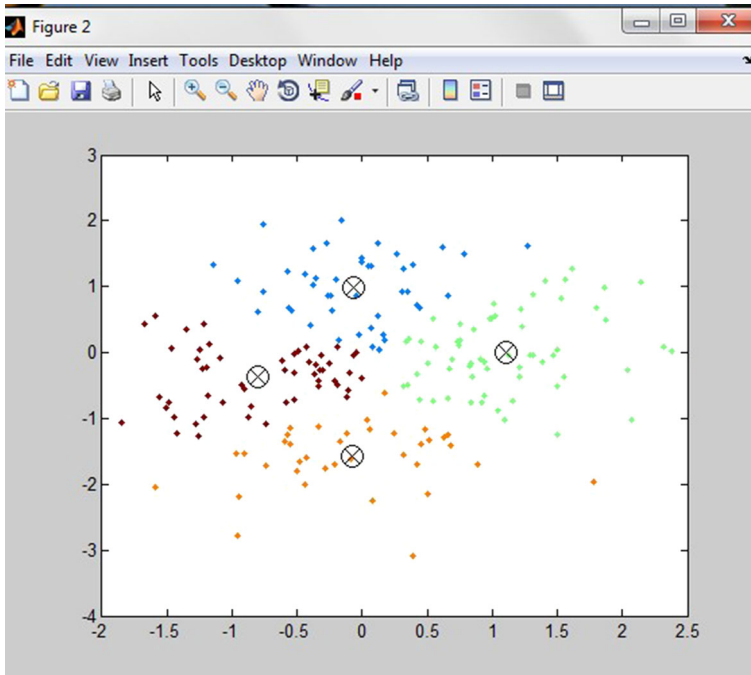**Fig. 1** Image data in search space before optimization

Fig. 2 Image data after Euclidean distance optimization

on the principle of back propagation [21] with high response. In MLP, the weights (fitting parameters) are adjusted so that the sum of square error between the target and the present output is minimized. MLP learns by processing the N-dimensional input vector x such that the error function is minimized to produce the M dimensional output vector. The error function

$$E(w) = \mathbf{1/2} \sum \|y(x_i, w) - d_i\|^2 \tag{14}$$

with $y(x_i, w)$ being the output, w the adaptive weight follows gradient algorithm for minimizing the error. The weight update equation for MLP is,

$$w(k+1) = w(k) + \rho q(k) \tag{15}$$

where ρ is the adaptation coefficient and $q(k)$ is the minimization direction in the kth iteration.

Both these classifiers have their own inherent advantages and differ based on the learning strategies [22]. However literature shows that fusion of classifiers gives better classification accuracy than the individual classifiers [23].

Among the different classifier fusion schemes, this research work uses the fusion of SVM and MLP classifiers with three different fusion schemes, namely Bayes method, Decision Template scheme and Dempster Schafer method. The decision template fusion scheme for a class j gives the average of the decision profiles of the elements in the training set Y in the class j. The decision template of the class Y is a matrix whose (m, n)th element is,

$$DT_{j(m,n)}(Y) = \sum F(Y_{k,j})DT_{m,n}(Y_k) / \sum F(Y_{k,j}), \quad m = 1, 2, \ldots P, \, n = 1, 2, \ldots Q \quad (16)$$

where $F(Y_{k,j})$ is the indicator function with value 1 if Y_k is a member of the class and 0 otherwise. When an input vector is presented, the decision template calculates a similarity measure, which acts as a support for the vector in that class. Similarity measure may be in terms of consistency measure or inclusion indices.

Bayesian inference works on the conditional probability. It is highly scalable and has maximum likelihood training. For classifier fusion, Bayes method takes the weights as classifier's posterior probability. Accordingly,

$$Class(x) = \arg\max \sum P(Y_k|T) \cdot P'Y_k(y = c_i|x) \quad (17)$$

here the probability of classifier Y_k is $P(Y_k|T)$.

The Dempster Schafer fusion scheme is similar to decision template, but the output is in terms of probable proximities rather than similarity measure between the decision profile and the decision template. The proximity between the decision template DTj (for ith row and jth class) and every classifier input $Di(x)$ represented as a matrix norm is,

$$D_{j,i}(x) = \frac{(1 + \|DTj - Di(x)\|)^2}{\sum (1 + \|DTk - Di(x)\|)^2} \quad (18)$$

These three methods (Bayes method, Decision Template scheme and Dempster Schafer method) are seen to be competitive with good classification accuracies compared to other fusion strategies [23].

4 Implementation and Results

4.1 Image Data Base

The image data base used for this research involves the BOSS (Break Our Steganographic System) database which has two sets of database named the BOSS base and BOSS Rank sets [24]. The BOSS Base has 1000 uncompressed raw images taken from different cameras like Panasonic, Canon, EOS, Pentax, Nikon and Leica. The raw format of these images is .cr2 or .dng and later converted to.png format. All these images are 512×512 gray scale images. For this work, 100 raw images from BOSS Base have been taken and processed with the spatial steganographic method (described in Sect. 2) to create 100 stego images. These stego images have the embedded information. The combined set of 200 images is used for feature extraction, optimization and then classification.

4.2 Experimental Frame Work

All algorithms are implemented with MATLAB. The embedding algorithm used to obtain the stego image is the SUNIWARD simulator

Algorithm 1: Optimized Feature Extraction

Input: 100 cover images and 100 stego images.

Output: 848 features from 200 images.

Start

For i = 1 to number of cover images

    Read cover image

    Extract features (34671)

    Apply optimization algorithm

        Select number of clusters

        Set minimum distance metric.

Run optimization and get output as 848 features.

Repeat for all cover images

For i = 1 to number of cover images

    Read cover image

    Create stego image for a specific payload

    Extract features (34671)

    Apply optimization algorithm

        Select number of clusters

        Set minimum distance metric.

Run optimization and get output as 848 features.

Repeat for all cover images

Concatenate cover and stego features to get 200×848 features.

End

Algorithm 2:   Classification

Input: 200×848 features

Output: Classification accuracies

Start

Set values of No. of classes

    Set values for partitions

        Run classification algorithm

            Do data partitioning

            Perform training of classifiers

            Perform testing of classifiers

            Find classification accuracies

Out the classifier accuracies

End

Table 2 Classifier accuracies of individual classifiers

No. of cluster	Feature size	Payload	SVM	MLP
K = 4	200 × 848	0.1	0.7232	0.9866
		0.3	0.7368	0.9933
		0.5	0.7201	0.9900
		0.7	0.7200	0.9883
Average classification accuracy			0.725025	0.98955

Table 3 Classifier accuracies of fusion classifiers

No. of cluster	Feature size	Payload	Bayes	Decision template	Dempster Schafer
K= 4	200 × 848	0.1	0.7232	0.9933	0.4998
		0.3	0.7368	0.9917	0.5002
		0.5	0.7201	0.9934	0.5000
		0.7	0.7200	0.9917	0.5000
Average classification accuracy			0.725025	0.992525	0.5

The classifications based on Euclidean distance metrics is enumerated in Tables 2 and 3. The classifier accuracy is the percentage of correct prediction of image (as either a stego image or a clean image). Classification accuracy is

$$Accuracy = (TP + TN)/(TP + TN + FP + FN) \qquad (19)$$

where TP is True Positive, TN is True Negative, FP is False Positive and FN is False Negative [25]. For the Euclidean distance optimization, the classification accuracies are good for a payload of 0.5. Maximum accuracies occur for MLP and the fusion scheme based on Decision Template. Comparing individual classifiers, MLP classifier gives better results compared to SVM as indicated in Fig. 3 and Decision Template classifier gives the best results among all compared classifiers as shown in Fig. 4.

Comparison of the best among individual and the best among fusion classifiers is shown in Fig. 5.

Thus Euclidean distance based optimization with Decision Template fusion classification scheme seems to be a good choice for better image steganalysis than the existing work. The comparison of the results with the existing state of art is shown in Table 4.
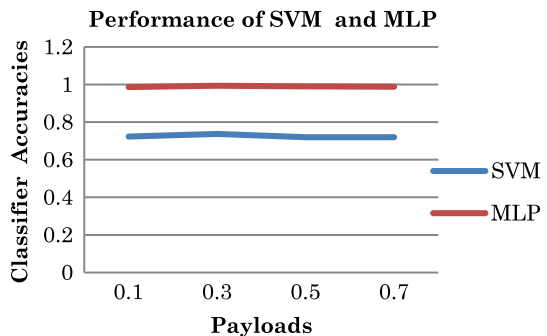
Fig. 3 Comparison of SVM and MLP

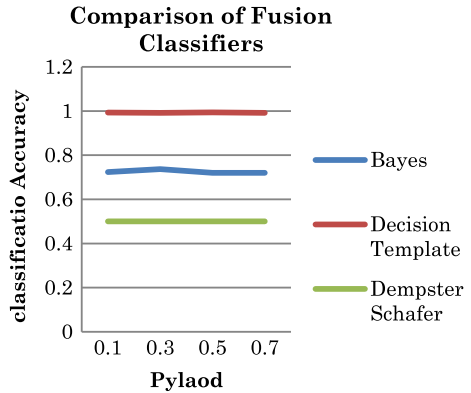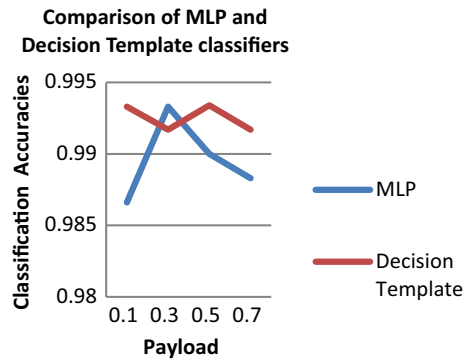**Fig. 4** Comparison of fusion classifiers

Comparison of Fusion Classifiers

Fig. 5 MLP versus decision template classifiers

Comparison of MLP and Decision Template classifiers

5 Discussion and Conclusion

As steganalysis without prior knowledge of the steganographic technique is the need, universal image steganalysis with a rich set of image features has been implemented and tested in this work. All possible features of an image in spatial domain are extracted. The changes in statistical dependencies among neighboring pixels have been modeled as noise residuals to create the feature space. The exact model has been created with four dimensional co-occurrence matrices of the quantised and truncated noise residues. A large dimension feature set is extracted to get information about the embedding and is further reduced by an unsupervised clustering based optimization technique to identify the most appropriate features for classification. The classifiers implemented include the individual classifiers SVM, MLP and three fusion classifiers (Bayes, Decision Template and Dempster Schafer fusion schemes). It has been identified that MLP performs better than SVM but is not superior to fusion classifiers. Comparing all the fusion classifiers, Decision template based fusion method gives the best classification accuracy (99.25%). Thus the implemented unsupervised optimization method for extracting best image features combined with Decision template fusion classification scheme provides the best classification of stego and clear images as compared to the existing work in this field. Future scope of this research is steganalysis with genetic based optimization techniques as they select more appropriate features for accurate classification.

Table 4 Comparison of the results with the existing state of art

| Steganographic domain	Researcher (s) name	Type of feature set	Performance measure and it's value	Method of classification	Value of error measure	Percentage accuracy
SPATIAL	Fridrich and Kodovsky [6]	34,671	Detection error	Ensemble with FLD	6%	94.00%
	Fridrich et al. [26]	33,963	Detection error	Ensemble with FLD	41.4%	58.6%
	Chonev and Ker [27]	25,000	Classification accuracy	FLD	–	85%
	This research	34,671	Average percentage classification accuracy	Unsupervised optimization and single classifiers (SVM, MLP)	–	98.95% For MLP
				Unsupervised optimization and FUSION classifiers (Bayes, Decision template, Dempster Schafer)	–	99.25% For Decision Template fusion classifier

Appendix

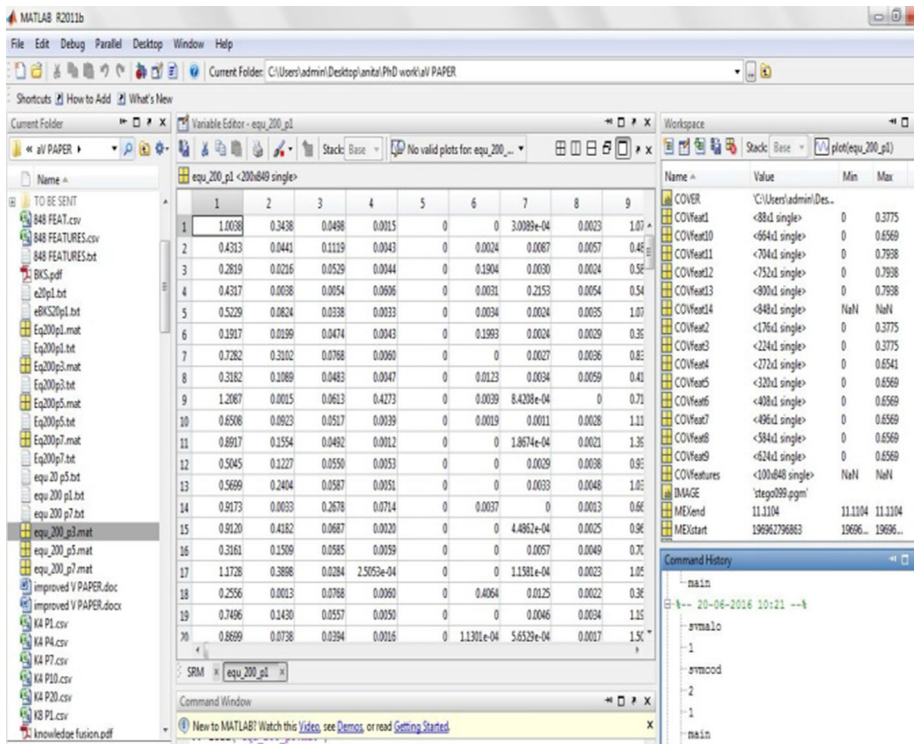See Figs. 6 and 7.
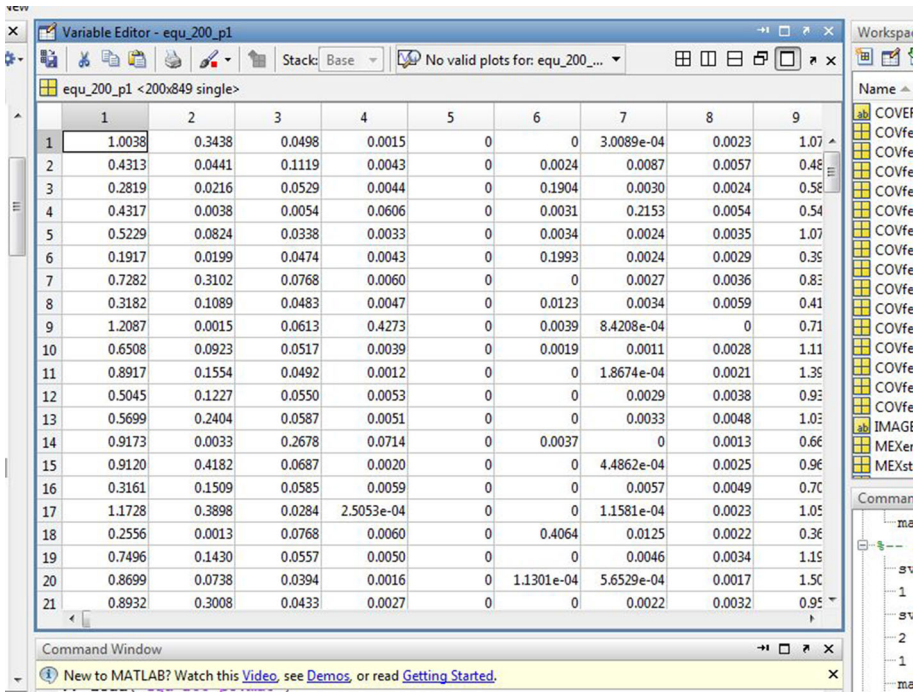
Fig. 6 Image features extracted

Fig. 7 Image features extracted in .mat file

References

1. Cachin, C. (2004). An information-theoretic model for steganography. *Information and Computation Elsevier, 192*(1), 41–56.
2. Pevny, T., Bas, P., & Fridrich, J. (2010). Steganalysis by subtractive pixel adjacency matrix. *IEEE Transactions on Information Forensics and Security, 5*(2), 215–224.
3. Pevny, T., Filler, T., & Bas, P. (2010). Using high-dimensional image models to perform Highly undetectable steganography. *International Workshop on Information Hiding, 6387,* 161–177.
4. Holub, V., Fridrich, J., & Denemark, T. (2014). Universal distortion function for steganography in an arbitrary domain. *EURASIP Journal on Information Security.* http://jis.eurasipjournals.com/content/2014/1/1. Accessed December 25, 2016.
5. Holub, V., & Fridrich, J. (2012). Designing steganographic distortion using directional filters. In *Fourth IEEE international workshop on information forensics and security* (pp. 234–239).
6. Fridrich, J., & Kodovsky, J. (2011). Rich models for steganalysis of digital images. *IEEE Transactions on Information Forensics and Security, 7*(30), 868–882.
7. Filler, T., Judas, J., & Fridrich, J. (2011). Minimizing additive distortion in steganography using syndrome-trellis codes. *IEEE Transactions on Information Forensics and Security, 6*(3), 920–935.
8. Goljan, M., Fridrich, J., Holotyak, Y. (2006). New blind steganalysis and its implications. In *Electronic imaging* (p. 607201-13).
9. Zou, D., Shi, Y. Q., & Xuan, W. S. G. (2006). Steganalysis based on Markov model of thresholded prediction-error image. In *Proceedings IEEE international conference on multimedia and expo* (pp. 1365–1368).
10. Ker, A. D., & Bohme, R. (2012). Revisiting weighted stego-image steganalysis. In *Electronic imaging* (p. 681905-17).

11. Kodovsky, J., Fridrich, J., & Holub, V. (2012). Ensemble classifiers for steganalysis of digital media. *IEEE Transactions on Information Forensics and Security, 7*(2), 432–444. https://doi.org/10.1109/tifs.2011.2175919.
12. Holub, V., Fridrich, J., & Denemark, T. (2013). Random projections of residuals as an alternative to co-occurrences in steganalysis. *IS&T/SPIE Electronic Imaging.* https://doi.org/10.1117/12.1000330.
13. Du, K. L. (2010). Clustering—A neural network approach. *Journal Neural Networks, 23*(1), 89–107.
14. Shi, Y. Q., Chen, C., & Chen, W. (2006). A Markov process based approach to effective attacking JPEG steganography. In: *Proceedings of the 8th international workshop information hiding* (pp. 249–264). Springer, Berlin.
15. Lyu, S., & Farid, H. (2004). Steganalysis using color wavelet statistics and one-class vector support machines. *Proceedings of SPIE Security Steganography Watermarking of Multimedia Contents, 5301,* 35–45.
16. Yu, H. F., Hsieh, C. J., Chang, K. W., & Lin, C. J. (2012). Large linear classification when data cannot fit in memory. *ACM Transactions on Knowledge Discovery From Data, 5*(4), 1–23. https://doi.org/10.1145/2086737.2086743.
17. Sri Lalitha, Y., & Govardhan, A. (2015). Improved Text Clustering with Neighbours. *International Journal of Data Mining & Knowledge Management Process (IJDKP), 5*(2), 23–37. https://doi.org/10.5121/ijdkp.2015.5203.
18. Sheikhan, M., Moin, M. S., & Pezhmanpour, M. (2010). Blind image steganalysis via joint co-occurrence matrix and statistical moments of contourlet transform. In *Intelligent systems design and applications* (ISDA) (pp. 368–372).
19. Avcibas, I., Kharrazib, M., Memon, N., & Sankurd, B. (2005). Image steganalysis with binary similarity measures. *EURASIP JASP, 17,* 2749–2757.
20. Pevny, T., & Fridrich, J. (2007). Merging Markov and DCT features for multi-class jpeg steganalysis. In *IS and T/SPIE EI* (Vol. 6505).
21. Windeatt, T. (2006). Accuracy/diversity and ensemble MLP classifier design. *IEEE Transactions on Neural Networks, 17*(5), 1194–1211.
22. Gader, P. D., Mohamed, M. A., & Keller, M. (1996). Fusion of handwritten word classifiers. *Pattern Recognition Letters—Special Issue on Fuzzy Set Technology in Pattern Recognition, 17*(6), 577–584. https://doi.org/10.1016/0167-8655(96)00021-9.
23. Kuncheva, L. I., Bezdek, J. C., & Duin, R. P. W. (2001). Decision templates for multiple classifier fusion: An experimental comparison. *Pattern Recognition, 34*(2), 299–314.
24. Bas, P., Filler, T., & Pevny, T. (2011). Break our steganographic system—The ins and outs of organizing BOSS. In *Proceedings of information hiding conference* (Vol. 6958, pp. 59–70).
25. Kantarcioglu, M., Xi, B., & Clifton, C. (2011). Classifier evaluation and attribute selection against active adversaries. *Data Mining and Knowledge Discovery—DATAMINE, 22*(1–2), 291–335. https://doi.org/10.1007/s10618-010-0197-3.
26. Fridrich, J., Kodovský, J., Goljan, M., & Holub, V. (2011). Steganalysis of content-adaptive steganography in spatial domain. In T. Filler, T. Pevný, A. Ker, & S. Craver (Eds.), *Information Hiding, 13th International Workshop, volume 6958 of Lecture Notes in Computer Science* (pp. 102–117). Prague, Czech Republic, May 18–20, 2011.
27. Chonev, V. K., & Ker, A. D. (2011). Feature restoration and distortion metrics. In *Multimedia Watermarking, Security and Forensics XIII, volume 7880 of Proc. SPIE, 2011* (pp. 0G01–0G14).

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Anita Christaline Johnvictor was born in Dindigul, India in 1974. She completed her B.E. Degree in Electronics and Communication Engineering in 1996 from Bharathidasan University, Trichy. She is recipient of GATE score during 2003. She received her M.E. Degree in Applied Electronics in the year 2006 from Anna University, Chennai. She is currently working as Assistant Professor in SRM University, Chennai where she is pursuing her Doctoral Degree. Her area of interest is Steganography and Steganalysis by computational intelligence. She has published many papers in this field. She is member IEEE and IET.

Ramesh Rangaswamy was born in Kanyakumari, India, 1976. He received the B.E. Degree in Electronics and Communication Engineering in 1998 and the M.E. degree in Communication Systems in 2000, both from Madurai Kamaraj University, India. He has been awarded Doctoral degree in the SRM University in the year 2009 for his research work on Testing the Stability of two dimensional recursive filters. He is currently working as a Professor in the Department of Electronics and Communication Engineering at Saveetha Engineering College, Chennai. India. His current research interests concern digital signal processing particularly to find the stability of two dimensional recursive digital filters.

Gomathy Chidambaram acquired a B.E. (with honors) degree in Electronics and Communication Engineering from Government College of Engineering, Tirunelveli, in the year 1986, and an M.S. degree in Electronics and Control Engineering from Birla Institute of Technology and Science, Pilani, in 1992. She also obtained an M.S. (by research) degree from Anna University in 2001. She obtained her Ph. D. in the area of Mobile Ad hoc networks from Department of Electronics and Communication Engineering, College of Engineering, Anna University, Chennai, India, in the year 2007. She has published over 70 research papers in National and International conferences and journals. Her areas of interest include mobile ad hoc networks, high-speed networks, and wireless sensor networks.

Vaishali Durgamahanthi was born in Pune, India, 1969. She received B.E. Degree in Electronics and Telecommunication Engineering in 1994 and the M.E. degree in Communication Systems in 2002, both from Pune University, India. She is a research scholar in the SRM University and her research work is progressing in the field of Image Processing with wavelet Transforms. She is currently working as a Asst. Professor in the Department of Electronics and Communication Engineering at SRM University, Chennai. India. She is member IEEE and IET.