Chapter 8
ICA and PCA–Based Cryptology

Sattar B. Sadkhan Al Maliky
University of Babylon, Iraq

Nidaa A. Abbas
University of Babylon, Iraq

ABSTRACT

Blind Source Separation (BSS) represented by Independent Component Analysis (ICA) has been used in many fields such as communications and biomedical engineering. Its application to image and speech encryption, however, has been rare. In this chapter, the authors present ICA and Principal Component Analysis (PCA) as a category of BSS-based method for encrypting images and speech by using Blind Source Separation (BSS) since the security encryption technologies depend on many intractable mathematical problems. Using key signals, they build a suitable BSS underdetermined problem in the encryption and then circumvent this problem with key signals for decoding. The chapter shows that the method based on the BSS can achieve a high level of safety right through building, mixing matrix, and generating key signals.

INTRODUCTION

At first we would like to touch on the subject of Principal Component Analysis (PCA) because it is older scientifically than ICA. Principal Component Analysis (PCA) is a statistical method that consists in reducing the dimensions of the data set through linear combinations. The idea is to search for linear combinations that allow separating at best the possible values of the data. In other words, PCA searches for linear combinations with the largest variances, and when several linear combinations are needed, it considers the variances in decreasing order of importance.

The power of PCA is that it reduces the dimension of the data set in such a way that it conserves the most interesting components of the data set. In the signal processing world the statistical transformation involved in PCA is often called the Karhunen-Loeve transformation, another often used name is the Hotelling transformation. The major disadvantage of this transformation is its huge computation time. As this transformation has no fixed basis vectors, in the case of a Fourier transformation, the basis vectors that we will obtain are linked to the data set. (Lilian Bohy & Jean-Jacques Quisquater, 2003)
On the other hand, Independent Component Analysis (ICA) belongs to a class of BSS methods to separate data into underlying informational components, where such data can take the form of images, sounds, telecommunication channels or stock market prices.

ICA defines a generative model for the observed multivariate data, which is typically given as a large database of samples. In the model, the data variables are assumed to be linear or nonlinear mixtures of some unknown latent variables, and the mixing system is also unknown. The latent variables are assumed nonGaussian and mutually independent and they are called the independent components of the observed data. These independent components, also called sources or factors, can be found by ICA.

ICA can be seen as an extension to principal component analysis (PCA) and factor analysis (FA). ICA is much more powerful technique, however, capable of finding the underlying factors or sources when these classic methods fail completely.

The data analysis by ICA could originate from many different kinds of application fields, including digital images and document databases, as well as economic indicators and psychometric measurements. Frequently, the measurements are given as a set of parallel signals or time series; the term BSS is used to characterize this problem. Typical examples are mixtures of simultaneous speech signals that have been picked up by several microphones, brain waves recorded by multiple sensors, interfering radio signals arriving at a mobile phone, or parallel time series obtained from some industrial process.

ICA is subjected to a variety of disciplines, such as statistics, signal processing, neural networks, applied mathematics, neural and cognitive sciences, information theory, artificial intelligence, and engineering. (A. Hyvrenen & E. Oja, 2001)

Blind Source Separation (BSS) is a technique of array signal processing which aims to recover a set of unknown mutually independent source signals from their observed mixtures without knowing the mixing coefficients.

ICA is a special case of BSS with linear and instantaneous mixture. BSS has hopeful applications to communications, biomedical engineering, cocktail party problem, etc. Its applications to signal encryption focused on the speech and image encryption (W. Kasprzak & A. Cichocki, 1996; Qiu-Hua Lin & Hua-Lou Liang, 2004).

With the rapid development of multimedia and networking technologies, the security of multimedia data becomes more and more important in many real applications. To meet such an increasing demand, during the past two decades many encryption schemes have been proposed for protecting multimedia data, including speech signals, images and videos. (Shujun Li & Guanrong Chen, 2008).

This chapter is devoted to study two statistical tools, namely the Principle Component Analysis (PCA) and the Independent Component Analysis (ICA) and their roles in cryptology. The basic features of the ICA and PCA models in many speech and image encryption applications are examine.

BACKGROUND

Cryptology, the study of cryptosystems, can be divided into three disciplines. Cryptography is interested in the design of cryptosystems; cryptanalysis studies the breaking of cryptosystems, while security evaluation is interested in the design of different methods suitable to evaluate the security (complexity) of the crypto system under consideration. These three disciplines are closely related; when setting up a cryptosystem, the analysis of its security plays an important role and evaluation of its strength against analysis is also an important aspect.

Modern communication techniques, using computers connected through networks, make all data even more vulnerable to these threats. Also, new issues have come up that were not relevant