INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
ISSN 1841-9836, 10(5):678-685, October, 2015.

PCA Encrypted Short Acoustic Data Inculcated in Digital Color Images

S.H. Karamchandani, K.J. Gandhi, S.R. Gosalia,
V.K. Madan, S.N. Merchant, U.B. Desai

Sunil H. Karamchandani*
1. Indian Institute of Technology, Bombay
Mumbai, India.
*Corresponding author: skaramchandani@rediffmail.com

Krutarth J. Gandhi, Siddharth R. Gosalia
D. J. Sanghvi College of Engineering
Mumbai, India.
gandhi.krutarth@yahoo.co.in
siddharth_destiny@hotmail.com

Vinod K. Madan
Kalasalingam University
Krishnankoil (TN), India
klvkmadan@gmail.com

Shabbir N. Merchant
Indian Institute of Technology, Bombay
Mumbai, India
merchant@iitb.ac.in

Uday B. Desai
Indian Institute of Technology, Hyderabad
Hyderabad, India
ubdesai@iith.ac.in

Abstract: We propose develop a generalized algorithm for hiding audio signal using image steganography. The authors suggest transmitting short audio messages camouflaged in digital images using Principal Component Analysis (PCA) as an encryption technique. The quantum of principal components required to represent the audio signal by removing the redundancies is a measure of the magnitude of the Eigen values. The aforementioned technique follows a dual task of encryption and in turn also compresses the audio data, sufficient enough to be buried in the image. A 57Kb audio signal is decipher from the Stego image with a high PSNR of 47.49 and a correspondingly low mse of 3.3266×10^{-6} with an equalized high quality audio output. The consistent and comparable experimental results on application of the proposed method across a series of images demonstrate that PCA based encryption can be adapted as an universal rule for a specific payload and the desired compression ratio.
Keywords: Colour image steganography, principal component analysis, eigen thresholding, Pareto analysis.

1 Introduction

Steganography or Stego in IT parlance, in Greek means "covered writing" and is defined by Markus Kahn [1,2] as the art and science of communicating in a way which hides the existence of the communication. In contrast to Cryptography, where the enemy is allowed to detect, intercept and modify messages without being able to violate certain security premises guaranteed by a

cryptosystem, the goal of Steganography is to hide messages inside other harmless messages in a way that does not allow any enemy to even detect that there is a second message present. A modern steganographic system should defeat detection even by a machine. It replaces bits of useless or unused data in computer files such as graphics, sound, text, HTML with bits of different invisible information. This hidden information can be plain text, cipher text, sound, and even images. Ideally steganography can be used for any communication channel. However in real life the cover media generally used are multimedia objects such as image, video, and audio files. The reasons include that cover media should be large compared to the size of the secret message, and the methods so far developed have less than one percent of the cover size, and the indeterminacy in the cover is necessary to achieve the necessary security. Large objects without inderminacy like the value π at a very high precision are not suitable. The transmitting data should be plausible as in the modern digital society dependence on audio and image files are so prevalent. It is desirable that a steganography system should have a good embedding capacity, be secure and robust.

Modern steganography uses a number of techniques such as masking and filtering, algorithms and transformations, and least significant bit insertion [3]. In masking and filtering, the information is hidden inside of an image using digital watermarks that include information such as copyright, ownership, or licenses. It adds an attribute to the cover the image thus extends the amount of information presented. In algorithms and transformations technique data is hidden in mathematical functions that are often used in compression algorithms and facilitates to hide the secret message in the data bits in the least significant coefficients. The least significant bit insertion is the most common and popular method of the modern day steganography, and it uses LSB of a picture's pixel information keeping the overall image distortion to a minimum while the message is spaced out over the pixels in the images. This technique works best when the image file is larger than the message file. We propose to inculcate audio information in color images following the encryption of the short audio messages using the Principal Component Analysis (PCA). The age old adage of PCA incepts as an encrypting tool diversifying from its myth as a compression standard.

The framework of the paper is as follows. Literature review of steganography related to audio embedding is detailed in Section 2. The subsequent Section 3 discusses the proposed algorithm elaborating the PCA encryption of audio data followed by the implementation of steganography and steganalysis modules in Section 4. Simulation results are tabled in Section 5 with the conclusions drawn in Section 6.

2 Related Steganography Techniques for Audio Concealment

Majority of the steganography techniques disguise either image or text within an audio or image file. The major contributions concerned with encryption of acoustic files, particularly short audio message in images are described herewith. The authors of [4] discuss a traditional steganography technique which involves substitution of the least significant bit of each pixel of the cover image with the encrypted bits of the audio file. An arbitrary color bit stream of the RGB image is considered as the envelope where the kilo byte audio information is saved. The ciphering of the audio signals is loosely based on the very basic fundamentals of logic design. The encryption is performed [4] using the Boolean operations on the bit 5 and bit 6 of color information of the individual pixels given by (1)

$$h = (b_5 \oplus b_6) \oplus m \tag{1}$$

where "h" represents the encoded bit generated from the X−OR operation of the audio bit "m" with the bits of color information obtained from the image. The encoded bit then replaces the

LSB of the respective color pixels in the cover image. Since the exclusive disjunction of the three logical variables is associative and reversible the audio bit is recovered from (2)

$$m = (b_5 \oplus b_6) \oplus h \tag{2}$$

The process of steganalysis involves decoding of the secret "key" deposited in the header files. The knowledge of the header file exposes the entire stego process making it prune to brute force attacks. The algorithm also puts a cap on the length of the audio file to be embedded in the image. Moreover, the technique proposed in this paper does not give any information regarding the evaluation of the performance parameters. Exploiting the color information The authors [5–8] propose a steganography technique in which Discrete Cosine Transform (DCT) is performed on the RGB planes of the cover image by converting them to grayscale and creating blocks of size 8×8. The least significant bits of the frequency coefficients of B plane are swapped with the bits of the audio file [9]. Manipulation of the color component in the individual blocks leads to a computationally extensive algorithm. In the process the low frequency coefficients are also being manipulated which when converted back to spatial domain, the changes can be easily perceived by the human eye.

The PCA technique incorporated in our proposed algorithm is a method of compression which is camouflaged as an encryption technique for the audio signal. The proposed technique performed in the spatial domain and suggests a robust encryption algorithm, thus overcoming the drawbacks of the existing methods.

3 Proposed Steganography and Steganalysis Model for Short Audio Data

We propose to camouflage short audio signals encrypted by their immanent principal components within the digital color images.The wave or "wav", the most commonly used audio file format accommodates the audio information exclusively in the later 43 bytes. The header files, followed by four bytes representing the length of the file precede the audio data. The proposed model is sub divided in two modules: steganography and steganalysis for encryption and deciphering the data as illustrated in Figure 1

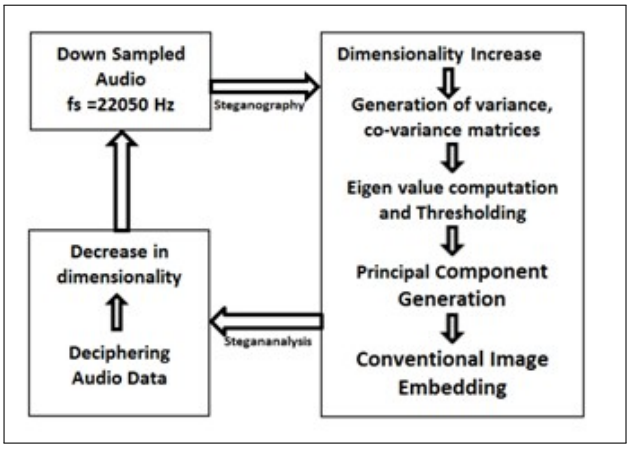

Figure 1: Proposed Model for smuggling audio information in image

3.1 Steganography: Manipulation of the embedded audio

The encryption is based on the Gaussian assumption that an N- dimension data evolves N directions of variance. The one dimension audio signal is trapped in a 2-D array and is represented as (3)

$$Z = WX \tag{3}$$

Where "W" indicates the weight vector required to minimize correlation between the two dimensions. This sets up the basis for the PCA technique which not only provides a good representation of the signal but also reduces the redundancy in the data. The feature similarity is obtained by decorrelating the covariance matrix as (4)

$$z_1^t z_2 = 0 \tag{4}$$

Diagonalization of the covariance matrix suppresses cross-dimensional co-activity. The decorrelation (whitening) of the weights is obtained as a diagonalization problem given in (5)

$$WCOV(X)W^t = I \tag{5}$$

where COV (x) denotes the covariance of the input data and I represents the identity matrix. The solution of the resulting covariance matrix is a function of the eigenvectors and eigenvalues of covariance of X. The decorrelating matrix W contains vectors that normalize the input's variance also called as the principal components while the audio signal gets scaled to a well developed Gaussian curve with unit variance in all dimensions. The decorrelating matrix is used to find the directions of maximal variance in the audio data. The variance of each principal component is reflected in the Eigen values obtained from the as a solution of diagonalization algorithm.The audiosigna lX_e,is obtained in (6) as a scrambled waveform with its redundancies eliminated.

$$X_e = XW \tag{6}$$

The scrambled lX_e is then conventionally embedded in the image using the missionary LSB algorithm.

3.2 Steganalysis: validation of the recovered audio

Steganalysis is art of reconnoitering the hidden information in a medium. The retrieval of speech signal from its redundant components is an essential criterion to discuss the robustness of the system. The generated principal components are used for deciphering the audio data using (7)

$$X = X_e W \tag{7}$$

The audio samples are then retrieved by reducing the dimensionality of the resultant $2-$D array X using the Pareto analysis.

4 Implementation of Proposed Scheme

The dominant frequency band where human speech cognition is most sensitive is around 3 kHz. The epiglottis transmits the vowel sounds at about 100 Hz. The audio message of size 57 Kb is disguised as a "wav" file in a 24-bit ".bmp" format color image of size $256 \times 256 \times 3$. Each sample is recorded in a 16 bit stereo.The audio signal sampled at 44100 samples/sec is down sampled to 22050 Hz, causing the removal of samples with frequencies above 10.025 Khz. The preprocessing eliminates whatever random noise existing in the wav file.

4.1 Proposed algorithm

The algorithm for hiding N audio samples by is structured by selecting the prime principal components. The Eigen vector corresponding to the highest Eigen value is the first principal component. By thresholding the Eigen value, the least important Eigen vectors can be eliminated and a matrix of retained Eigen vectors in the order of significance is formed whose transpose is multiplied with the transpose of audio sample matrix to obtain final matrix. The least significant bit of each pixel of the cover image is replaced by the bits of binary value of each element of the final matrix. This technique reduces the number of bits to be hidden as compared to the actual number of bits to be hidden.

5 Simulation Results

Short hidden message is embodied as an audio signal. The PCA performed on the final matrix contains the original data expressed in terms of retained Eigen vectors that are orthogonal to each other, thus compressing the original data.

5.1 Eigen thresholding

The Pareto analysis is used to determine the total number of Eigen values corresponding to the principal Eigen vectors to be considered for audio retrieval. This process we terms as Eigen thresholding. A Pareto chart in Figure 2 (a) plots the individual variances against the percentage contribution of each Eigen value. The first ten principal components contribute to about 95% information of the audio data. Figure 2 (b) shows the range of the Eigen values obtained from the covariance matrix. From the plot, it is indicated that the final 20% of the Eigen values represent the actual content of information in the signal. A hard threshold and the number of principal components are determined by analyzing the plot of the Eigen values. The audio data is therefore encrypted by compressing it into the fewer principal components. By

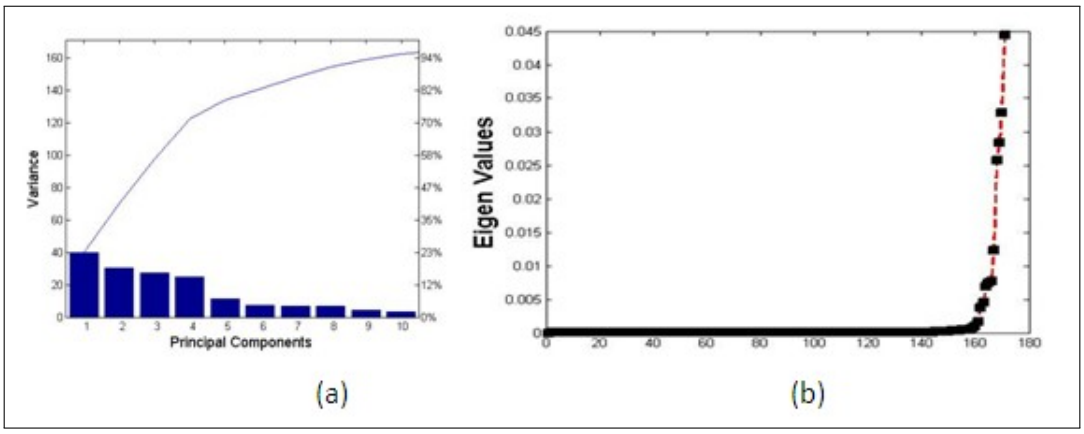

Figure 2: Proposed Model for smuggling audio information in image

thresholding the Eigen value, the principal components with least contribution have been ignored due to redundancy in theinformation. Table 1 illustrates the quantum of principal components and compression ratio obtained for different Eigen thresholds.

The compression ratio between the data to be hidden and original audio signal are calculated as (8)

$$CompressionRatio = \frac{OriginalData - CompressedData}{OriginalData} \tag{8}$$

Table 1: Number of principal components and Compression Ratios achieved for various thresholds of Eigen value for cover image (Barbara)

*E*igen Threshold	*P*rincipal Components Array	*C*ompression Ratio
10^{-2}	171×5	0.9707
10^{-3}	171×11	0.9355
0.8×10^{-3}	171×12	0.9297
10^{-4}	171×31	0.8183
0.75×10^{-4}	171×35	0.7948

The number of principal components exhibits an inverse relation with the Eigen threshold as observed in Table 1, causing the compression ratio to decrease. For threshold value 10−−4, sufficient number of principal components is retained to obtain the retrieved audio message with no distortion. Decrease in the number of principal components cause significant Eigen vectors to be ignored, resulting in misinterpretation of the derived audio. For threshold value greater than 10−−4, the additional principal components do not provide any information of the message and are of lesser significance. Also, the compression ratio will be decreased due to the additional data of lesser significance to be hidden in cover image. Hence the optimum value for thresholding the Eigen values is 10−−4 with ideal compression ratio of 0.8183.

5.2 Performance parameters for steganalysis

Performance parameters such as Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR), Normalized Absolute Error (NAE) [10], Maximum Difference (MD) and Structural Content (SC) are used to determine the quality of stego image and are calculated as (9)− (13). Table 2 charts the performance parameters of stego image for various thresholds of Eigen value.

$$MSE = \frac{1}{MN}\sum_{j=1}^{N}\sum_{k=1}^{N}(x_{j,k} - x'_{j,k})^2 \tag{9}$$

$$PSNR = 10\log\frac{255^2}{MSE} \tag{10}$$

$$SC = \frac{\sum_{j=1}^{M}\sum_{k=1}^{N}(x_{j,k})^2}{\sum_{j=1}^{M}\sum_{k=1}^{N}(x'_{j,k})^2} \tag{11}$$

$$MD = \max(\left|(x_{j,k} - x'_{j,k}\right|) \tag{12}$$

$$NAE = \frac{\sum_{j=1}^{M}\sum_{k=1}^{N}\left|(x_{j,k} - x'_{j,k})\right|}{\sum_{j=1}^{M}\sum_{k=1}^{N}\left|(x'_{j,k})\right|} \tag{13}$$

The cover image x of size M × N and the stego image of size x′ of size M × N, and x_jk and are pixel located at j−th row and k−th column of images x and x′ respectively. The maximum difference (MD) is the maximum value of absolute difference between a pixel located at j−th row and k−th column of cover image and Stego image and it is same, equal to 0.0039 for all thresholds of Eigen value. The value of structural content also remains same, equal to unity which implies the cover image and stego image are similar. As the threshold for Eigen values is decreased to retain the significant principal components, the size of final matrix increases thus increasing the number of bits to be replaced in cover image. Hence, the value of MSE and NAE increase causing a decrease in the values of PSNR. The above analysis and performance parameters

Table 2: Performance Evaluation parameters for different thresholds of Eigen value for multiple images

*E*igen *T*hreshold	*I*mage	*M*ean Square *E*rror (MSE)	*P*eak Signal to *N*oise Ratio (PSNR)	*N*ormalized Absolute *E*rror (NAE)
10^{-2}	*Lena*	5.4097×10^{-4}	48.13	2.7513×10^{-4}
	Barbara	5.4645×10^{-4}	48.13	3.3014×10^{-4}
	Cameraman	5.4535×10^{-4}	48.13	2.9832×10^{-4}
10^{-3}	*Lena*	1.1895×10^{-6}	47.86	6.0493×10^{-4}
	Barbara	1.1864×10^{-6}	47.86	7.1679×10^{-4}
	Cameraman	1.1882×10^{-6}	47.86	6.5001×10^{-4}
0.8×10^{-3}	*Lena*	1.2975×10^{-6}	47.83	6.5986×10^{-4}
	Barbara	1.2955×10^{-6}	47.83	7.8267×10^{-4}
	Cameraman	1.27555×10^{-6}	47.83	6.9776×10^{-4}
10^{-4}	*Lena*	3.3462×10^{-6}	47.49	1.7×10^{-3}
	Barbara	3.3404×10^{-6}	47.49	2×10^{-3}
	Cameraman	3.3188×10^{-6}	47.49	1.8×10^{-3}
0.75×10^{-4}	*Lena*	3.7630×10^{-6}	47.45	1.9×10^{-3}
	Barbara	3.7475×10^{-6}	47.45	2.3×10^{-3}
	Cameraman	3.7528×10^{-6}	47.45	2.1×10^{-3}

indicate successful implementation of hiding audio message in an image. The threshold for the Eigen values can be easily visualized by the Pareto chart which provides percentage contribution of the individual Eigen values. Figure 3 shows the stego image adjacent with the recovered audio signal.

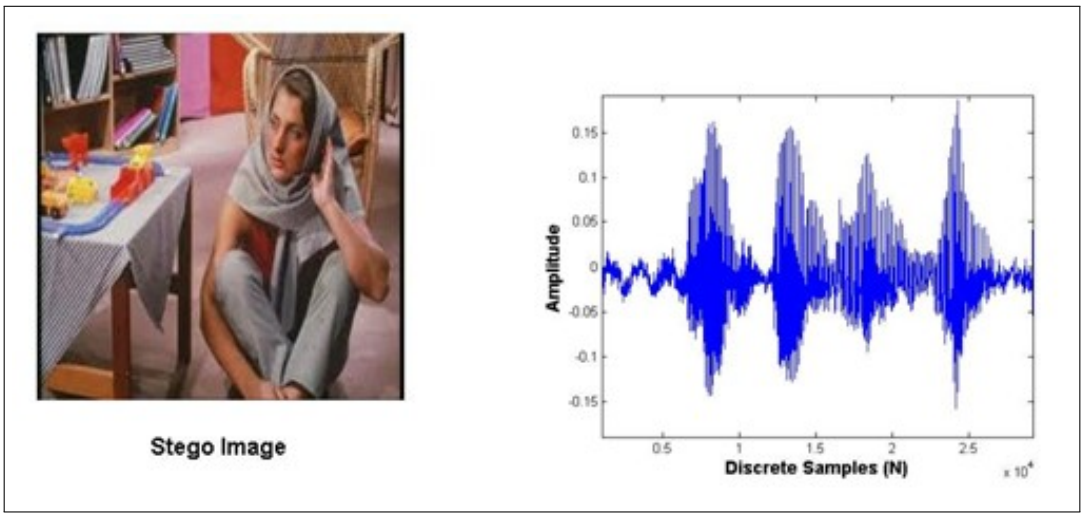

Figure 3: Stego image (Barbara) for Eigen threshold 0.0001 alongside retrieved audio sample

6 Conclusions and Future Works

We have developed a generalized procedure for concealing diminutive information using image steganography. The quantum of the principal components selected is a measure of the magnitude of the Eigen values. A high PSNR of 47.49 with reverse MSE value of 3.3266×10^{-6} is obtained

for the stego image along with a well audible recovered audio signal using the encrypted PCA technique. The results demonstrated in Table 2 illustrates comparable values of the parameters PSNR, NAE and MSE across a set of images concluding that the proposed method can be custom-made universally across all sets of images. For steganography, principal components emerge as a method of encryption while simultaneously performing compression of the audio signal. The future scope would involve adaptive thresholding of the Eigen value for principal component selection with the payload as a constraint.

Bibliography

[1] http://www.jjtc.com/stegdoc/index2.html

[2] Bohme, R. (2010); *Advanced Statistical Steganalysis*, Springer-Verlag, Berlin.

[3] Samarth, K.N.; Poornapragna, M.S. (2013); A Novel Technique of hiding an audio message in an image, *International Conference on Electronics and Communication Engineering*, ISBN: 978-93-83060-04-7, 170-172

[4] Khalil, M. I. (2011); Image Steganography: Hiding Short Audio Messages within Digital Images, *J. Computer Science and Technology*, ISSN: 1860-4749, 11(2): 68-73.

[5] Abd, E. H.; Abdulwahed, H. J.; Mohammed, H. A. (2012); The Use of Discrete Cosine Transformation (DCT), *Information Hiding Process*, J Kerbala University, 10(1):45-51.

[6] Shahana, T. (2013); A Secure DCT Image Steganography based on Public-Key Cryptography, *International Journal of Computer Trends and Technology (IJCTT)*, 4(7): 2039-2043.

[7] Bucerzan, D.; Ratiu, C.; Manolescu, M. J. (2013); SmartSteg: A New Android Based Steganography Application, *International Journal of Computers Communications & Control*, 8(5):681-688.

[8] Singh, K.M.; Chanu, Y.J.; Tuithung, T; (2014), Steganalysis of $\pm k$ Steganography based on Noncausal Linear Predictor, *International Journal of Comuters Communications & Control*, 9(5):623-632.

[9] Othman, S. E. (2012); Hide and Seek: Embedding Audio into RGB 24-bit Color Image Sporadically Using Linked List Concepts, *IOSR Journal of Computer Engineering (IOSRJCE)*, ISSN 2278-0661, 4(1): 37-44.

[10] WAVE and AVI Codec Registries - RFC 2361, Microsoft Corporation (June 1998).