Speech Scrambling Based on Principal Component Analysis

Dr. Nidaa A. Abbas, *Member, IEEE*
drnidaa_muhsin@ieee.org

Abstract

Scrambling methods are considered as important methods that provide the communication systems a specified degree of security, depending on the used technique to implement the scrambling method. There are many traditional scrambling methods used in single dimension such as time or frequency domain scrambling. This paper proposes the application of analog speech scrambler (also called analog speech encryption) using statistical method called Principal Component Analysis (PCA). Practically the descrambling process is implemented using traditional PCA algorithm and taking into consideration the test cases of many input speech signals. The objective test using Linear Predictive Coding (LPC) and Signal-to-Noise Ratio (SNR) where applied to evaluate the proposed system. It is shown via simulations that the proposed technique is more robust in the case of 8KH frequency samples.

Index Term- ***Analog speech, LPC, PCA, Scramblers, SNR.***

I. Introduction

As speech communications become more and more widely used and even more vulnerable, the importance of providing a high level of security is dramatically increasing. As such, a variety of speech encryption techniques have been introduced. The analogue encryption has been one of the most popular encryption techniques widely used in speech communication. In general, there are four main categories: frequency-domain scrambling (e.g., the frequency inverter and the band splitter), time-domain scrambling (e.g., the time element scrambling), amplitude scrambling (also known as the masking technique that covers the speech signal by the linear addition of pseudorandom amplitudes), and two-dimensional scrambling that combines the frequency-domain scrambling with the time-domain scrambling [1, 2]. Besides, there are many other analogue speech encryption methods in the transform domain, e.g., fast Fourier transform, discrete cosine transform and wavelet transform, etc. [3]–[5].

Recently, some new speech encryption methods including chaotic cryptosystem [6] and encryption using circulant transformations [7] have also been developed.

Modern analog speech scrambling systems have successfully tackled the issues of voice security by using new technological methods that have been borrowed from high fidelity digital audio bit rate reduction systems, namely MPEG audio [8].

Any analog speech scrambling algorithm needs to satisfy the following requirements:-

1. The scrambled speech should be unintelligible.
2. The scrambled speech should occupy the same bandwidth as does the original speech signal. That is, the scrambling process should be a bandwidth preserving operation.
3. It should be difficult to decrypt, if the decryption key is not available. In other words, it should be cryptanalytically strong or secure.
4. The communication delay caused by the scrambling process must be as small as possible.
5. The recovered speech at the receiver end should be of good quality and should preserve both the intelligibility of the speech and the characteristics of the speaker. [9]

Principal component analysis is a useful statistical technique that has found application in fields such as face recognition and image compression, and is a common technique for finding patterns in data of high dimension.

PCA it is useful when you have obtained data on a number of variables (possibly a large number of variables), and believe that there is some redundancy in those variables. In this case, redundancy means that some of the variables are correlated with one another, possibly because they are measuring the same construct. Because of this redundancy, you believe that it should be possible to reduce the observed variables into a smaller number of principal components (artificial variables) that will account for most of the variance in the observed variables [10].

This paper proposed the possibility of application of analog speech scrambler using statistical method techniques which is called Principal component analysis. Practically the descrambling process is implemented using traditional PCA algorithm. This paper contains four sections. Section two introduces the PCA technique, while section three describes the proposed system. Section four provides the conclusions.

Manuscript received September 13, 2009.
N. A. Abbas , with University of Babylon, Iraq. She is now with the Department of Computer Science, (phone: +9647801805510, e-mail: drnidaa_muhsin@ieee.org).

II. Principal Component Analysis (PCA)

Principal component analysis (PCA) is a mainstay of modern data analysis - a black box that is widely used. PCA has been called one of the most valuable results from applied linear algebra and is used abundantly in all forms of analysis - from neuroscience to computer graphics - because it is a simple, non-parametric method of extracting relevant information from confusing data sets. With minimal additional effort PCA provides a roadmap for how to reduce a complex data set to a lower dimension to reveal the sometimes hidden, simplified structure that often underlie it [11].

In their simplest forms, they assume that observed data can be represented by a linear combination of some unknown hidden factors called sources. The model is mathematically described as [12]:

$$X(t) = AS(t) \quad (1)$$

where $s(t) = [s_1(t),...,s_m(t)]^T$ is a $m \times 1$ column vector collecting the source signals, similarly vector $x(t)$ collects the n observed signals, A is an $n \times m$ matrix of unknown mixing coefficients, $n \geq m$, and t is the time index. This linear model represents a useful description for many applications.

PCA exploits the assumption that hidden source signals are not mutually related. PCA reconstructs the source signals **s** by decorrelating observed signals **x**.

PCA is a linear transformation that can be used to reduce, compress or simplify a data set. It does this by transforming the data to a coordinate system so that the greatest variance of the data by a projection of the data ends up on the first component (coordinate), the next one in line on the magnitude of variance ends up on the second component and so on. This way one can choose not to use all the components and still capture the most important part of the data.

Under the zero mean assumption, the Uncorrelatedness assumption is formally expressed as:

$$R_S = E[SS^T] = \Lambda \quad (2)$$

Where $\mathbf{R_S}$ represents covariance matrix, $\mathbf{\Lambda}$ is some diagonal matrix and E denotes mathematical expectation. The PCA transform **W** is designed such that the transformed data matrix:

$$Z = WX \quad (3)$$

has uncorrelated components, i.e., $\mathbf{R_Z} = \mathbf{\Lambda}$. Having this in mind, we derive the PCA transform from:

$$R_Z = E[ZZ^T] = WE[XX^T]W^T = WR_xW^T = \Lambda \quad (4)$$

From equation (4), we can recognize that the PCA transform **W** is nothing else but matrix of eigenvectors **E** obtained through eigen-decomposition of the data covariance matrix $\mathbf{R_X}$, i.e,

$$W = E^T \quad (5)$$

A special form of the PCA transform is whitening or sphering transform that makes transformed signals uncorrelated with unit variance. This is formally expressed as $\mathbf{R_Z} = \mathbf{I}$, where **I** represents identity matrix. This is equivalent to writing equation (4) as:

$$R_Z = E[ZZ^T] = WE[XX^T]W^T = \Lambda^{-1/2}WR_XW^T\Lambda^{-1/2} = I \quad (6)$$

Using equation (6), the whitening transform is obtained as:

$$w = \Lambda^{-1/2}E^T \quad (7)$$

A. *Properties and Limitations of PCA*

PCA is theoretically the optimal linear scheme, in terms of least mean square error, for compressing a set of high dimensional vectors into a set of lower dimensional vectors and then reconstructing the original set. It is a non-parametric analysis and the answer is unique and independent of any hypothesis about data probability distribution. However, the latter two properties are regarded as weakness as well as strength, in that being non-parametric, no prior knowledge can be incorporated and that PCA compressions often incur loss of information.

The applicability of PCA is limited by the assumptions made in its derivation. These assumptions are [13]:

1. Assumption on Linearity.

We assumed the observed data set to be linear combinations of certain basis. Non-linear methods such as kernel PCA have been developed without assuming linearity.

2. Assumption on the statistical importance of mean and covariance.

PCA uses the eigenvectors of the covariance matrix and it only finds the independent axes of the data under the Gaussian assumption. For non-Gaussian or multi-modal Gaussian data, PCA simply de-correlates the axes. When PCA is used for clustering, its main limitation is that it does not account for class separability since it makes no use of the class label of the feature vector. There is no guarantee that the directions of maximum variance will contain good features for discrimination.

3. Assumption that large variances have important dynamics

PCA simply performs a coordinate rotation that aligns the transformed axes with the directions of maximum variance. It is only when we believe that the observed data has a high signal-to-noise ratio that the principal components with larger variance correspond to interesting dynamics and lower ones correspond to noise.

Essentially, PCA involves only rotation and scaling. The above assumptions are made in order to simplify the algebraic computation on the data set.

B. Noise Measurement

Measurement of noise in any data set must be low or else, no matter the analysis technique, no information about a system can be extracted. There exists no absolute scale for noise but rather all noise is measured relative to the measurement. A common measure is the Signal-to-Noise Ratio (SNR) [14],

$$SNR = 10\log_{10}\frac{\sum_{n=-\infty}^{\infty} s^2(n)}{\sum_{n=-\infty}^{\infty}(s(n)-\hat{s}(n))^2}(dB) \qquad (8)$$

where n is the number of samples, s(n)is the amplitude of input speech signal and $s\hat{}(n)$ is the amplitude of reconstructed speech signal.

A high $SNR\ (\gg 1)$ indicates high precision data, while a low SNR indicates noise contaminated data.

C. LPC Distance

Is a tool used mostly in audio signal processing and speech processing for representing the spectral envelope of a digital signal of speech in compressed form, using the information of a linear predictive model. It is one of the most powerful speech analysis techniques, and one of the most useful methods for encoding good quality speech at a low bit rate and provides extremely accurate estimates of speech parameters [15].

$$d_{lpc}(c,e) = \ln\left(\frac{a_e R_c a_e^T}{a_c R_c a_c^T}\right) \qquad (9)$$

where Rc is the autocorrelation matrix of the clear speech block, vector a_c contains the LPC coefficients for the clear speech block and vector a_e contains the LPC coefficients for the scrambled speech block.

III. Proposed System

The steps of the proposed system are: -
1. Input the original signal
2. Split the original signal to segments, where each segment has an equal samples, and independent from each other.
3. Choose an interval for the random mixing matrix.
4. The result is the scramble speech, as in equation (1). Where the steps from 2-4 represent the scramble speech process.
5. Using Principal Component Analysis (PCA), this can be accomplished by scaling the vector elements by the inverses of the eigenvalues of the correlation matrix as in equation (7). The result is the descramble speech. Fig.. 1 represents the proposed system.

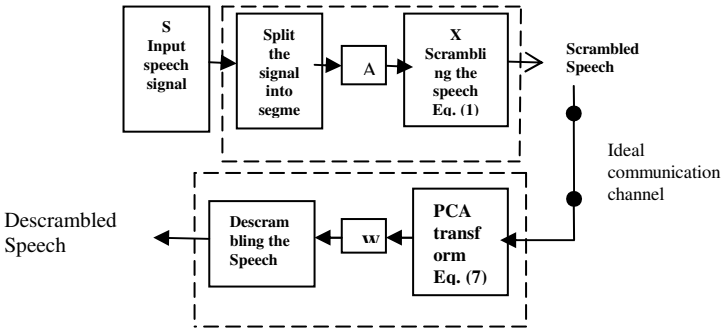

Fig. 1. The Proposed System.

To test the validity of the above proposed system, we take the speech signal, with the variety of interval of random mixing matrix, samples, and frequency samples. Two measures are used for the original and scrambled speech, these are: SNR, and LPC distance as shown in Table I.

Through tests we implement the experiments on different separate Arabic words. Fig. 2 up to Fig. 10 provides the original, scrambled and descrambled forms of the tested signals. While the Table I provide the performance parameters (SNR and LPC distance) measured at different sampling frequencies. It is evident from this table as the sampling frequency decreased the SNR and LPC increased.

Table I. Performance parameters (SNR and LPC distance) measured at different sampling frequencies and only for two sources.

Speech	Interval of random mixing matrix	Length of Segments	Sampling Frequency	SNR	LPC Distance	Time (Sec)
alk11	**[0.040314 0.5689 0.67709 -0.25565]**	**16064**	**44100**	**4.1510**	**0.000114**	**2.461**
mas01	**[-27.886 -37.435 -4.9772 -32.236]**	**17152**	**44100**	**4.7757**	**0.0135**	**2.858**
mas8	**[8.4336 56.186 27.783 20.493]**	**3111**	**8000**	**8.4134**	**0.0192**	**2.015**

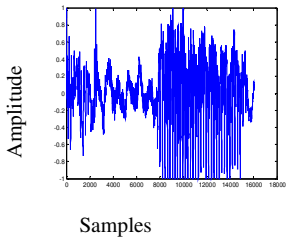

Fig. 2. the Original speech alk11

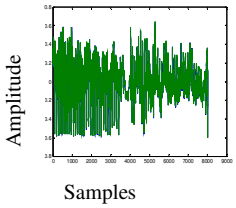

Fig. 3, the Scrambled alk11 Signal

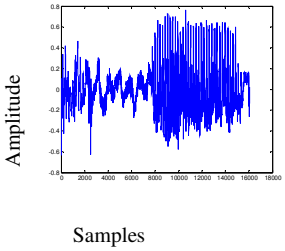

Fig. 4, the Descrambled alk11 Signal

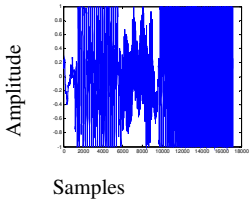

Fig. 5, the Original mas01 Signal

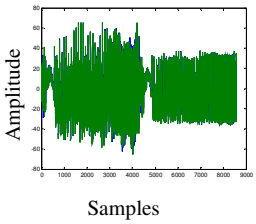

Fig. 6, the Scrambled mas01 Signal

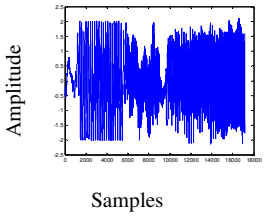

Fig. 7, the Descrambled mas01 Signal

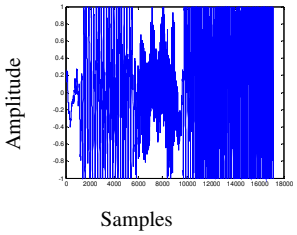

Fig. 8, the Original mas8 Signal

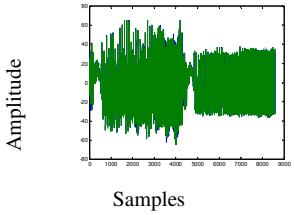

Fig. 9, the Scrambled mas8 Signal

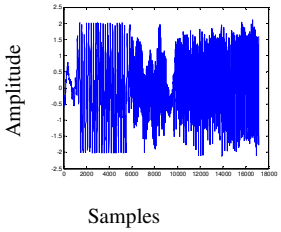

Fig. 10, the Descrambled mas8 Signal

IV. CONCLUSIONS

From the evaluation tests performed on the scrambled and descrambled speech, we conclude that the new proposed speech scrambling algorithm satisfies the followings:

- Unintelligibility of the tested scrambled speech.
- The proposed scrambling process is bandwidth preserving operation.
- It is very difficult to decrypt. Hence it is cryptanalytically strong or secure.
- Communication delay caused by scrambling process was small.
- The recovered speech signal at the receiving side was of the good quality (through hearing) and preserves both the intelligibility of the speech and the characteristics of the speaker.

REFERENCES

[1] Qiu-Hua Lin, Fu-Liang Yin, Tie-Min Mei, and Hualou Liang, "A Blind Source Separation Based Method for Speech Encryption", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—I: REGULAR PAPERS, VOL. 53, NO. 6, JUNE 2006
[2] H. J. Beker and F. C. Piper, "Secure Speech Communications". London, U.K.: Academic, 1985.
[3] B. Goldburg, S. Sridharan, and E. Dawson, "Design and cryptanalysis of transform-based analog speech scramblers," *IEEE J. Select. Areas Commun.*, vol. 11, no. 5, pp. 735–744, May 1993.
[4] A. Matsunaga, K. Koga, and M. Ohkawa, "An analog speech scrambling System using the FFT technique with high-level security," *IEEE J. Select. Areas Commun.*, vol. 7, no. 4, pp. 540–547, Apr. 1989.
[5] F. L. Ma, J. Cheng, and Y. M. Wang, "Wavelet transform-based analogue speech scrambling scheme," *Electron. Lett.*, vol. 32, no. 8, pp. 719–721, 1996.
[6] K. Li, Y. C. Soh, and Z. G. Li, "Chaotic cryptosystem with high sensitivity to parameter mismatch," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 50, no. 4, pp. 579–583, Apr. 2003.
[7] G. Manjunath and G. V. Anand, "Speech encryption using circulant transformations," *Proc. IEEE Int. Conf. Multimedia and Expo*, vol. 1, pp. 553–556, 2002.
[8] Max Power, Power Broadcasting, HireMe.geek.nz," An Overview of Structural Design Issues in Theoretical and Practical Implementations of Analog Speech Scrambling Systems", Spring 2003 (Original content).
[9] A.S. Bopardikar, "Speech Encryption using wavelet packet", Master thesis, Indian Institute of Science, 2000.
[10] Lindsay I Smith, "A tutorial on Principal Components Analysis", 2002.
[11] Jonathon Shlens, "A Tutorial on Principal Component Analysis", 2005
[12] Te-Ming Huang, Vojislav Kecman, and Ivica Kopriva,"Kernel Based Algorithms for Mining Huge Data Sets Supervised, Semi-supervised, and Unsupervised Learning", Springer-Verlag Berlin Heidelberg 2006
[13] Principal component analysis From Wikipedia, the free encyclopedia
[14] Signal-to-noise ratio From Wikipedia, the free encyclopedia
[15] Linear predictive coding From Wikipedia, the free encyclopedia

Dr. Nidaa A. Muhsin (M'08) was born in Basraha, Iraq in 1968. She received the B.Sc. in computer science from Basraha University, Iraq in 1990, and M.Sc. in computer science from University of Babylon, Iraq in 1999, and Ph.D. in computer science from the University of Technology, Iraq in 2006.

She is currently a lecturer in computer science department in University of Babylon. Her main research interests in adaptive signal processing, ICA, data compression, speech scrambling, Steganography, steganalysis.