DOKUZ EYLÜL UNIVERSITY

GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

BLIND SOURCE SEPARATION USING SUPPORT VECTOR MACHINES

by

Gülçin YAVAŞ

December, 2006

İZMİR

BLIND SOURCE SEPARATION USING SUPPORT VECTOR MACHINES

A thesis Submitted to the

Graduate School of Natural and Applied Sciences of Dokuz Eylül University

In Partial Fulfillement of the Requirements for the Degree of Master of Science in

Electrical and Electronics Engineering, Electrical and Electronics Engineering Program

by

Gülçin YAVAŞ

December, 2006

İZMİR

M.Sc THESIS EXAMINATION RESULT FORM

We have read the thesis entitled **"BLIND SOURCE SEPARATION USING SUPPORT VECTOR MACHINES"** completed by **GÜLÇİN YAVAŞ** under supervision of **Asst. Prof. Dr. NALAN ÖZKURT** and we certify that in our opinion it is fully adequate, in scope and in quality, as a thesis for the degree of Master of Science.

Yrd. Doç. Dr. Nalan ÖZKURT

Supervisor

Prof. Dr. Cüneyt GÜZELİŞ	Yrd. Doç. Dr. Radosveta SOKULLU
(Jury Member)	(Jury Member)

Prof.Dr. Cahit HELVACI
Director
Graduate School of Natural and Applied Sciences

ACKNOWLEDGEMENTS

I would like to thank my advisor Asst. Prof. Dr. Nalan Özkurt for her valuable and guidance and patient during the preparation of this thesis. I would like to Express my sincere appreciate to all friends and my dear friend Dilek İnceköse for their endless support and friendship. Mostly, I thank to my family for their endless support.

Gülçin YAVAŞ

BLIND SOURCE SEPARATION USING SUPPORT VECTOR MACHINES

ABSTRACT

Blind source separation (BSS) can be defined as the recovering of source signals from their mixture signals without any prior knowledge of the source and mixing environment. The type of the mixing environment may be linear or nonlinear and this determines the type of the BSS problem. Independent component analysis (ICA) method solves linear BSS by finding a representation for a multivariate data that minimizes the statistical dependency between signal components, but ICA can not find solution for nonlinear BSS. The kernel-based methods have been implemented to solve the nonlinear BSS problems.

The temporal predictability method that is based on the measure of temporal predictability solves the linear BSS problems, but it can not achieve the nonlinear BSS. The kernelized temporal predictability method is based on applying the kernel methods in support vector machines (SVMs) to the temporal predictability method.

In this thesis, the linear and nonlinear BSS applications have been simulated in Matlab numeric analysis software package by using the linear and kernelized temporal predictability methods, respectively. In addition to sound separation, the algorithms has been extended for the image separation problem newly in this study. The statistical performance of the algorithms have been simulated using different algorithm parameters like kernel degree, kernel type, image mass size and nonlinear mixing function.

Keywords: Blind source separation, independent component analysis, temporal predictability, kernelized temporal predictability**.**

DESTEK VEKTÖR MAKiNALARIYLA GÖZÜ KAPALI KAYNAK AYRIŞIRILMASI

ÖZ

Gözü kapalı kaynak ayrımı, kaynak sinyallerinin karışım sinyallerinden hiçbir kaynak sinyal ve karıştırıcı ortam bilgisi olmadan yeniden elde edilmesi olarak tanımlanabilir. Kaynak karıştırıcı ortam tipi doğrusal veya doğrusal olmayabilir, bu bilgi gözü kapalı ayrımı probleminin tipini belirler. Bağımsız bileşen analizi çok değişkenli bir veri için sinyal bileşenleri arasında istatiksel bağımlılığı en az hale getirerek doğrusal gözü kapalı kaynak ayrıştırma problemini çözer, fakat doğrusal olmayan durumu çözemez.

Zamansal tahmin edilebilirlik bilgisinin ölçümü üzerine kurulmuş olan zamansal tahmin edilebilirlik metodu doğrusal gözü kapalı kaynak ayrıştırma problemlerini çözebilir fakat doğrusal olmayan problemi çözemez. Çekirdek tabanlı zamansal tahmin edilebilirlik metodu destek vektör makinelerindeki (DVM) çekirdek metotlarının zamansal tahmin edilebilirlik metoduna uygulanması şeklinde oluşturulmuştur.

Bu tezde, doğrusal ve doğrusal olmayan gözü kapalı kaynak ayrıştırılması uygulamalarının simülasyonları, doğrusal ve doğrusal olmayan zamansal tahmin edilebilirlik metotları kullanılarak Matlab nümerik analiz yazılım paketinde yapılmıştır. Bu çalışmada ses ayrıştırmasına ek olarak algoritmalar görüntü ayrıştırması içinde genişletildi. Algoritmaların istatiksel performans analizi farklı çekirdek dereceleri, tipleri ve görüntü hücre büyüklüğü ve doğrusal olmayan karışım fonksiyonları için yapılmıştır.

Anahtar Kelimeler: Gözü kapalı kaynak ayrıştırma, bağımsız bileşen analizi, zamansal tahmin edilebilirlik, kernelleştirilmiş zamansal tahmin edilebilirlik.

CONTENTS

Page

CHAPTER ONE
INTRODUCTION

1.1 Introduction

Blind source separation (BSS) can be explained as the problem of recovering original source signals from several observed mixtures signals. The observed mixtures are typically obtained at the outputs of sensors and each sensor receives a different combination of source signals. This problem is called as "cocktail party problem" in the literature. The lack of prior knowledge of mixing environment is compensated with the main working principle assumed that is the statistical independency of source signals. This approach makes BSS versatile for different types of applications (Cardoso, 1998).

The key working principle of BSS is the assumption of the statistical independency of source signals. Independent component analysis (ICA) finds a desired representation for a multivariate data and this representation minimizes the statistical dependency between source signal components. ICA finds solution for linear BSS problems but the mixing environment is generally nonlinear for most of the applications. For solving the nonlinear BSS problems, new algorithms based on kernel methods have been developed.

Nonlinear BSS problems are solved using the kernel-based learning algorithms which are summarized in (Müller, Rätsch, Tsuda & Schölkopf, 2001). The most popular kernel-based learning algorithms are support vector machines (SVMs), kernel Fischer discriminant and kernel principal component analysis. Kernel-based methods maps the original data to a higher dimension feature space by applying a nonlinear mapping function. The nonlinear mapping can be expressed as $\phi : R^N \to F$ and $x \to \phi(x)$. The source separation algorithm is now considered in F feature space instead of R^N. The curse of dimensionality is a key concept in statistics and states the fact that difficulty of the estimation problem increases in a drastic manner with the increase of the data dimension. This concept conflicts with the kernel-based methods but finding simpler classification function makes the learning in F easy. The functions used in F must be simpler because of the intractability of the algorithm in F. A linear function in F corresponds the nonlinear function in input space. The kernel function makes the computations easier using the kernel trick for scalar dot products.

The temporal predictability method is particular type of ICA method. The algorithm uses the measure of temporal predictability to separate the source signals. The algorithm solves the linear BSS problems (Stone, 2001). The application of the method has been extended for the nonlinear BSS problems via kernelizing the algorithm. Using the facility of kernel methods, the kernelized temporal predictability algorithm has been developed (Martinez & Bray, 2003).

In addition to separation of sound signals, blind source separation of image mixtures have also been studied in some applications. The motivation of applying source separation to image mixtures comes from the recovery of the corrupted writings or images. In literature, there are several works deals with the image recovery problem with different kinds of methods in (Liang, Ahmadi & Shridbar 1997) (Almeida &Almeida, 2004) (Kasturi, Bow, El-Masri, , Shah, Gattiker & Mokate, 1990). Different types of algorithms have been applied for different types of applications. Some of them are separation of text and image, image and writings etc.

1.2 Aim of the Thesis

In this thesis, the performance analysis of the linear and nonlinear source separation of sound mixtures using temporal predictability and kernelized temporal predictability based on the source codes developed in (Stone, 2001; Martinez and Bray, 2003) has been accomplished. In order to test algorithms different types of sound signals have been used: First signals are TIMIT speech signals that are taken randomly from the TIMIT database. TIMIT database can be requested from the web address of http://www.mpi.n1/world.ltg/corpara/timit/timit.html. The second type of the sound signals are the musical notes that were taken from the IOWA database (http://theremin.music.uiowa.edu/index.html) that has been rearranged by M. Erdal Ozbek in Inst. Supelec. and the third type of the sound signals are recorded by from daily song signals. The types of kernel functions and parameters have been analyzed using the simulation results of different data types. Different nonlinear source signals mixing functions have been simulated for all sound data types. FastICA method has been simulated for the linear BSS for the sound and image data sets to compare with the temporal predictability method.

The temporal predictability and kernelized temporal predictability based blind source separation methods have also been applied to image mixtures. The image data sets consist of the linear and nonlinear mixtures of different objects, objects and writings; and objects and scribbles. These sets are chosen in order to simulate the recovery of the corrupted images by scanning process or human corruptions. As in the sound separation case, different mixture types and different kernel parameters have been applied. Additionally, the image mass size has been considered as a performance parameter.

1.3 Outline of the Thesis

In Chapter 2, blind source separation, independent component analysis and support vector machines topics have been given to state the problem and solution approaches.

In Chapter 3, the temporal predictability method has been expressed in detail for both linear and the nonlinear case approaches.

In Chapter 4, FastICA method has been simulated for the linear and nonlinear BSS problem of sound and image data. The temporal predictability method and the kernelized temporal predictability method have been simulated for different types of sound data and image data. Different types of kernel function with kernel parameters ranging some interval have been simulated.

In Chapter 5, the temporal predictability method and the kernelized temporal predictability method have been simulated for different types image data. Different types of kernel function with kernel parameters ranging some interval have been simulated.

In Chapter 6, conclusions have been given.

CHAPTER TWO

BLIND SOURCE SEPARATION, INDEPENDENT COMPONENT ANALYSIS AND SUPPORT VECTOR MACHINES

2.1 Introduction

In this chapter, the main topics of the thesis that are blind source separation, independent component analysis and support vector machines have been explained. The blind source separation for the linear and the nonlinear mixing environments has been given in section 2.2. Independent component analysis (ICA) methods and its working principles have been expressed in section 2.3. One special ICA method FastICA has been introduced in section 2.3.1. ICA solves the linear BSS problems. For the nonlinear BSS, the kernel-based methods have been used. A support vector machines (SVMs) has been given in section 2.4 to introduce the kernel approach for nonlinear separation.

2.2 Blind Source Separation

Blind source separation is the problem of recovering of original source signals from mixtures of the source signals. There are a number of different types of signals emitting in physical environments like electrical signals, speaking people in the same room, mobile phones emitting the radio waves etc. The main point is that there are a number of receivers or sensors that collect that emitting signals and their positions are different so that each receiver or sensor collects a mixture of source signals with different weights for each source signal.

A set of source signals $\mathbf{s} = \{s_1, s_2, \ldots, s_K\}^t$, where ith row signal in $\mathbf{s}$ is s_i signal measured at n time points. A set of linear mixture signals $\mathbf{x} = \{x_1, x_2, \ldots, x_M\}^t$ where $M \geq K$. The signals in s can be formed with an MxK mixing matrix A: $\mathbf{x}=As$. If the rows of A are linearly independent, each s_i signal can be recovered with a 1xM matrix W_i by applying the formation $s_i=W_i x$. The problem to be pointed here to recover the source signals is finding an unmixing matrix $W= \{W_1|W_2|\ldots|W_K\}^t$ such that each row recovers a different signal y_i, y_i is a scaled version of source signal s_i.

If the number of the sources and the observed signals have been assumed as three, then the equations for the mixture signals in terms of s and A can be given as in the following equation.

$$\begin{aligned} x_1(t) &= a_{11}s_1(t) + a_{12}s_2(t) + a_{13}s_3(t) \\ x_2(t) &= a_{21}s_1(t) + a_{22}s_2(t) + a_{23}s_3(t) \\ x_3(t) &= a_{31}s_1(t) + a_{32}s_2(t) + a_{33}s_3(t) \end{aligned} \tag{2.1}$$

The a_{ij} are constant coefficients of the unknown mixing matrix A. The source signals are recovered from the mixtures signals by using the independency of source signals. By using the unmixing matrix W, the source signals are recovered.

$$\begin{aligned} s_1(t) &= w_{11}x_1(t) + w_{12}x_2(t) + w_{13}x_3(t) \\ s_2(t) &= w_{21}x_1(t) + w_{22}x_2(t) + w_{23}x_3(t) \\ s_3(t) &= w_{31}x_1(t) + w_{32}x_2(t) + w_{33}x_3(t) \end{aligned} \tag{2.2}$$

2.3 Independent Component Analysis

A central problem in neural network research, as well as in statistics and signal processing, is finding a suitable representation of the data, by means of a suitable transformation. It is important for subsequent analysis of the data, whether it be pattern recognition, data compression, de-noising, visualization or anything else, that the data is represented in a manner that facilitates the analysis. A recently developed linear transformation method is the Independent Component Analysis (ICA) method that seems to capture the essential structure of the data in many applications.

Independent Component Analysis (ICA) is a computational technique for revealing hidden factors that underlie sets of measurements (Hyvärinen, 1999). ICA assumes a statistical model whereby the observed multivariate data, typically given as a large database of samples, are assumed to be linear or nonlinear mixtures of some unknown latent variables. The mixing coefficients are also unknown. The latent variables are nongaussian and mutually independent and they are called the independent components of the observed data. By ICA method independent components can be found. Thus ICA can be assumed to be an extension to Principal Component Analysis and Factor Analysis.

2.3.1 Motivation to understand Independent Component Analysis (ICA) method

Imagining a room where two people are speaking simultaneously. There are two microphones, which are held in different locations. The microphones give two recorded time signals, which could be denoted by $x_1(t)$ and $x_2(t)$, with x_1 and x_2 the amplitudes, and t the time index. Each of these recorded signals is a weighted sum of the speech signals emitted by the two speakers, which we denote by $s_1(t)$ and $s_2(t)$. We could express this as a linear equation:

$$x_1(t) = a_{11}s_1 + a_{12}\,s_2 \quad (1) \qquad (2.3)$$
$$x_2(t) = a_{21}s_1 + a_{22}\,s_2 \quad (2).$$

a_{11}, a_{12}, a_{21} and a_{22} are some parameters that depend on the distances of the microphones from the speakers. It would be very useful if you could now estimate the two original speech signals $s_1(t)$ and $s_2(t)$, using only the recorded signals $x_1(t)$ and $x_2(t)$. This is called the cocktail-party problem. Any time delays or other extra factors from the simplified mixing model are omitted.

An illustration can be made to exemplify the cocktail party problem as follows; the two original source signals are shown in (Figure 2.1.)

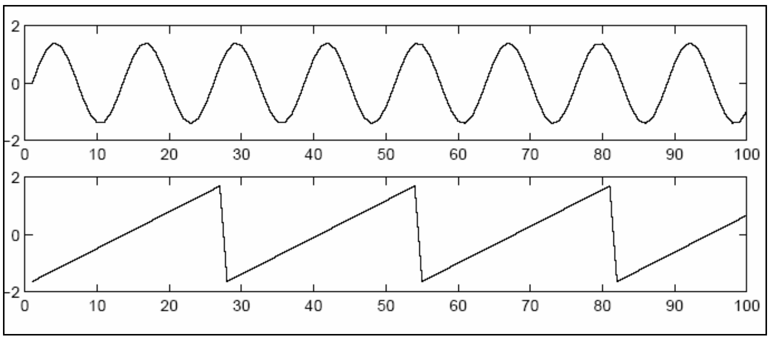

Figure 2.1 The source signals

The mixed signals of the two original source signals looks like below:

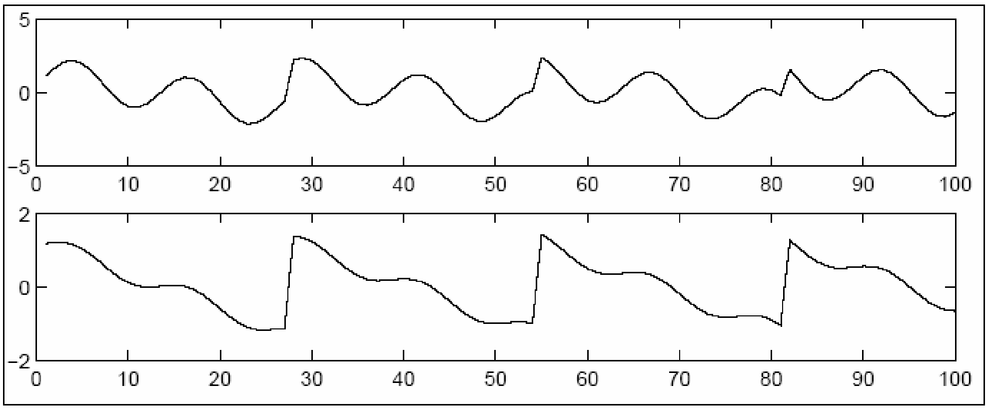

Figure 2.2 The mixture signals

The problem is to recover the source signals using only the observed mixture signals. If the parameters a_{ij} were known, the linear equation in (1) could be solved by classical methods. One approach to solving this problem would be to use some information on the statistical properties of the signals $s_i(t)$ to estimate the a_{ii}. The question is that assuming $s_1(t)$ and $s_2(t)$, at each time instant t, are statistically independent. This is not an unrealistic assumption in many cases, and it need not be exactly true in practice. The recently developed technique of independent component analysis (ICA), can be used to estimate the a_{ij} based on the information of their independence, which allows us to separate the two original source signals $s_1(t)$ and $s_2(t)$ from their mixtures $x_1(t)$ and $x_2(t)$ (Hyvärinen, 1999). Figure 2.3 represents the two signals estimated by the ICA method. As can be seen, these are very close to the original source signals (their signs are reversed, but this has no significance.) Independent component analysis was originally developed to deal with problems that are closely related to the cocktail-party problem.

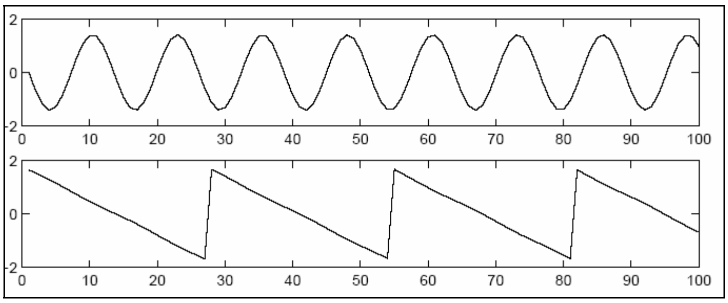

Figure 2.3 The recovered source signals

It is assumed that the data vector x is distributed according to the ICA model as x=As, it is a mixture of independent components. All the independent components are assumed to have identical distributions. To estimate one of the independent components, it is considered that a linear combination of the xi, y=w^{T}x=$\sum_i w_i x_i$, where w is a vector to be determined. If w were one of the rows of the inverse of A, this linear combination would actually equal to one of the independent components.

2.3.2 FastICA

FastICA algorithm has been implemented depending on the principles of measuring nongaussianity and it has been extended with different types of algorithms.By using the FastICA graphic user interface package, the sound and image data applications have been simulated for both linear and nonlinear blind source separation applications in section 4.2. The figure of FastICA package has been represented in Figure 2.4.

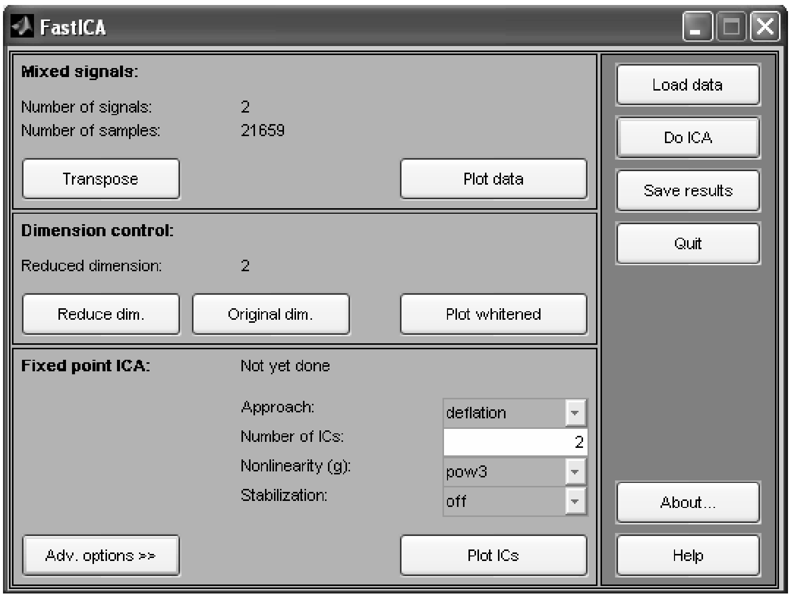

Figure 2.4 FastICA package user interface

The setting parameters of the FastICA software package are approach, number of ICs, nonlinearity function (g) and stabilization. The four different types of nonlinearity functions are available, pow3 that is the third power of the term, tanh, gauss and skew that is the square of the term.

The approach parameter is decorrelation approach. It can be set to deflation or symmetric. The number of ICs defines the number of independent components to be estimated. The nonlinearity item determines the nonlinear function that will be used in fixed-point algorithm.

2.4 Support Vector Machines

The main idea of support vector machine is to construct a hyper plane as the decision surface that enables maximizing the distance between negative and positive examples is maximized. The distance term is referred as margin of separation.

Linearly separable data can be separated by in infinite number of linear hyper planes that can be written as $f(x, w) = w^T x + b$ .The problem is to find the optimal separating hyper plane. Optimal separating plane is the one that has the largest margin.

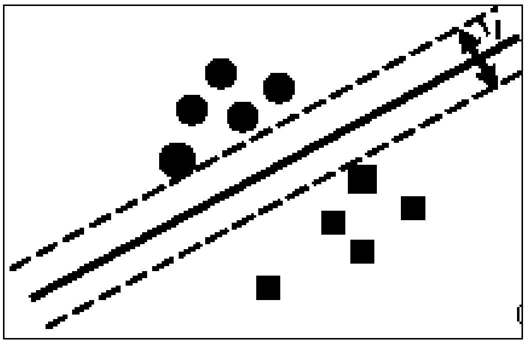

Figure 2.5 Optimal Hyper Plane with Margin

Margin is defined by w as follows, $M=\frac{2}{\|w\|}$.

The problem is to solve:

Margin Maximization $J=w^T w= ||w||^2$.

Correct Classification $y_i (w^T w+b) \geq 1$.

This is a classic QP(Quadratic programming) problem with constraints that ends in forming and solving of a primal and / or dual Lagrangian. The optimal canonical separating hyper plane (OCSH), a separating hyper plane with the largest margin specifies support vectors, training data closest to it, which satisfy y_i $(w^Tw+b) \equiv 1$, $j=1,N_{sv}$. At the same, the OCSH must separate data correctly .It satisfy the inequalities

$$y_i\,(w^Tw+b) \geq 1 \qquad i=1,L.$$

Where L denotes a number of training data and N_{sv} stands for a number of support vectors. Maximization of M means a minimization of $||w||$.

$w^Tw= \sqrt{w_{12}+w_{22}+\ldots+w_{n2}}$ leads to a maximization of a margin M. Because square root is a monotonic function, its minimization is equivalent to a minimization of function itself.

However, the hyper planes cannot be the solutions when the decision boundaries are non-linear. Support Vector Machines should be constructed by mapping input vectors nonlinearly into a high dimensional feature space, and by constructing OCSH in the high dimensional feature space.

Non linear Support Vector Classification :

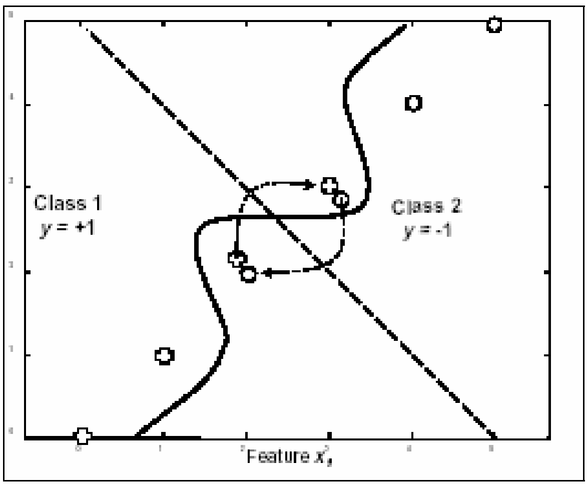

Figure 2.6 Non linear Separation of Class1 and Class2

Mapping the input data into feature space on a classic XOR (nonlinearly separable) problem. Many different nonlinear discriminant functions that separate 1-s from 0-s can be drawn in a feature plane. Suppose, the following one.

$f(x)=x_1+x_2-2x_1x_2-1/3 \qquad x_3=x_1x_2 \qquad f(x)= x_1+x_2-2x_3-1/3.$

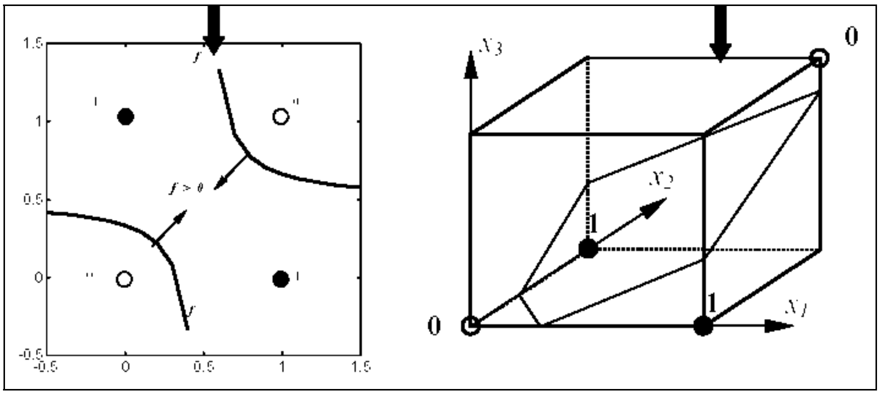

Figure 2.7 Non linear Separation of Class1 and Class2

The hyper plane in feature space has been converted to the linear form by applying the nonlinear function. The key has been represented here is the kernel trick of the SVMs and has been applied all kernel-based algorithms.

The kernel trick transforms any algorithm that solely depends on the dot product between two vectors. Wherever a dot product is used, it is replaced with the kernel function. Thus, a linear algorithm can easily be transformed into a non-linear algorithm. This non-linear algorithm is equivalent to the linear algorithm operating in the range space of φ. However, because kernels are used, the φ function is never explicitly computed (Aizerman, Braverman & Rozonoer, 1964).

CHAPTER THREE - TEMPORAL PREDICTABILITY METHOD AND KERNELIZED TEMPORAL PREDICTABILTY METHOD

3.1 Introduction

In chapter 2, blind source separation (BSS) problem has been expressed and independent components analysis method has been expressed briefly as a solution to BSS problem. A special ICA algorithm, FastICA method has been expressed in section 2.3.1 (Hyvärinen, 2003).

The temporal predictability method for linear blind source separation and the kernelized temporal predictability method for nonlinear blind source separation have been given in section 3.2 and 3.3, respectively. In chapter 4 and 5, all these methods have been simulated for different types of source signals.

3.2 Temporal Predictability Method

A linear mixture of statistically independent source signals has the property that the temporal predictability of any signal mixture is less than (or equal to) that of any of its component source signals. This property is used to recover the original source signals from any linear mixture signal by finding an unmixing matrix that maximizes a measure of temporal predictability for each recovered signal (Stone, 2003).

The method for recovering the source signals is based on the following principle assumed: the temporal predictability of a signal mixture x_i is usually less than that of any of the source signals contributing the mixture signal x_i. A measure F (W_i, x) predictability of temporal is defined to estimate the relative predictability of a signal y_i recovered by a matrix W_i, where $\mathbf{y_i}=W_i x$. If the source signals are more predictable, then the value of W_i which maximizes the predictability of an extracted signal $\mathbf{y_i}$ gives one of the source signal.

The definition of signal predictability F used is:

$$\mathrm{F(W_i,x)=log}\frac{\mathrm{V(W_i,x)}}{\mathrm{U(W_i,x)}}=\log\frac{\mathrm{V_i}}{\mathrm{U_i}}=\log\frac{\sum_{\tau=1}^{n}(\bar{y}_\tau-y)^2}{\sum_{\tau=1}^{n}(\tilde{\bar{y}}_\tau-y)^2}. \quad (3.1)$$

Where y_τ = Wi x_τ is the value of y at time τ, and x_τ is a vector of K signal mixture values at time τ (Stone, 2001). The term U_i reflects the fact that y_τ is predicted by a short-term moving average of values in y. The term V_i is a measure of the overall variability in y, as measured by which y is predicted by a long-term moving average y_τ of values in y. The long-term and short-term predicted values are expressed in equation 3.2.

$$\begin{aligned}\tilde{y} &= \lambda_s\tilde{y}_{(\tau-1)}+(1-\lambda_s)y_{(\tau-1)}: \qquad 0\le\lambda_s\le 1\\ \bar{y} &= \lambda_L\bar{y}_{(\tau-1)}+(1-\lambda_L)y_{(\tau-1)}: \qquad 0\le\lambda_L\le 1.\end{aligned} \quad (3.2)$$

The relation between a half-life h and the parameter λ is defined as $\lambda=2^{-1/h}$. The half life of h_L of λ_L is much longer than the corresponding half life of h_S of λ_S.

Maximizing the numerator term **V** of temporal predictability measure **F** to maximize the predictability measure term results in a high variability with no constraint on temporal structure. And minimizing the denominator term **U** to maximize **F** results in a DC signal. V_i/U_i term can be maximized only if the following two conditions are satisfied: (1) **y** has a nonzero range and (2) the values in **y** change slowly over time.

3.2.1 Extracting a signal

A scalar signal mixture y_i formed via $y_i=W_ix$. The equation can be written as:

$$F=\log\frac{\mathrm{W_i\bar{C}W_i^t}}{\mathrm{W_i\tilde{C}W_i^t}} \quad (3.3)$$

In the equation 3.3, $\bar{C}$ is an MxM matrix of long-term covariance between signal mixtures and $\tilde{C}$ is a corresponding matrix of short-term covariances. The long-term and short-term covariance matrices between the ith and jth mixtures are defined as:

$$\begin{aligned} \tilde{C}_{ij} &= \sum_{\tau}^{n} (x_{i\tau} - \tilde{x}_{i\tau})(x_{j\tau} - \tilde{x}_{j\tau}) \\ \bar{C}_{ij} &= \sum_{\tau}^{n} (x_{i\tau} - \bar{x}_{i\tau})(x_{j\tau} - \bar{x}_{j\tau}). \end{aligned} \tag{3.4}$$

The derivative of **F** with respect to the W_i is:

$$\Delta_{W_i} F = \frac{2W_i}{V_i}\bar{C} - \frac{2W_i}{U_i}\tilde{C}. \tag{3.5}$$

And to find the maximum point the left side of the equation 3.5 is equaled to zero, so the equality below is get.

$$W_i\bar{C} = \frac{V_i}{U_i} W_i\tilde{C}. \tag{3.6}$$

Extrema in **F** is the values of W_i and solutions for W_i can be obtained as eigenvectors of the matrix ($\tilde{C}^{-1}\bar{C}$), with corresponding eigenvalues $\gamma_i = V_i/U_i$ (Stone, 2001). Using the eigenvalue function eig($\tilde{C}$, $\bar{C}$). Applying this for all of the mixture signals' components get the all source signals recovered.

The mixing matrix A is obtained by using *randn* function Matlab. The short-term and long-term half-lives are used in the following applications are h_S=1 and h_L=9000 (Stone 2001).

3.2.2 Examples

In order to show the capability of the Stone's algorithm, some examples which are given in (Stone, 2001) will be introduced in this section. The detailed application testing the performance of the algorithm will be given in Chapter 4 and 5.

3.2.2.1 Separation of Mixture of Signals Having Different Probability Density Functions

The three source signals s= $\{s_1|s_2|s_3\}^t$ are a super Gaussian signal that is a gong signal, a sub Gaussian signal that is a sine wave and a Gaussian signal. Three source signals are mixed by a random mixing matrix A to get the three signal mixtures: **x**=As. Each signal consists of 3000 samples. And using the temporal predictability measurement technique which is implemented by J.V. Stone, the following correlation coefficients are resulted.

Table 3.1 The correlation coefficients between original source signals and recovered signals of simulations

Recovered Signals	Source Signals		
	s_1	s_2	s_3
y_1	0.000	0.001	1.000
y_2	1.000	0.000	0.000
y_3	0.042	0.999	0.002

3.2.2.2 Separation of Sound Mixtures

In this example, there are two sets of eight sounds: male and female voices and classical music with and without singing. The source signals are sampled at 44100 Hz and each signal consists of 50000 data points. The absolute values of correlation coefficients between the recovered signals and original source signals are shown in the following table.

Table 3.2 The correlation coefficients between eight source signals and recovered eight signal by the method (Stone, 2001)

Signal Recovered	Source Signals (voices)							
	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8
y_1	0.001	0.008	0.004	0.003	0.028	0.046	**0.988**	0.143
y_2	**0.994**	0.013	0.003	0.001	0.001	0.017	0.016	0.109
y_3	0.179	0.001	0.162	0.011	0.037	0.102	0.134	**0.956**
y_4	0.015	0.012	0.007	**0.999**	0.024	0.032	0.004	0.004
y_5	0.004	0.020	**0.993**	0.000	0.021	0.008	0.007	0.109
y_6	0.010	0.003	0.026	0.018	0.021	**0.992**	0.044	0.111
y_7	0.027	**0.999**	0.012	0.002	0.000	0.010	0.009	0.002
y_8	0.015	0.003	0.027	0.022	**0.998**	0.020	0.021	0.043

Note: Each source signal has a high correlation coefficient with only one recovered signal.

Table 3.3 The correlation coefficients between eight source signals and recovered eight signal by the method (Stone, 2001)

Signal Recovered	Source Signals (classical music)							
	s_1	s_2	s_3	s_4	s_5	s_6	s_7	s_8
y_1	0.096	**0.993**	0.001	0.001	0.019	0.030	0.035	0.043
y_2	0.088	0.032	0.000	0.002	0.000	**0.991**	0.038	0.086
y_3	0.006	0.070	0.005	0.001	0.016	0.094	0.085	**0.990**
y_4	0.028	0.044	0.018	0.002	0.068	0.034	**0.991**	0.093
y_5	0.005	0.007	**0.998**	0.066	0.006	0.009	0.021	0.010
y_6	**0.995**	0.083	0.001	0.008	0.007	0.055	0.018	0.007
y_7	0.000	0.001	0.075	**0.997**	0.002	0.002	0.000	0.000
y_8	0.008	0.027	0.022	0.012	**0.995**	0.006	0.098	0.019

3.3 Kernelized Temporal Predictability Method

The linear blind source separation problem has been solved by the temporal predictability method by Stone's algorithm (Stone, 2001). The nonlinear blind source separation has been solved by applying the kernel method in support vector machines (SVMs) to Stone's algorithm (Martinez & Bray, 2003). The nonlinear mixtures of signals have been separated by using the kernelized temporal predictability method.

The linear blind source separation solution is based on the idea of the independency of source signals. But the nonlinear transform of independent source signals forms independent signals again, without any constraints the original source signals can not be separated. That's why the linear blind source separation methods can not separate the nonlinear mixtures of source signals. In (Martinez & Bray, 2003), Stone's algorithm has been extended for nonlinear blind source separation.

The key points of kernel method have been given in support vector machines in section 2.4. The input signals have been mapped onto a higher dimension by applying nonlinear mapping but working this high dimension data causes huge dimension problems. This problem has been prevented by working in the feature space and as will be seen the method has allowed to linear blind source separation in feature space.

In section 3.2, the long-term and short-term covariance matrices have been represented as:

$$\begin{aligned} \bar{C}_x &= \frac{1}{l}\sum_i \bar{\mathrm{x}}_{\mathrm{i}}\bar{\mathrm{x}}_{\mathrm{i}}^T \\ \tilde{C}_x &= \frac{1}{l}\sum_i \tilde{\mathrm{x}}_{\mathrm{i}}\tilde{\mathrm{x}}_{\mathrm{i}}^T \end{aligned} \tag{3.7}$$

The long-term and short-term running averages that have been expressed in long-term and short-term covariance matrice equations have been defined as below:

$$\begin{aligned} \bar{\mathrm{x}}_i &= \mathrm{x}_i - \frac{1}{l}\sum_{k=1}^{l} \mathrm{x}_k \\ \tilde{\mathrm{x}}_i &= \mathrm{x}_i - \frac{1}{2m+1}\sum_{k=i-m}^{i+m} \mathrm{x}_k \end{aligned} \tag{3.8}$$

By applying the nonlinear transformation as depicted in Figure 3.1, the kernelized temporal predictability method has been formed (Martinez & Bray, 2003).

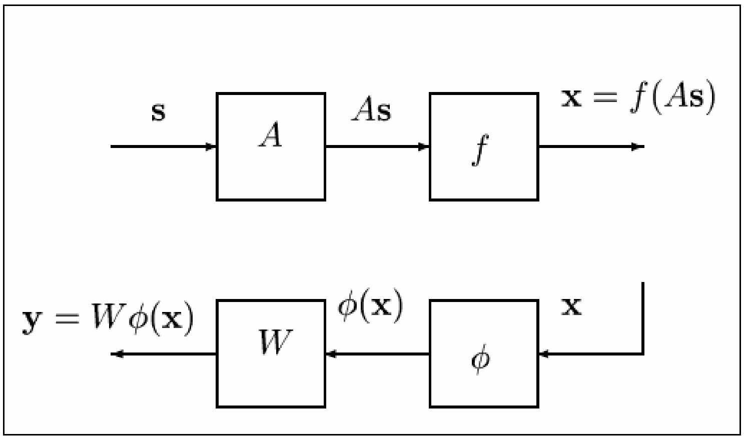

Figure 3.1 The kernelized temporal predictability method model

The upper path is the mixing process chain, the original source signal has been transformed by a nonlinear function f, x=f(As). The bottom path is the unmixing model of the algorithm. If $\varphi = f^{-1}$ and $W = A^{-1}$, then the sources are able to be separated, successfully.

The covariance matrices that are evaluated in the algorithm have been represented below:

$$\begin{aligned} \bar{C} &= \frac{1}{l}\sum_i \overline{\phi(\mathrm{x}_i)}\ \overline{\phi(\mathrm{x}_i)}^T \\ \tilde{C} &= \frac{1}{l}\sum_i \widetilde{\phi(\mathrm{x}_i)}\ \widetilde{\phi(\mathrm{x}_i)}^T . \end{aligned} \tag{3.9}$$

The terms included as long-term and short-term running averages in the unmixing process have been given as:

$$\begin{aligned} \overline{\phi(\mathrm{x}_i)} &= \phi(\mathrm{x}_i) - \frac{1}{l}\sum_i^l \phi(\mathrm{x}_k) \\ \widetilde{\phi(\mathrm{x}_i)} &= \phi(\mathrm{x}_i) - \frac{1}{2m+1}\sum_{k=i-m}^{i+m} \phi(\mathrm{x}_k). \end{aligned} \tag{3.10}$$

This representation of the data has shown the data in the feature space. The problem has been caused here is the dimensionality of the data. A mapping function with a high degree (i.e. fifth-order polynomial) maps the input data (256-dimension) onto a feature space with 10^{10} or higher dimensions. Use of training data directly instead of mapped data has been offered to cancel this dimensionality problem. The solution w that separates the sources has been expressed in terms of the training data.

$$w = \sum_{i=1}^{l} \alpha_i \phi(\mathrm{x}_i) \tag{3.11}$$

The temporal predictability measuring function have been given 3.12 and putting the solution w into the F function the numerator and the denominator of the F can be expressed in terms of the kernel function.

$$F = \frac{w^T \bar{C}_x w}{w^T \tilde{C}_x w} \tag{3.12}$$

$$\begin{aligned} \mathrm{w}^T \bar{C} \mathrm{w} &= \frac{1}{l} \sum_{i,j} \alpha_i \alpha_j \phi(\mathrm{x}_i)^T \sum_k \overline{\phi(\mathrm{x}_k)} \, \overline{\phi(\mathrm{x}_k)}^T \phi(\mathrm{x}_j) \\ &= \frac{1}{l} \sum_{i,j} \alpha_i \alpha_j \sum_k \phi(\mathrm{x}_i)^T \overline{\phi(\mathrm{x}_k)} \phi(\mathrm{x}_j)^T \overline{\phi(\mathrm{x}_k)} \\ &= \frac{1}{l} \alpha^T \bar{K} \bar{K}^T \alpha. \end{aligned} \tag{3.13}$$

The long term kernel $\bar{\mathrm{K}}$ is a matrix with *lxl* dimension with terms

$$\bar{K}_{ij} = \phi(\mathrm{x}_i)^T \overline{\phi(\mathrm{x}_j)}. \tag{3.14}$$

The denominator of the F function and $\tilde{\mathrm{K}}$ have been expressed as below

$$\mathrm{w}^T \tilde{C} \mathrm{w} = \frac{1}{l} \alpha^T \tilde{K} \tilde{K}^T \alpha. \tag{3.15}$$

The term *l* has been indicated as the number of observations. Support vector machines in terms of the kernel function have been expressed as a dot product form $k(x, y) = \phi(x)^T \phi(y)$. In the same way a kernel matrix K has been defined with its entries and has been represented in detail as in the following:

$$
\begin{aligned}
\bar{K}_{ij} &= \phi(\mathrm{x}_i)^T \overline{\phi(\mathrm{x}_j)} \\
&= \phi(\mathrm{x}_i)^T \left\{ \phi(\mathrm{x}_j) \; - \frac{1}{l} \sum_k \phi(\mathrm{x}_k) \right\} \\
&= \phi(\mathrm{x}_i)^T \phi(\mathrm{x}_j) - \frac{1}{l} \sum_k \phi(\mathrm{x}_i)^T \phi(\mathrm{x}_k) \\
&= K_{ij} - \frac{1}{l} \sum_{k=1}^{l} K_{ik}.
\end{aligned} \tag{3.16}
$$

In the same approach the kernel matrix for short-term has been given as:

$$
\tilde{K}_{ij} = K_{ij} - \frac{1}{2m+1} \sum_{k=i-m}^{i+m} K_{ik} \,. \tag{3.17}
$$

The generalized eigenvalue problem that has been expressed in terms of the kernel matrices is $\bar{K}\bar{K}^T \alpha = \lambda \tilde{K}\tilde{K}^T \alpha$. The parameter of λ is the corresponding largest eigenvalue. For recovering a source signal, the linear projection $y = w^T \phi(x)$ has been evaluated. The recovering equation has been expressed in terms of the defined term $w = \sum_{i=1}^{l} \alpha_i \phi(\mathrm{x}_i)$ as $y = \sum_{i=1}^{l} \alpha_i k(x_i, x)$.

If the generalized eigenproblem has been carried out on the entire training dataset, then the size of the kernel matrices would have been *l*x*l* and the solution have depended on the entire dataset. The reduced size of the kernel matrices has been enabled by using p number of training data. Using p number of training data that has been participated the solution has been indicated as support vectors in SVM method. Then the solution has been expressed as $y = \sum_{i \in SV}^{l} \alpha_i k(x_i, x)$. The SV indicates the set of support vectors (Martinez & Bray, 2003).

Using the set of support vectors has constrained the solution in the subspace that has been spanned by the support vectors. If the entire training data has the statistical redundancy, the constrained subspace solution will facilitate the solution way.

The kernelization of the temporal predictability method has been given with equations in this section. The kernelized temporal predictability method has been constituted by the following steps:

1. A subset of p support vectors has been found by the feature selection algorithm for the selected kernel k(x,y).
2. Kernel matrices for long-term and short-term with size of *p*x*l* have been computed, between the support vectors and the entire training data
3. The generalized eigenvalue problem $\bar{K}\bar{K}^T\alpha = \lambda\tilde{K}\tilde{K}^T\alpha$ has been solved.
4. The sources have been recovered from the eigenvectors solution α.

.

CHAPTER FOUR
SEPARATION OF SOUND MIXTURES

4.1 Introduction

As stated in the previous chapter, the temporal predictability algorithm (Stone, 2001) can be used to recover the linearly mixed source signals. However this algorithm cannot separate the nonlinear mixtures which can be caused by the recording devices or the environmental properties. In (Martinez & Bray, 2003) a new algorithm which uses the kernel approach has been proposed to recover the nonlinear mixtures. In this chapter, both temporal predictability and kernelized temporal predictability algorithms have been applied to several different types of sound mixtures.

The kernelized temporal predictability method which has been expressed briefly in section 3.3 has been implemented to recover the source signals from nonlinear mixture signals. This approach has extended the recovering capability for different type of signals. The parameters of kernel type, kernel degree, number of feature vectors and their effect on the performance parameters that are mean and standard deviation of correlation coefficients have been simulated and the results have been observed.

In order to test algorithms different types of sound signals have been used: First signals are TIMIT speech signals that are taken randomly from the TIMIT database. The second type of the sound signals are the musical notes that were taken from the IOWA database that has been rearranged by M. Erdal Ozbek in Inst. Supelec. and the third type of the sound signals are recorded by from daily song signals.

4.2 Separation of Data Mixtures using FastICA Method

FastICA method that has been explained in section 2.3.1 has been simulated to recover the source signals of different data types in this section. Three sets of sound data that are speech, musical notes and recorded song signals have been used as source signals.

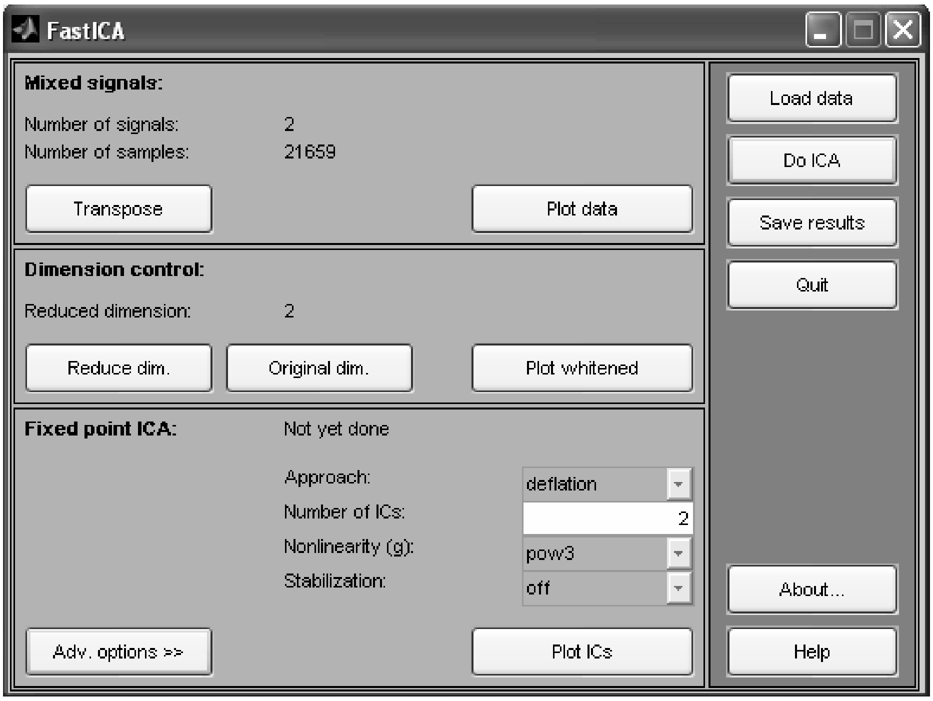

Figure 4.2 FastICA graphic interface in Matlab

FastICA graphic user interface has been represented in Figure 4.1. The FastICA parameters that the data loading, dimension control, nonlinearity function and numbers of independent components are used to adjust the FastICA algorithm. They can be adjusted by clicking on them. There are several choices for parameters to do the ICA, but only one of them has been used for each data type in this section.

4.2.1 Separation of Sound Signals using FastICA

Two speech signals from TIMIT database that are s_1= dr1_fetbo_sa2 and s_2= dr2_mcewo_sx182, the two musical notes from IOWA database and the two recorded song signals have been used as source signals. The linear and nonlinear mixtures of them have been used as mixture signals in FastICA algorithm. The linear mixing has been formed by a random mixing matrix created by randn in Matlab and the nonlinear mixing has been defined as square root function in Matlab.

The linear mixture of speech signals have been separated by using FastICA with settings seen in the previous figure. The figure of the source and the recovered signals has been represented in the next figure.

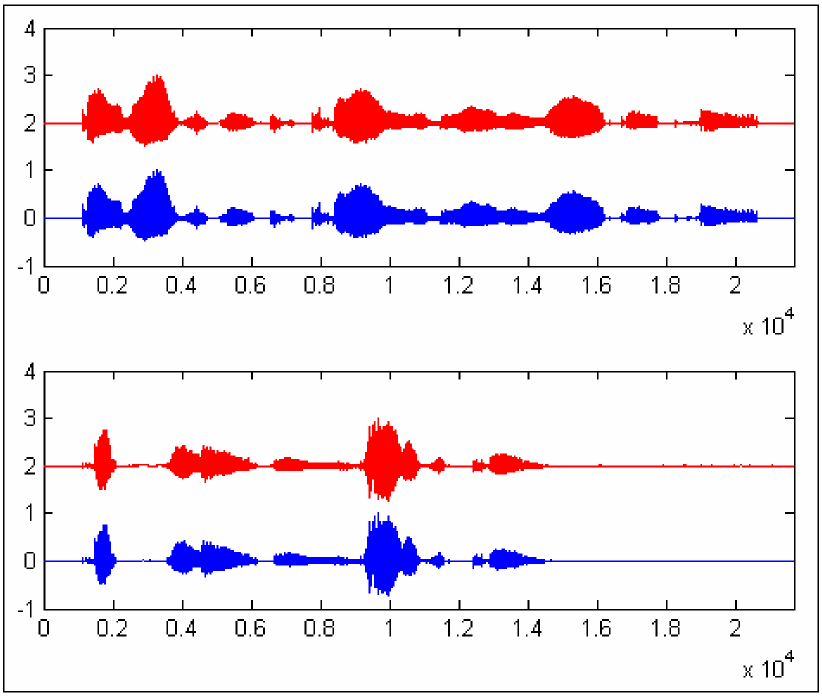

Figure 4.3 The source (bottom) and recovered (top) signals

The linear mixture of speech signals have been separated successfully as seen in the figure above. The recovered speech signals have been represented at top trace and the source signals at bottom trace.

The correlation coefficients between the source and the recovered signals for different sound data types have been given in the following table.

FastICA algorithm mentioned in section 2.3.1 has not been able to separate the nonlinear mixtures of sound data types as seen from the corresponding correlation coefficients in Table 4.1. It has been able to separate the linear mixtures of sound data types with correlation coefficients almost equal to unity.

Table 4.1The correlation coefficients between original source signals and recovered signals of FastICA simulations for nonlinear mixing type

Sound Data Type	Mixing type	FastICA Nonlinearity	Recovered Signals	Correlation Coefficients s_1	s_2
Speech signals	Nonlinear	Pow3	y_1	0.0170	0.9998
			y_2	0.9999	0.0107
	Nonlinear	Tanh	y_1	0.0012	0.0054
			y_2	0.0063	0.0022
	Nonlinear	Gauss	y_1	0.0012	0.0054
			y_2	0.0063	0.0022
	Nonlinear	Skew	y_1	0.0013	0.0054
			y_2	0.0063	0.0022
Musical note signals	Nonlinear	Pow3	y_1	0.0009	0.0001
			y_2	0.0055	0.0004
	Nonlinear	Tanh	y_1	0.0009	0.0001
			y_2	0.0055	0.0004
	Nonlinear	Gauss	y_1	0.0009	0.0001
			y_2	0.0055	0.0004
	Nonlinear	Skew	y_1	0.0001	0.0009
			y_2	0.0013	0.0053
Recorded song signals	Nonlinear	Pow3	y_1	0.0003	0.0015
			y_2	0.0001	0.0008
	Nonlinear	Tanh	y_1	0.0003	0.0015
			y_2	0.0001	0.0008
	Nonlinear	Gauss	y_1	0.0003	0.0015
			y_2	0.0001	0.0008
	Nonlinear	Skew	y_1	0.0015	0.0003
			y_2	0.0001	0.0008

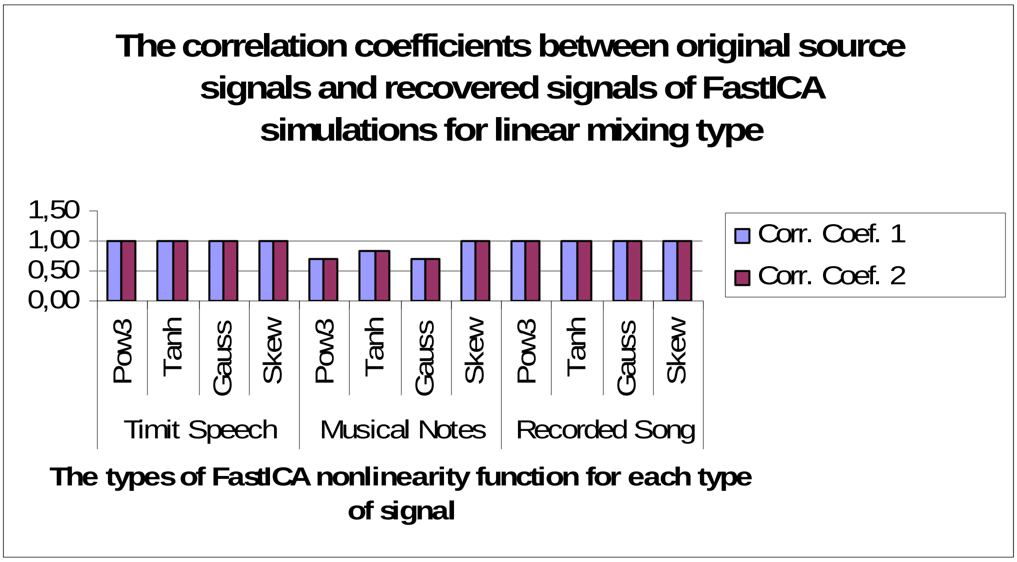

Figure 4.8 The correlation coefficients between original source signals and recovered signals of FastICA simulations for linear mixing type

4.2.2 Separation of Images using FastICA

FastICA algorithm has been used to separate the linear and nonlinear mixtures of image sources in this section. Five sets of different images have been used as source images. The image sources have been converted to row vector from matrix form to apply the algorithm. The type of the source data, the mixing type of the source data, FastICA nonlinearity and the corresponding correlation coefficients have been given in the following table.

The sizes of the images in Table 4.2 are 250x450 and 300x400 for motorcycle-plane and house-bridge, respectively. The three sets of images in Figure 4.5 are face-writing, face-object and face-scribble images. They have much smaller sizes than the images in Figure 4.4. The sizes of the face-writing, face-object and face-scribble images are 100x90, 96x92 and 100x90, respectively.

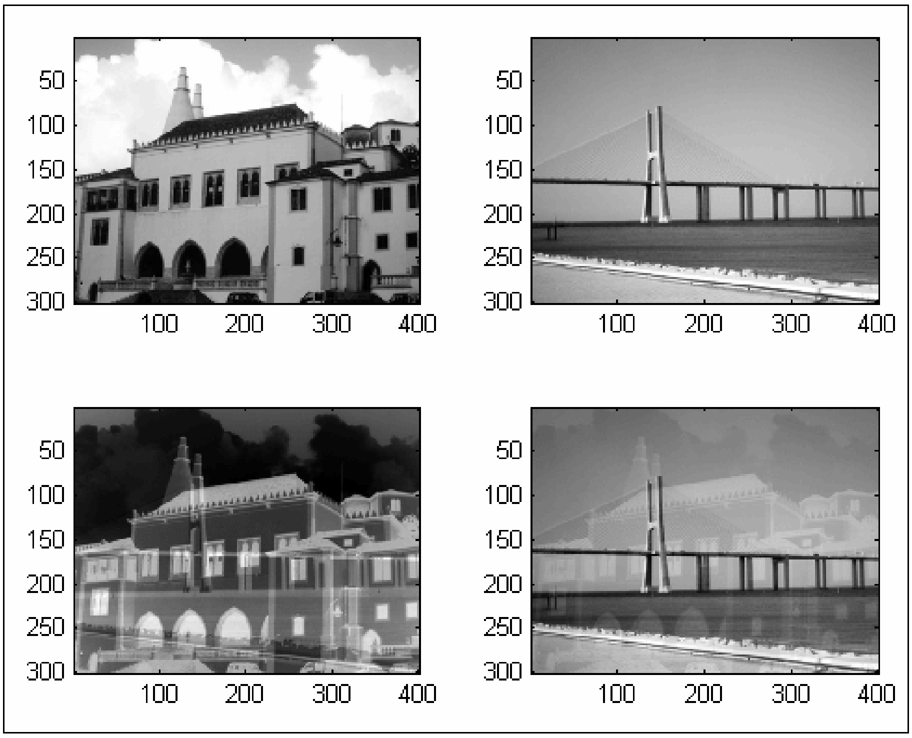

Figure 4.4 The source and recovered house-bridge images

The source and recovered house-bridge images have been represented in Figure 4.3. The FastICA has been simulated for the tanh nonlinearity and the correlation coefficients for this simulation are 0.9522 and 0.9930. The correlation coefficients for all FastICA nonlinearity have been given in the performance tables.

The FastICA algorithm has been applied for the linear mixtures of house-bridge and motorcycle-plane images. In section 4.3 the linear mixtures of these images have been used to simulate the temporal predictability method that is used to separate the linear mixtures of signals, so that a comparison can be made between FastICA and temporal predictability methods. For the same reason the linear and nonlinear mixtures of face-writing, face-object and face-scribble images have been used to simulate FastICA algorithm. The nonlinear mixing has been created as square root function in Matlab.

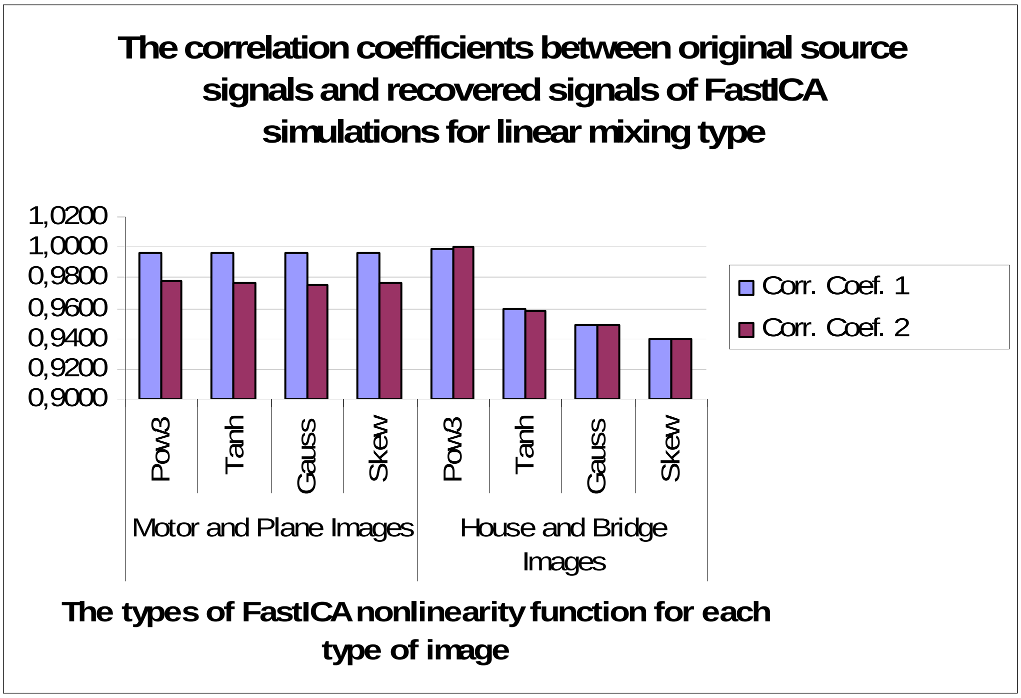

Figure 4.5 The correlation coefficients between original source images and recovered images for FastICA simulations

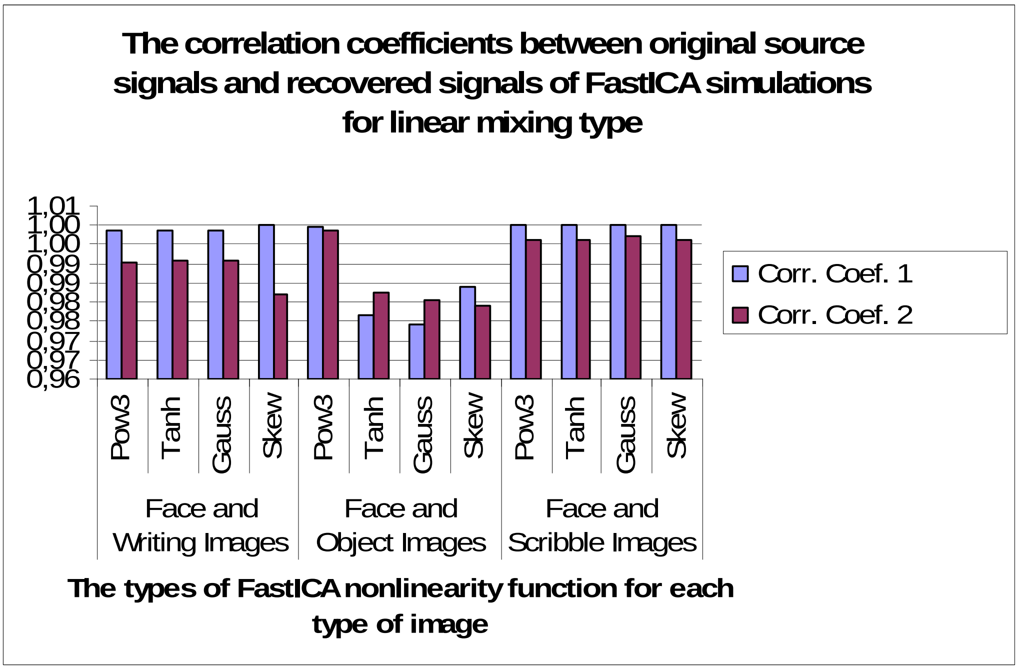

Figure 4.6 The correlation coefficients between original source images and recovered images for linear mixing type FastICA simulations

The correlation coefficients in the previous figure have represented the performance of FastICA for its different nonlinearity functions detailed in section (2.3.1). The values have been almost good for all nonlinearity functions but the type of the images has affected the performance. For house-bridge image set the pow3 nonlinearity function has got the best results but there has not been much difference between correlation coefficients of the motorcycle-plane image set for different nonlinearity functions.

Table 4.2 The correlation coefficients between original source images and recovered images for FastICA simulations

Data Type	Mixing type	FastICA Nonlinearity	Recovered Signals	Correlation Coefficients s_1	s_2
Face and Writing Images	Nonlinear	Pow3	y_1	0.0143	0.0041
			y_2	0.0255	0.0136
	Nonlinear	Tanh	y_1	0.0040	0.0143
			y_2	0.0136	0.0256
	Nonlinear	Gauss	y_1	0.0143	0.0041
			y_2	0.0255	0.0136
	Nonlinear	Skew	y_1	0.0001	0.0148
			y_2	0.0058	0.0284
Face and Object Images	Nonlinear	Pow3	y_1	0.0071	0.0069
			y_2	0.0043	0.0023
	Nonlinear	Tanh	y_1	0.0070	0.0070
			y_2	0.0023	0.0043
	Nonlinear	Gauss	y_1	0.0070	0.0070
			y_2	0.0023	0.0043
	Nonlinear	Skew	y_1	0.0065	0.0075
			y_2	0.0045	0.0019
Face and Scribble Images	Nonlinear	Pow3	y_1	0.0093	0.0126
			y_2	0.0101	0.0097
	Nonlinear	Tanh	y_1	0.0127	0.0092
			y_2	0.0098	0.0100
	Nonlinear	Gauss	y_1	0.0093	0.0126
			y_2	0.0101	0.0097
	Nonlinear	Skew	y_1	0.0029	0.0154
			y_2	0.0003	0.0140

The linear mixtures of images have been separated by FastICA successfully as seen from the correlation coefficients in Figure 4.4. The nonlinearity function of FastICA algorithm has affected the correlation coefficients a little.

4.3 Separation of Sound Data Mixtures using Temporal Predictability

The first applications include the separation of linear sound mixtures using the temporal predictability method and the next applications are nonlinear blind source separation of the same data mixtures. Three sets of sound data types have been simulated in linear and kernel blind source separation approach. They are speech signals from TIMIT database, musical note signals from IOWA database and recorded song signals.

4.3.1 Separation of Linear Mixture of Two Speech Signals using Temporal Predictability

Two speech signals from TIMIT database (a female and a male speech sentence) s_1= dr1_fetbo_sa2 and s_2= dr2_mcewo_sx182 have been used as source signals. The s_1 contains the sentence of "Don't ask me to carry an oily rag like that" and the s_2 contains the sentence of "Thick glue oozed out of the tube". The size of the s_1 signal is 21659x1 and the s_2 is a 22887x1 vector. A compromise value of 21659x1 has been used for both source signals and this value is sufficient to get the whole words in the speech signals. They have been mixed by a random mixing matrix **A** that has been created by randn function in Matlab. The temporal predictability algorithm that has been created by J.V. Stone (Stone, 2001) used as basis program and the new programs have been implemented to simulate the new data types and observe the performance in a general approach.

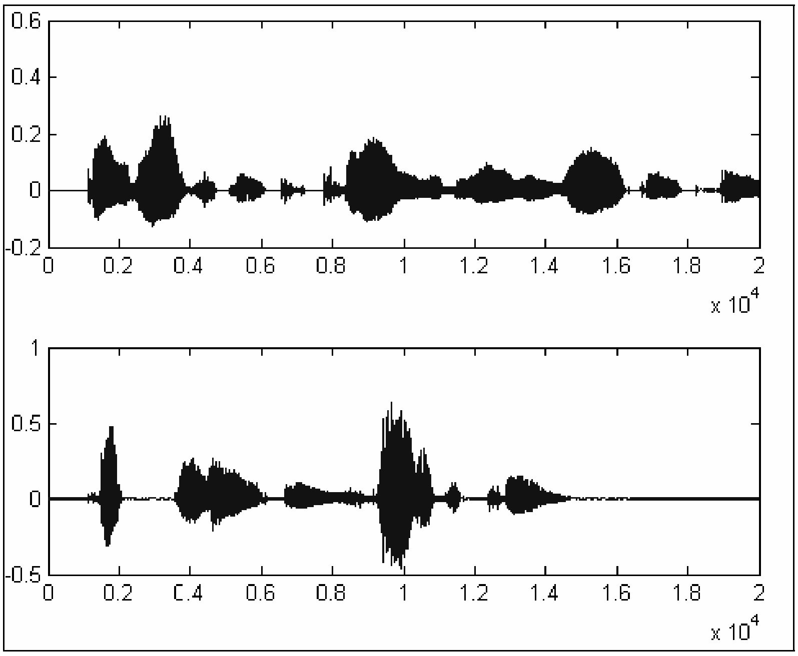

Figure 4.7 Speech Source Signals

The figure of the source signals that have been represented in the figure above and the linear mixtures of speech signals have been shown in the following figure.

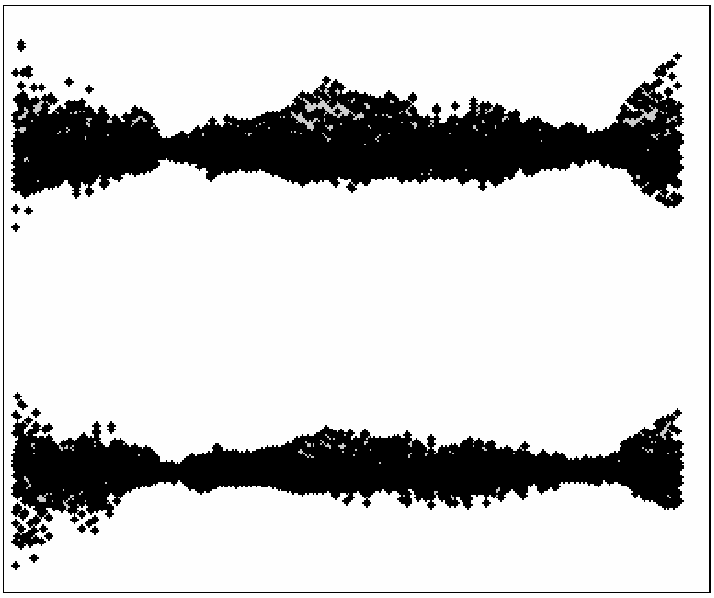

Figure 4.8 Mixtures of Speech Signals

The linear mixture signals in the figure above have been obtained by **x**=As transformation, and the original source signals have been recovered using the temporal predictability algorithm. In the following figure, the recovered signals from the linear mixture signals that are at the top and the source signals have been represented at the bottom. As seen from the figure, the algorithm has been able to recover the source signals successfully.

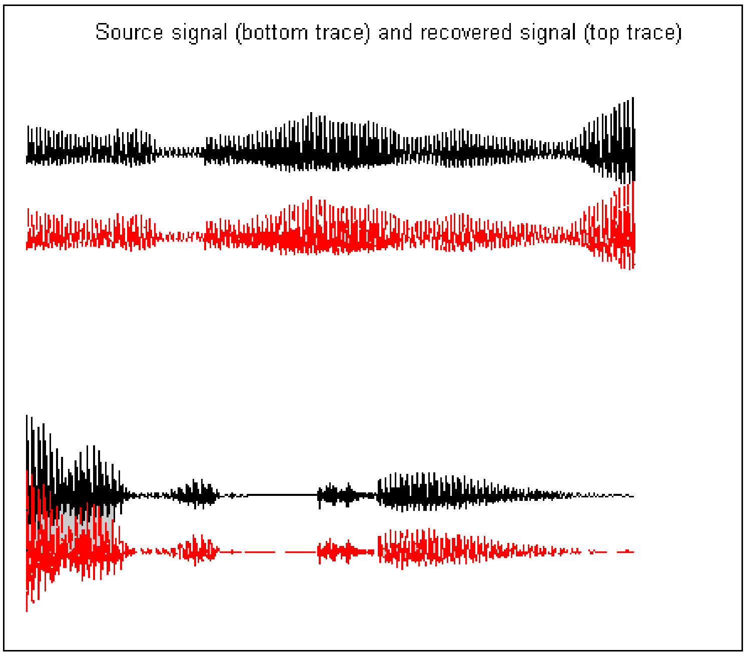

Figure 4.9 Recovered (top trace) and source signal (bottom trace)

The signal recovering program has been simulated for 1000 times to get the statistical parameters that are mean and standard deviation of the correlation coefficient matrix. The results are as follows:

Table 4.3 The mean of correlation coefficients between original source signals and recovered signals

Recovered Signals	Source Signals	
	s_1	s_2
y_1	0.9998	0.0189
y_2	0.0127	0.9999

The mean of correlation coefficient matrix in the table above has proven the success of the algorithm, because the correlation coefficients between the source and the recovered signals have values adequately approximate to unity (0.9999 and 0.9998). The maximum correlation coefficient value in the first row is the first source signal; this indicates the fact that the first recovered signals is $s_1$and the second recovered signal is s_2.

The standard deviation values of the correlation coefficient matrix in the next figure have very minor values so that it can be said that the algorithm is stable for the separation of source signals from random linear mixtures.

Table 4.4 The Standard deviation of correlation coefficients between original source signals and recovered signals

Recovered Signals	Source Signals s_1	s_2
y_1	$0.1144\ 10^{-13}$	0.001910^{-13}
y_2	$0.0025\ 10^{-13}$	0.015610^{-13}

The power spectrums of the source, mixture and recovered signals have been represented in the following figure. The two signals at the first row are source signals, the mixture signals are at the second row and the third row has the recovered signals.

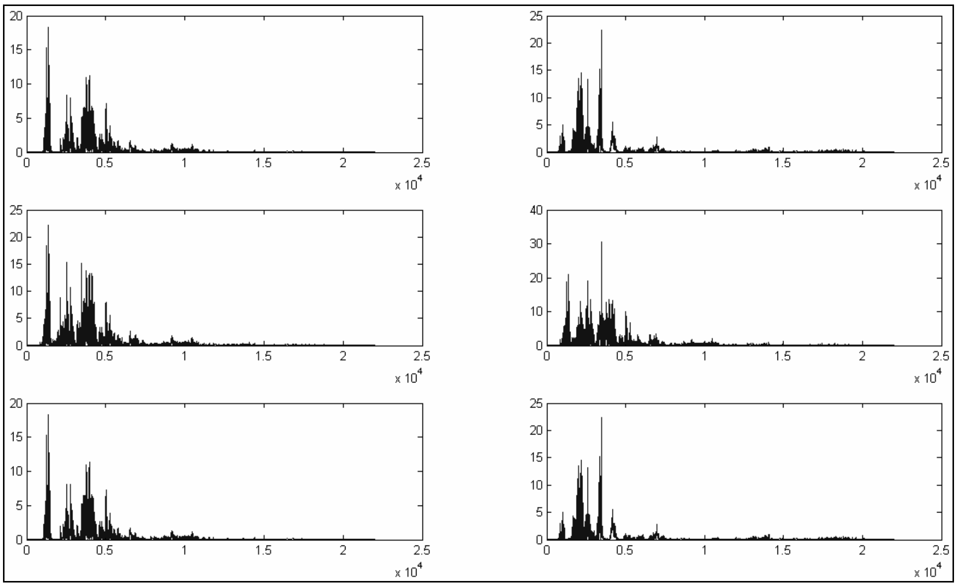

Figure 4.10 Power spectrums of spectrums of the source, mixture and recovered signals

The power spectrum of the source signals have not changed after the linear mixing as seen in above. The recovered signals at the third row have the same power spectrum with the source signals at the first row.

4.3.2 Separation of Three Musical Notes Using Temporal Predictability

Three musical notes from IOWA database, s_1= AltoSax.NoVib.ff.A4, s_2= BbClar.ff.A4 and s_3= SopSax.NoVib.ff.A4 have been used as source signals. They have been mixed by a random mixing matrix **A** that has been created by randn function in Matlab. The source signals have been recovered by the implemented programs based on basis temporal predictability programs. Statistical parameters of the simulations have been given to represent the performance of the algorithm.

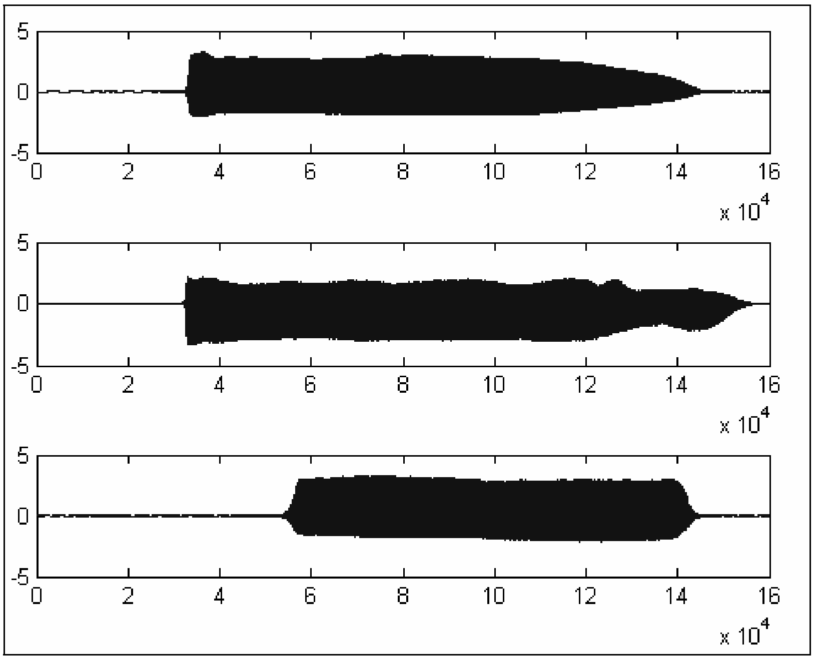

Figure 4.11 Source signals taken from the Iowa database

The source signals as depicted in the figure above have been mixed by the random mixing matrix **A** and the figure of mixture signals has been represented in the next figure.

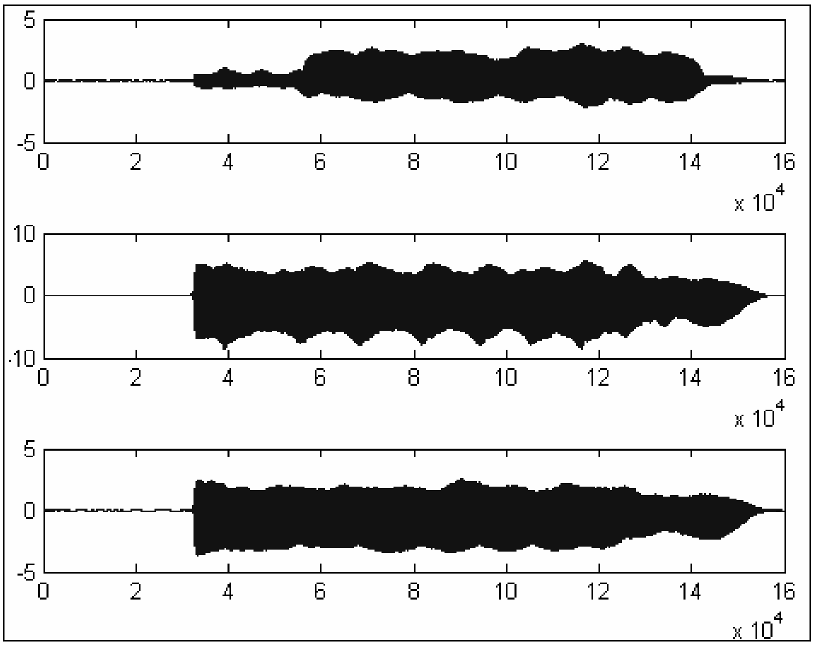

Figure 4.12 Mixture signals of musical notes taken from Iowa database

The source signals have been recovered by the temporal predictability algorithm based programs in the figure above. The algorithm has been simulated for one thousand times to observe the performance parameters of mean and standard deviation of correlation coefficients. The statistical parameters of the simulation in the following tables have given the performance stability of the algorithm.

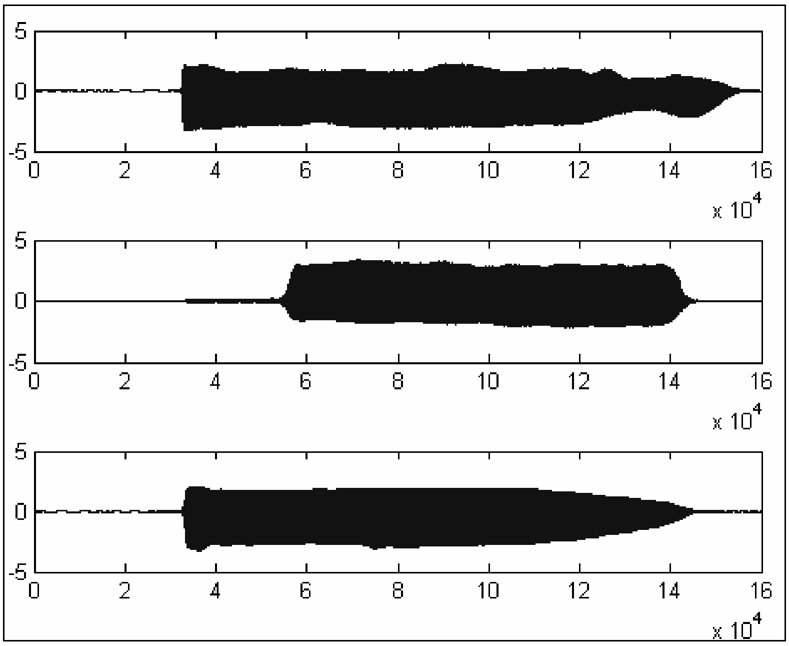

Figure 4.13 Recovered signals from the linear mixture signals

Table 4.4 The mean of correlation coefficients between original source signals and recovered signals

Recovered Signals	Source Signals		
	s_1	s_2	s_3
y_1	0.0004	0.0124	0.9999
y_2	0.9953	0.0966	0.0047
y_3	0.0022	0.9991	0.0414

Table 4.5 The Standard deviation of correlation coefficients between original source signals and recovered signals

Recovered Signals	Source Signals		
	s_1	s_2	s_3
y_1	$0.0000\ 10^{-14}$	$0.0021\ 10^{-14}$	$0.1336\ 10^{-14}$
y_2	$0.3562\ 10^{-14}$	$0.0056\ 10^{-14}$	$0.0004\ 10^{-14}$
y_3	$0.0002\ 10^{-14}$	$0.4675\ 10^{-14}$	$0.0139\ 10^{-14}$

The mean of correlation coefficient matrix in Table 4.5 has given the order of source signals within the recovered signals. For instance; the first row of the matrix in the table has its maximum value for s_3 source signal; it means the first recovered signal has corresponded to the third source signal s_3. The maximum value in each row of the correlation coefficient matrix indicates the recovered source signal identity. This source identification is the same for all simulation examples.

The mean values of correlation matrix in the indicated table have proven that the algorithm have been able to recover the all three source signals from linear mixtures successfully. The correlation values between the source and the recovered signals are adequate that are 0.99 for three signals. The standard deviation values of correlation coefficient matrix in Table 4.6 are very minor and this has proven temporal predictability algorithm's performance stability for recovering musical note signals from linear mixtures of these signals.

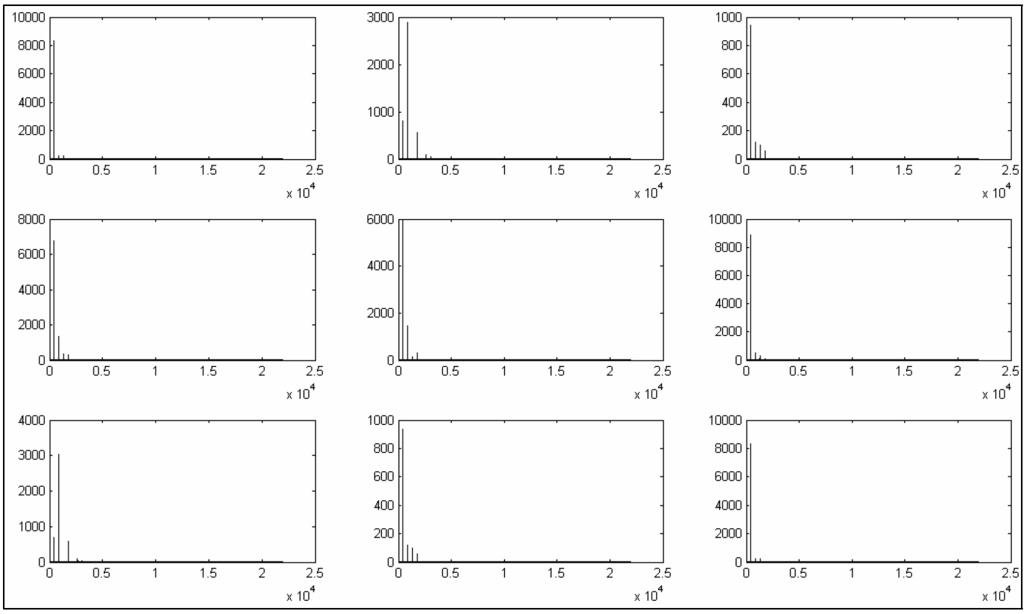

Figure 4.14 Power spectrums of spectrums of the source, mixture and recovered signals

The power spectrums of the source, mixture and recovered signals have been represented in the first, second and third rows in Figure 4.14, respectively. The power spectrums of the recovered signals have identified the order of source signals from the recovered signals. As seen the first recovered signals is the s_2, the second recovered signals is s_3 and the third one is s_1. This fact has been confirmed by the correlation matrix of this simulation as given below:

Table 4.6 The mean of correlation coefficients between original source signals and recovered signals

Recovered Signals	Source Signals		
	s_1	s_2	s_3
y_1	0.0004	0.0124	0.9999
y_2	0.9953	0.0966	0.0047
y_3	0.0022	0.9991	0.0414

4.3.3 Separation of Three Recorded Song Signals Using Temporal Predictability

In this simulation example, three recorded song signals for one minute have been used as source signals and the signals have been linearly transformed by using a random mixing matrix **A** in Matlab. The temporal predictability based programs have been used to recover the source signals from the linear mixture signals. The simulation has been simulated for one thousand times to observe the mean and standard deviation of correlation coefficients matrix.

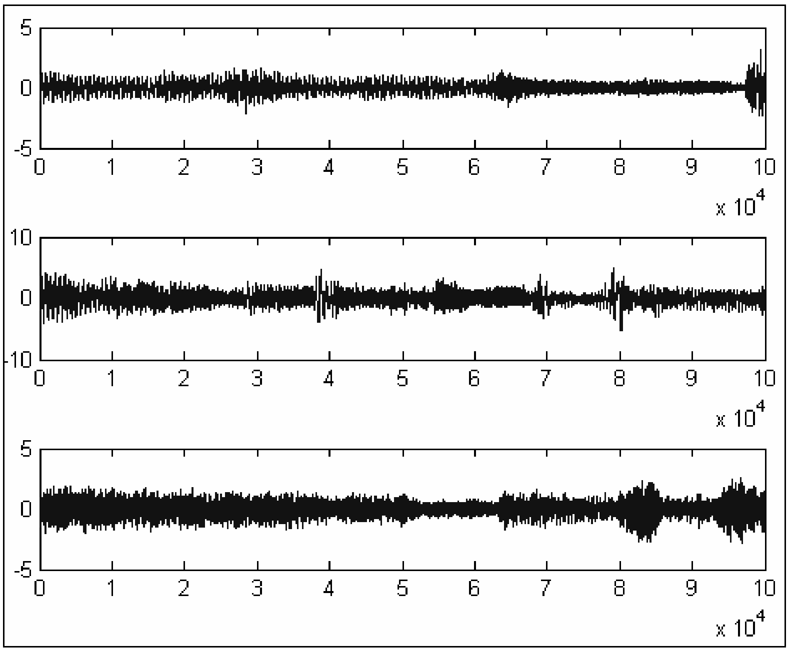

Figure 4.15 Source signals recorded from songs for one minute

The source signals of three recorded song signals have been represented in the previous figure. These signals have a structure that is different than the previous two types of source signals that are speech and musical note signals. The three recorded song signals that have been sampled at 44100 Hz have spread on overall range while the speech and musical note signals have spread some parts of the their overall range. Thus, the performance analysis for a spread signal has been implemented. (So that different data type's performance has been able to be tested for the temporal predictability method.)

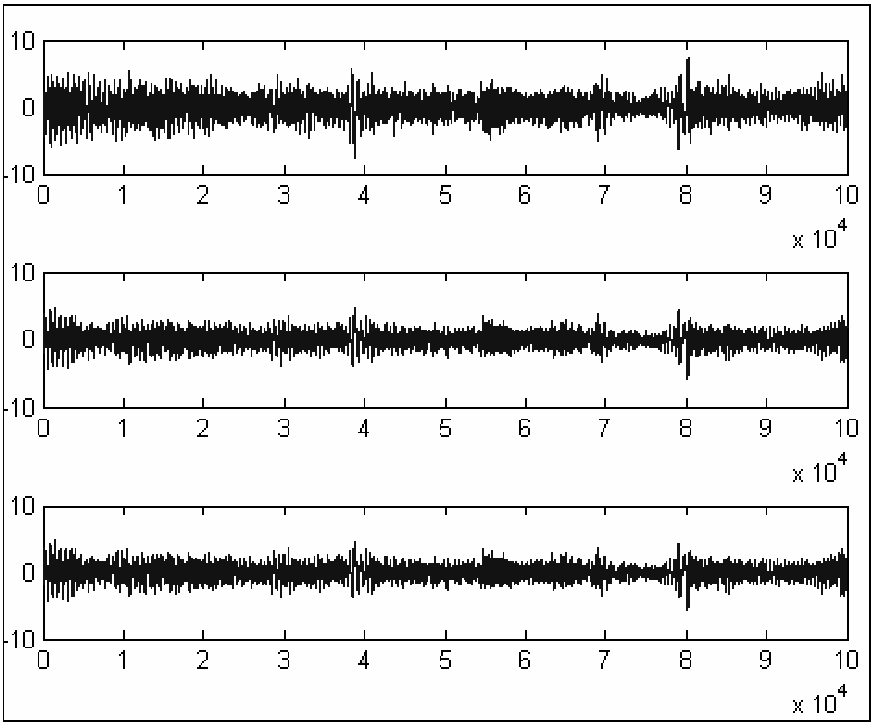

Figure 4.16 Mixture signals of three recorded song signals

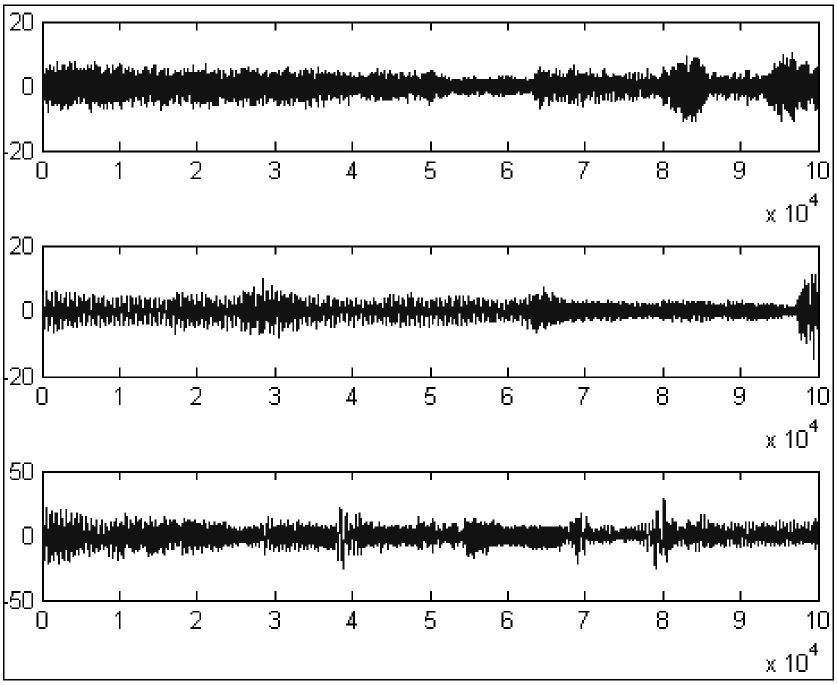

Figure 4.17 Recovered three song signals from linear mixture signals

As seen in Figure 4.17 the source signals have been recovered successfully. The mean of the correlation coefficient matrix in Table 4.8 have been given. As seen the first row maximum value is at s_2, it indicates the first recovered signal has corresponded to the second source signal. The second and third recovered signals have been third and first source signals, respectively.

Table 4.7 The mean of ccorrelation coefficients between original source signals and recovered signals

Recovered Signals	Source Signals		
	s_1	s_2	s_3
y_1	0.0119	0.9996	0.0264
y_2	0.0059	0.0215	0.9998
y_3	0.9997	0.0208	0.0100

Table 4.8 The Standard deviation of ccorrelation coefficients between original source signals and recovered signals

Recovered Signals	Source Signals		
	s_1	s_2	s_3
y_1	$0.0016\ 10^{-13}$	$0.1400\ 10^{-13}$	$0.0019\ 10^{-13}$
y_2	$0.009\ 10^{-13}$	$0.0045\ 10^{-13}$	$0.0944\ 10^{-13}$
y_3	$0.1622\ 10^{-13}$	$0.0004\ 10^{-13}$	$0.0016\ 10^{-13}$

The correlation coefficients for three recovered signals are adequately approximate to unity and the standard deviation values resulted from the simulations in Table 4.9 have been very minor. These results have proven the good performance and stability of the temporal predictability algorithm for recovering source signals from linear mixtures.

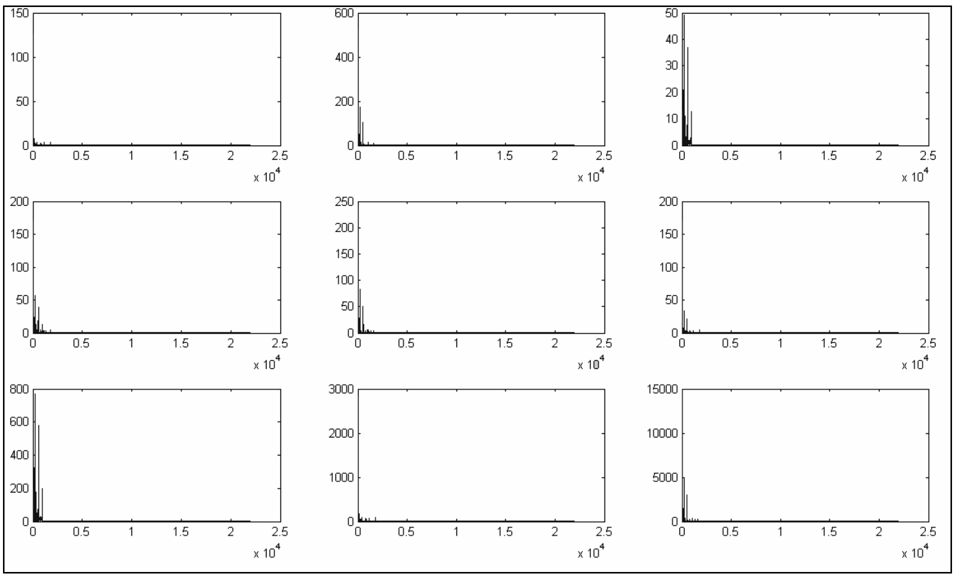

Figure 4.18 Power spectrums of spectrums of the source, mixture and recovered signals

The power spectrums of the source, mixture and recovered signals have been represented in the first, second and third rows in the figure above, respectively. The power spectrums of the recovered signals have identified the order of source signals from the recovered signals. As seen the first recovered signals is the s_3, the second recovered signals is s_1 and the third one is s_2. This fact has been confirmed by the correlation matrix of this simulation as given in Table 4.10.

Table 4.9 The mean of correlation coefficients between original source signals and recovered signals

Recovered Signals	Source Signals		
	s_1	s_2	s_3
y_1	0.0119	0.9996	0.0264
y_2	0.0059	0.0215	0.9998
y_3	0.9997	0.0208	0.0100

4.4 Separation of Sound Data Mixtures using Kernelized Temporal Predictability

4.4.1 Separation of Nonlinear Mixture of Two Speech Signals using Kernelized Temporal Predictability

Two speech signals from TIMIT database have been used as source signals. Two types of nonlinear mixing that are square root and a feedforward multilayer perceptron functions have been applied to the source signals. The implemented kernelization of the temporal predictability method in section 3.3 has been used to recover the source signals from nonlinear mixtures signals. The results have been observed for different kernel functions that are polynomial and sigmoid with different kernel parameters. The source, mixture and recovered signals have been depicted in the following figures for the polynomial kernel function having degree 2.

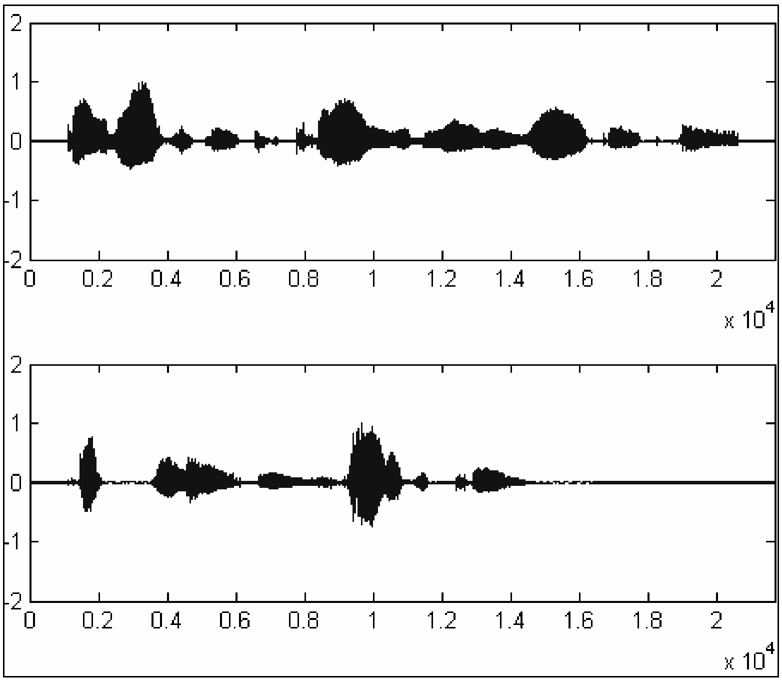

Figure 4.19 Two speech source signals from Timit database

The source signals of two speech signals from Timit database have been represented in the figure above and the nonlinear mixture signals that have been created from square root function have been depicted in the next figure.

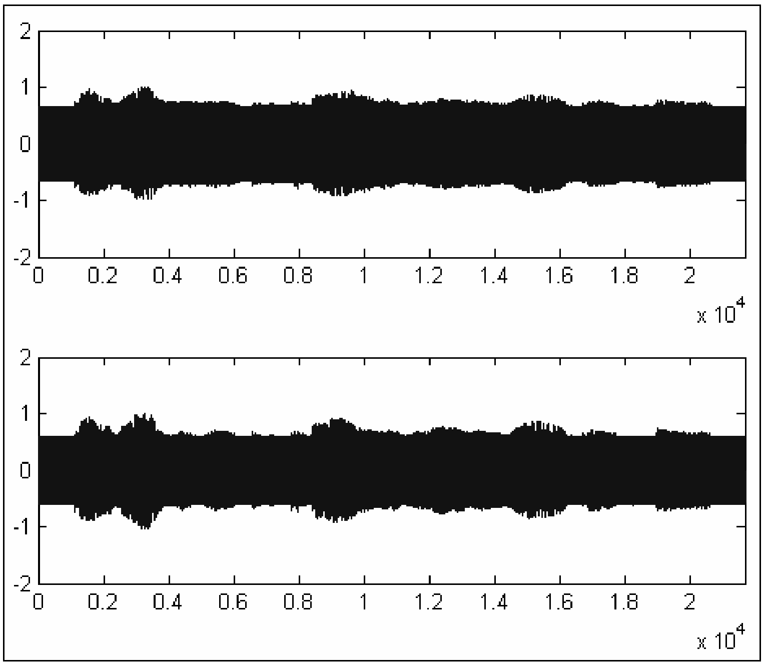

Figure 4.20 The nonlinear mixture signals of the source signals

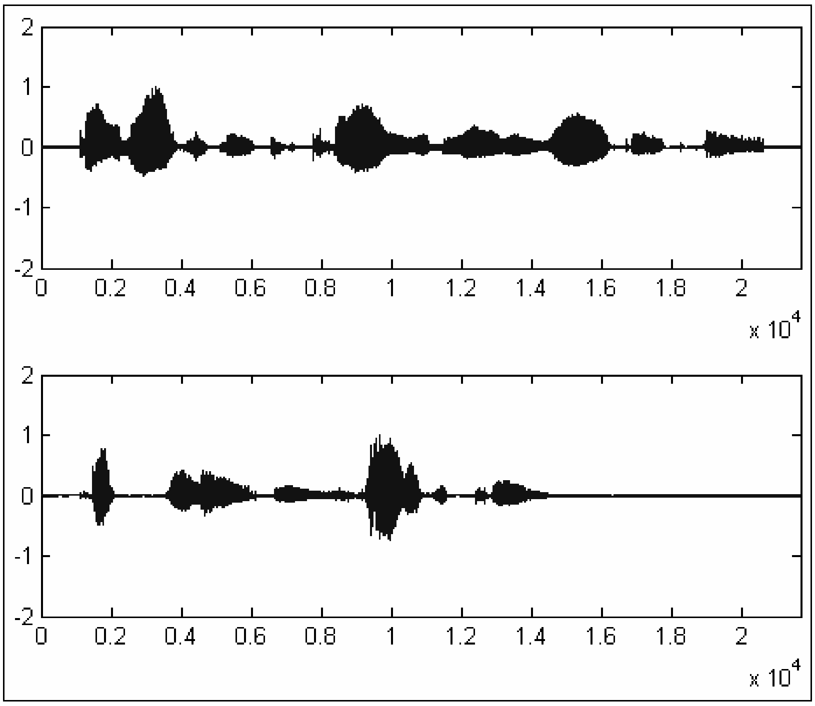

Figure 4.21 The recovered signals from nonlinear mixture signals

The kernelized temporal predictability algorithm has been simulated for the polynomial kernel functions having different degrees. The program has been simulated for one hundred times for each kernel degree to analyze the performance of the kernel degree. As seen from the results of the simulations for each polynomial kernel degree in the table below, the number of feature vectors has been changing depending on the polynomial kernel function's degree. For these simulations, square root nonlinear mixing function has been used.

There is an important point about the number of feature vectors that is the number of the feature vectors can be restricted by the feature vector selection algorithm. An upper limit can be put so that the algorithm extracts the number of feature vectors as defined an upper limit if the number of feature vectors are greater than the limit value.

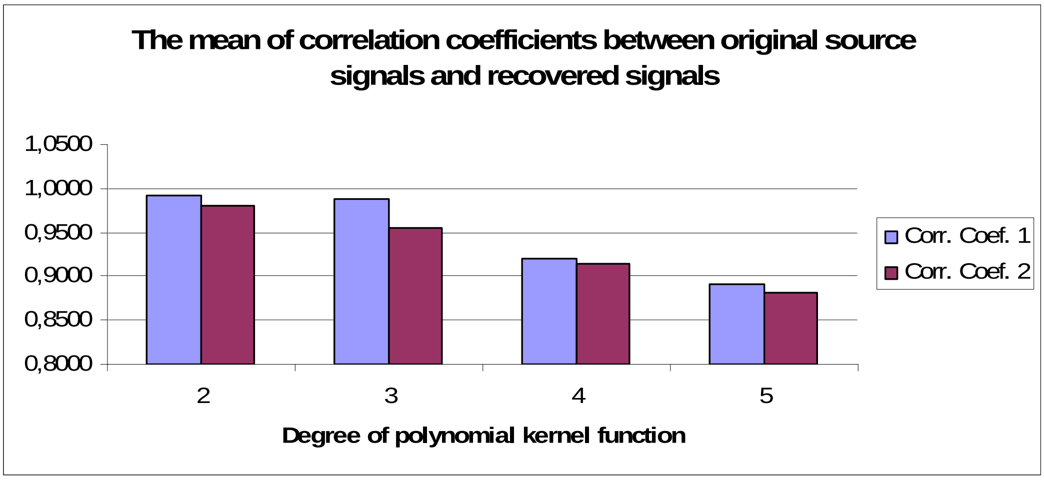

Figure 4.22 The mean of correlation coefficients between original source signals and recovered signals

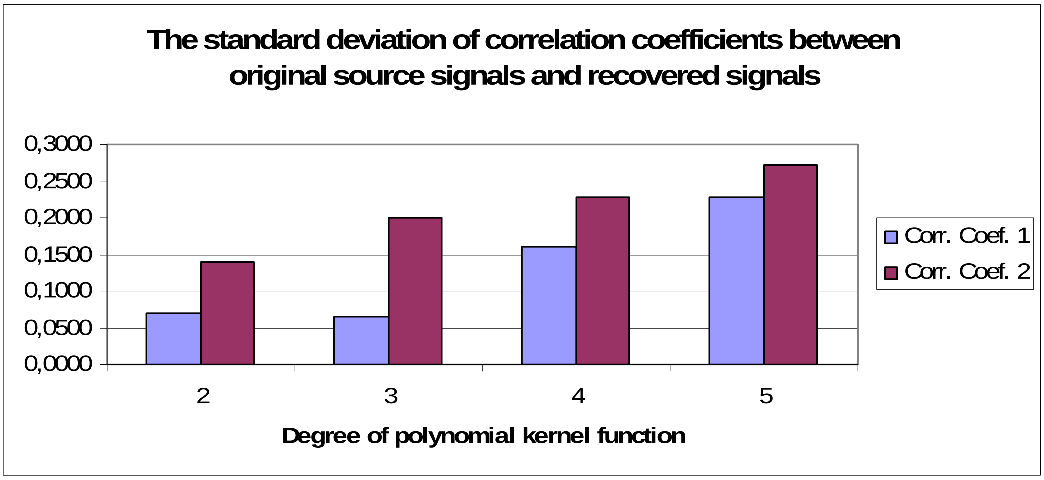

Figure 4.23 The Standard deviation of correlation coefficients between original source signals and recovered signals

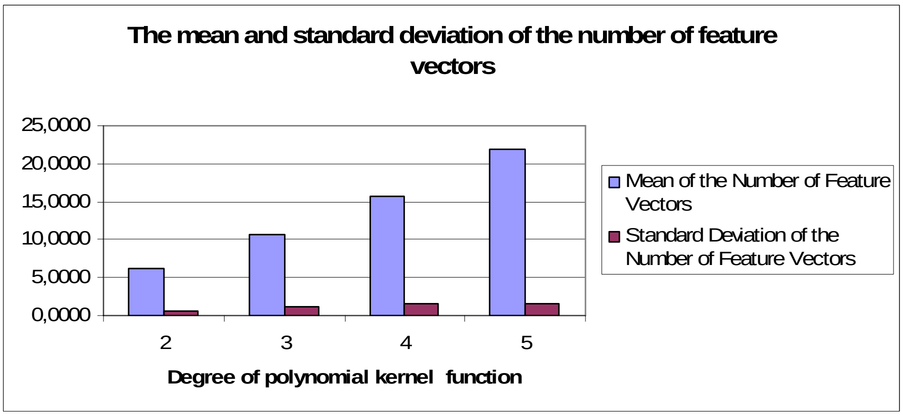

Figure 4.24 The mean and standard deviation of number of the feature vectors

The value of the performance parameters are related with the degree of polynomial kernel and the number of feature vectors. The number of the feature vectors has increased depending on the increase of the degree of kernel function. The mean values of the correlation coefficients between the source and the recovered signal have decreased as the polynomial kernel function's degree has increased.

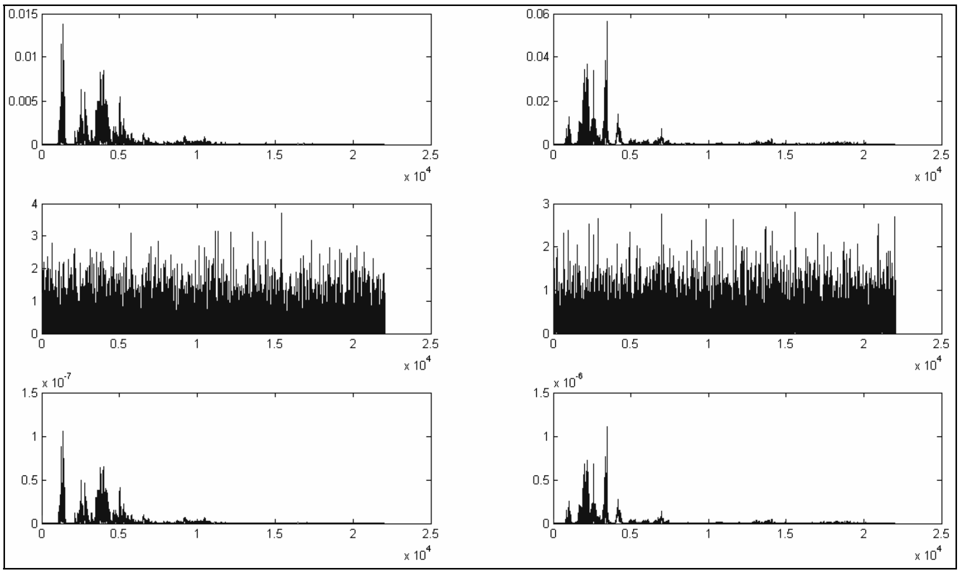

Figure 4.25 Power spectrums of spectrums of the source, mixture and recovered signals

The power spectrums of the source, mixture and recovered signals have been represented in the first, second and third rows in the figure above, respectively. The power spectrums of the nonlinear mixtures of source signals have changed as seen in the third row. The power spectrums of the recovered signals have identified the order of source signals. The order of the source signals have not changed as seen in the figure, the first recovered signal is the s_1 and the second one is s_2. This fact has been confirmed by the correlation coefficients of the simulation that has been used for evaluating power spectrums in Figure 4.25. The correlation coefficients of the simulation have been given in the table below.

Table 4.10 The correlation coefficients between original source signals and recovered signals

Recovered Signals	Source Signals s_1	Source Signals s_2
y_1	0.9999	0.0151
y_2	0.0213	0.9997

The sigmoid kernel function has been simulated for the kernelized temporal algorithm to observe the kernel affect. The p_1 and p_2 are parameters of sigmoid kernel function in (3.3).

Table 4.11 The Standard deviation of correlation coefficients between original source signals and recovered signals for square root nonlinear mixing and sigmoid kernel functions

The square root nonlinear mixing type	Sigmoid Parameters p_1	p_2	Mean of Correlation Coefficients' Matrix		Standard Deviation of Correlation Coefficients' Matrix		Mean of the Number of Feature Vectors	Standard Deviation of the Number of Feature Vectors
	1	2	0.0304	0.0244	0.0361	0.0263	2	0
			0.0373	0.0286	0.0288	0.0262		
	1	1	0.0192	0.0158	0.0173	0.0125	2	0
			0.0246	0.0210	0.0159	0.0140		
	2	5	0.0655	0.0786	0.0610	0.0685	2	0
			0.0758	0.1291	0.0506	0.0614		
	5	10	0.0893	0.0711	0.0941	0.0898	2	0
			0.1268	0.1495	0.1218	0.1215		

As seen from the results of the simulation in the table above, the sigmoid kernel function has not been able to separate the source signals. The power spectrums of source, mixture and recovered signals have been given in the following figure.

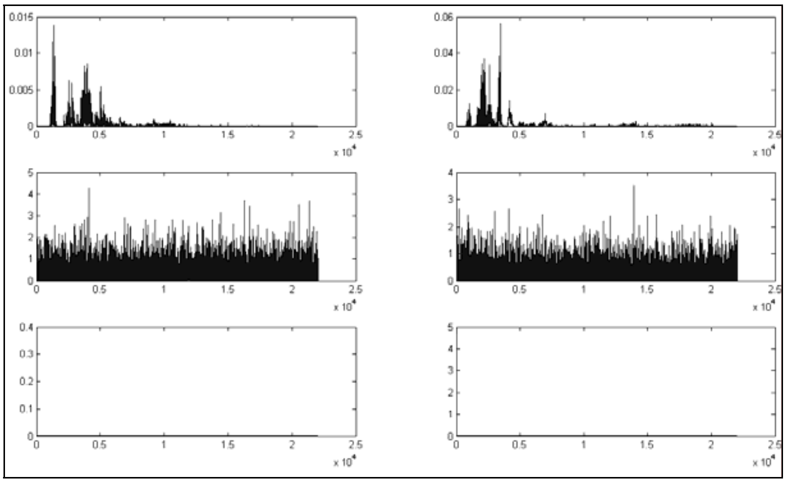

Figure 4.26 Power spectrums of the source, mixture and recovered signals

The power spectrums of the source, mixture and recovered signals have been represented in the first, second and third rows in the, respectively. The source signals certain power spectrum have changed by the nonlinear mixture defined as square root function. The sigmoid kernel function has not been able to recover the source signals as seen in the third row of the figure. This fact has been already observed by the correlation coefficients in Table 4.12.

The nonlinear mixing type of feed forward multilayer perceptron function has been applied to the source signals to observe the affect of the nonlinear mixing type. In the following example, the forward multilayer perceptron nonlinear mixing has been applied to the source signals and the results of the kernelized temporal predictability method based simulations have been observed for the polynomial and sigmoid kernel functions. The mean of correlation coefficients have been collected for different degrees of polynomial kernel function in Table 4.13.

Table 4.12 The mean of correlation coefficients between original source signals and recovered signals for feed forward multilayer nonlinear mixing type and polynomial kernel

	Polynomial Degree	Mean of Correlation Coefficients' Matrix		Standard Deviation of Correlation Coefficients' Matrix		Mean of the Number of Feature Vectors	Standard Deviation of the Number of Feature Vectors
The feed forward multilayer nonlinear mixing type	2	0.3918	0.4937	0.2964	0.3304	7	1.7466
		0.5648	0.4558	0.3164	0.3004		
	3	0.4191	0.6221	0.2491	0.2793	10.5000	1.8722
		0.5442	0.4838	0.2934	0.2657		
	4	0.4045	0.4751	0.3066	0.3073	13.3700	2.5728
		0.4211	0.5044	0.2843	0.2769		
	5	0.3105	0.4981	0.2628	0.2891	16.1500	3.2671
		0.3669	0.5169	0.3018	0.2533		

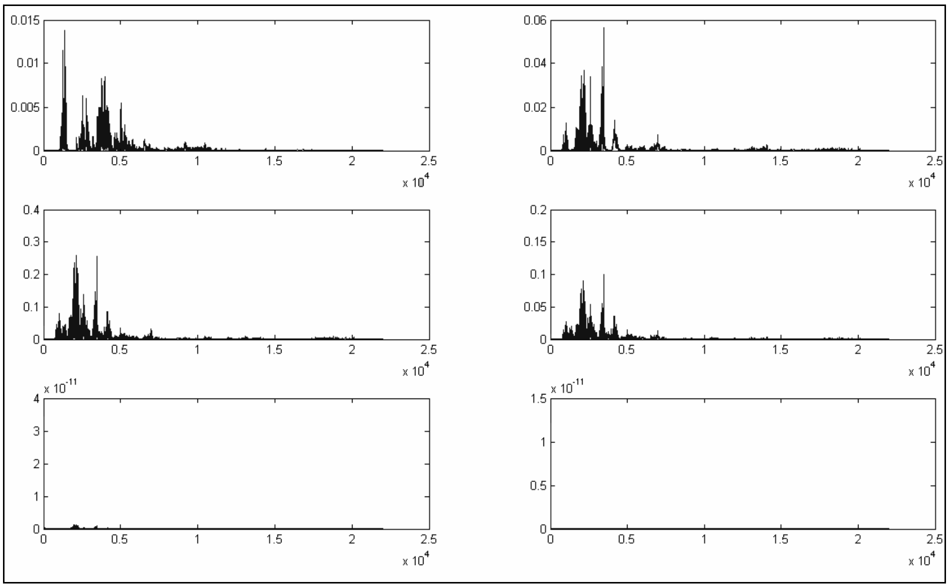

Figure 4.27 Power spectrums of spectrums of the source, mixture and recovered signals

The kernelized temporal predictability method with polynomial kernel function for the feedforward nonlinear mixing type have not formed good performance as seen from the correlation coefficients in the previous table. The power spectrum of the source, mixture and the recovered signals have been represented in the first, second and the third row of the Figure 4.27, respectively. As seen from the power spectrum of recovered signals the kernelized temporal predictability with polynomial kernel function algorithm have not been able to separate the spectrum components for the feedforward nonlinear mixing type that has been expressed in detail in section 3.3. The following correlation coefficients in the table next have been collected by the kernelized temporal predictability method with sigmoid kernel function for the feedforward nonlinear mixing technique. The correlation coefficients have shown that the algorithm with sigmoid kernel function has not been able to separate the sources from defined type of nonlinear mixtures.

Table 4.13 The mean of correlation coefficients between original source signals and recovered signals

	Sigmoid Parameters		Mean of Correlation Coefficients' Matrix		Standard Deviation of Correlation Coefficients' Matrix		Mean of the Number of Feature Vectors	Standard Deviation of the Number of Feature Vectors
The feedforward multilayer nonlinear mixing type	p_1	p_2						
	1	2	0.3604	0.3898	0.2074	0.1538	2	0
			0.6356	0.4151	0.1319	0.1538		
	1	1	0.4731	0.2747	0.1829	0.1180	2	0
			0.6589	0.3532	0.1530	0.1461		
	2	5	0.2181	0.4513	0.2178	0.1857	2	0
			0.5132	0.5571	0.1428	0.1983		
	5	10	0.0994	0.5344	0.1277	0.1721	2.0400	0.2814
			0.3788	0.6824	0.1277	0.1325		

4.3.5 Separation of Nonlinear Mixture of Two Musical Notes using Kernelized Temporal Predictability

Two musical note signals from IOWA database that are AltoSax.NoVib.ff.A4 and SopSax.NoVib.ff.A4 files have been used as source signals and they have been mixed by a nonlinear mixing type. The implemented kernelized temporal predictability method has been used to recover the source signals. The results have been observed for different kernel functions and kernel degrees. The figures of the source, mixture and recovered signals depicted below have been resulted for the polynomial kernel function of degree two.

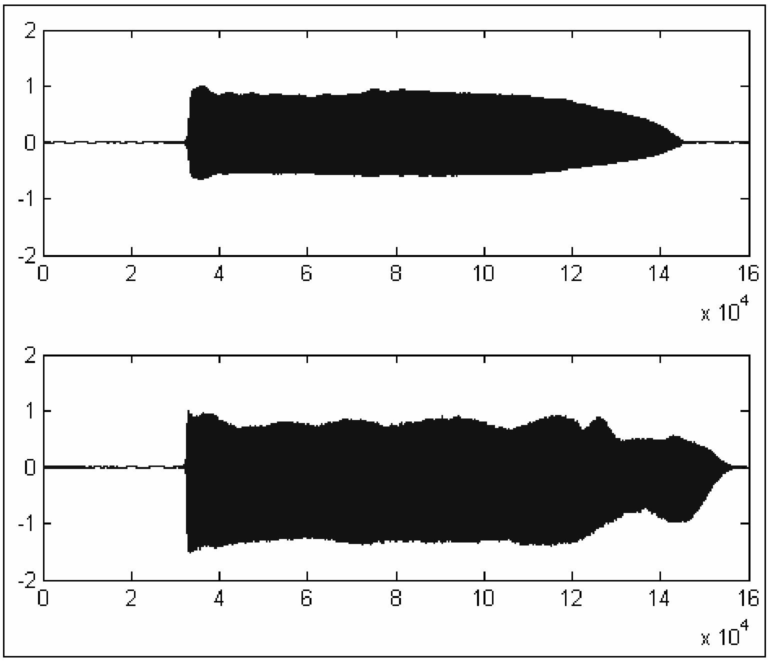

Figure 4.28 Two musical note source signals from Iowa database

The two musical note signals have been represented in the figure above. Both signals have the size of 160000x1 as vectors. The source signals have been mixed nonlinearly by a nonlinear function implemented as a square root function in Matlab and the nonlinear mixture signals have been represented in the next figure.

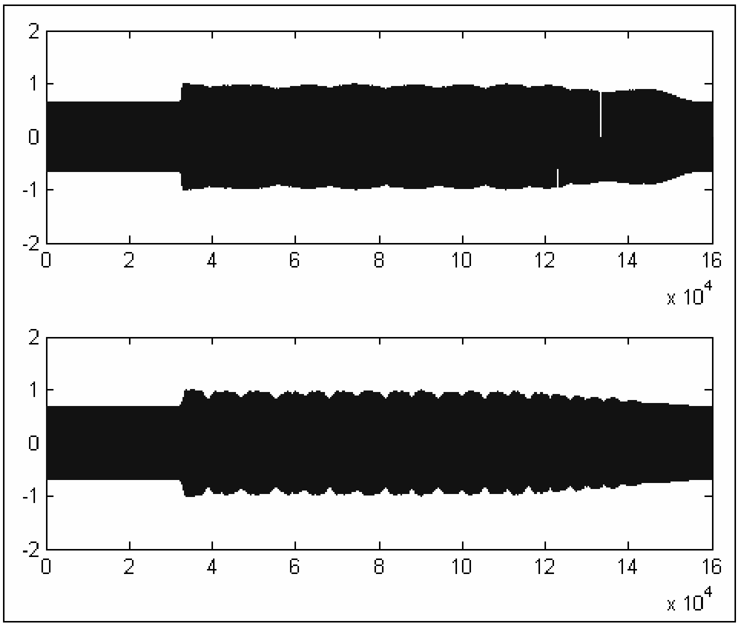

Figure 4.29 The nonlinear mixture signals of two musical note signals from Iowa database

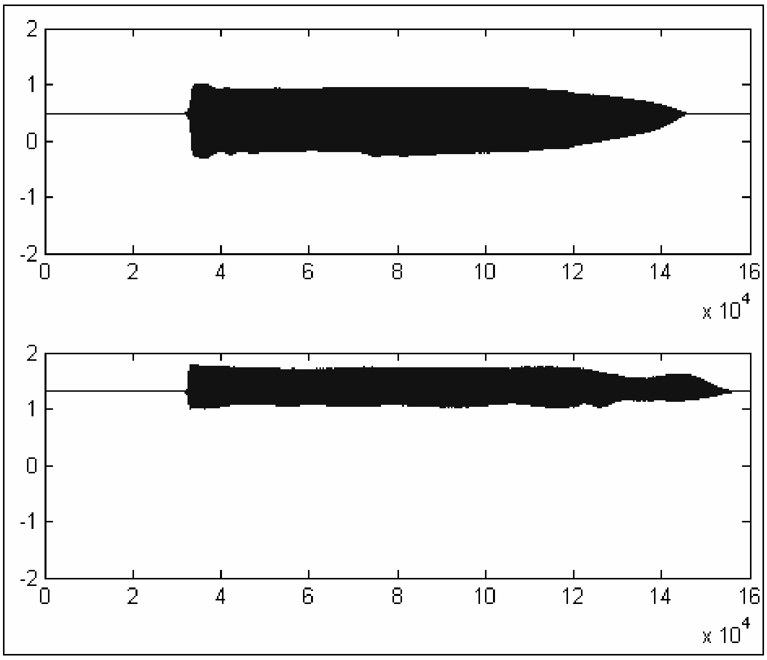

Figure 4.30 The recovered signals from nonlinear mixture signals

For observing the relation between degree of polynomial kernel and number of feature vectors, a table that contains the mean and standard deviation of correlation coefficients and the number of feature vectors have been created by simulating the kernel temporal predictability program for the polynomial kernel function having different degrees. For this purpose, the program has been simulated for one hundred times for each kernel degree.

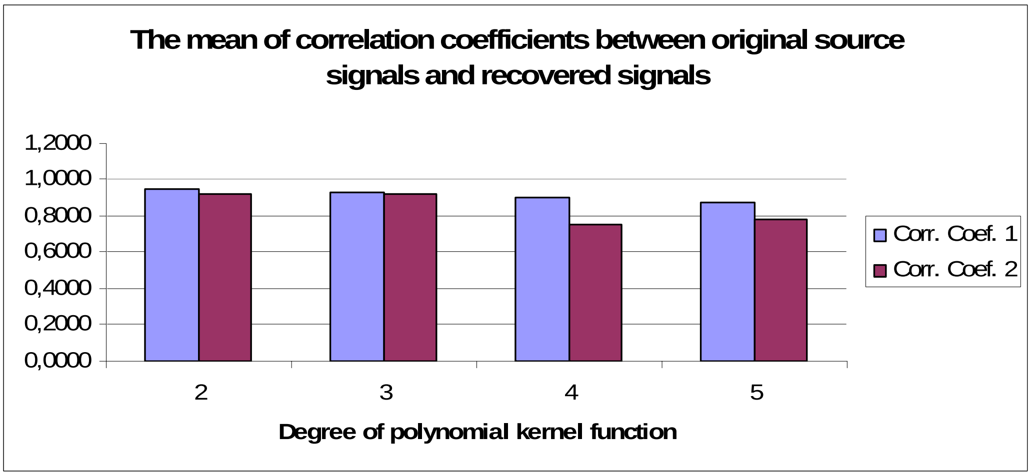

Figure 4.31 The mean of correlation coefficients

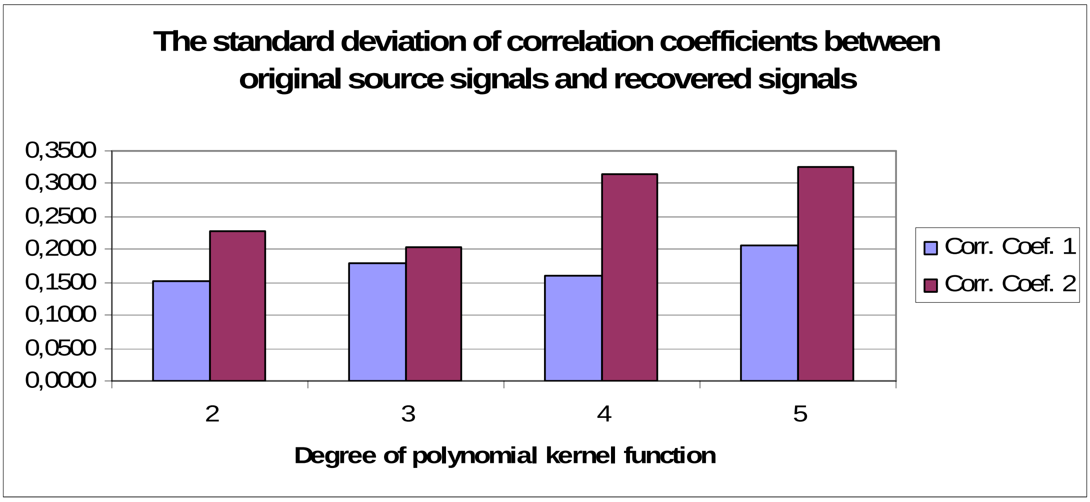

Figure 4.32 The Standard deviation of correlation coefficients

The mean of correlation coefficients have decreased when the polynomial kernel function's has increased. The maximum value of correlation coefficient has been get for polynomial degree of 2. The value is not as good as the results for speech signals. The reason for this can be musical notes own structure. As seen from the results, the number of feature vectors has not much increased according to the kernel degree.

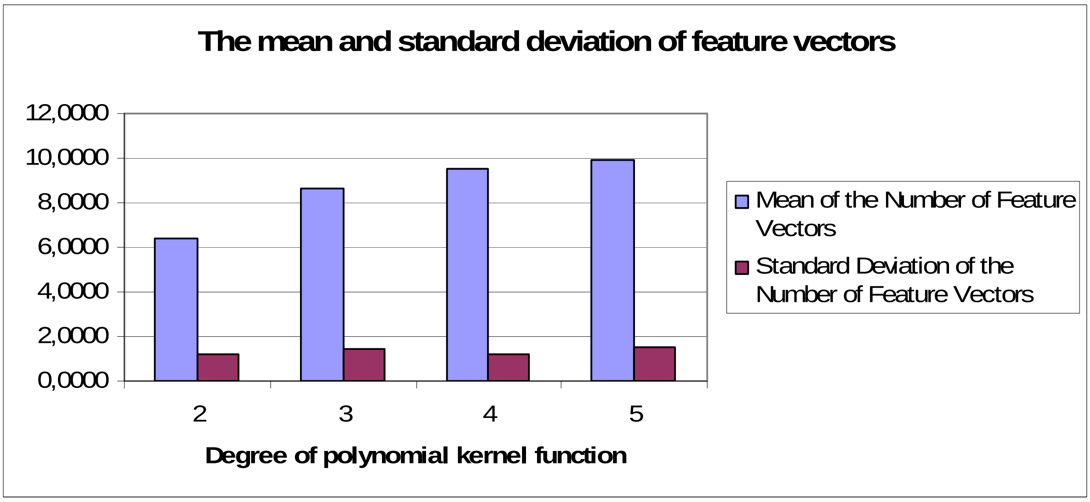

Figure 4.33 The Standard deviation of the number of feature vectors

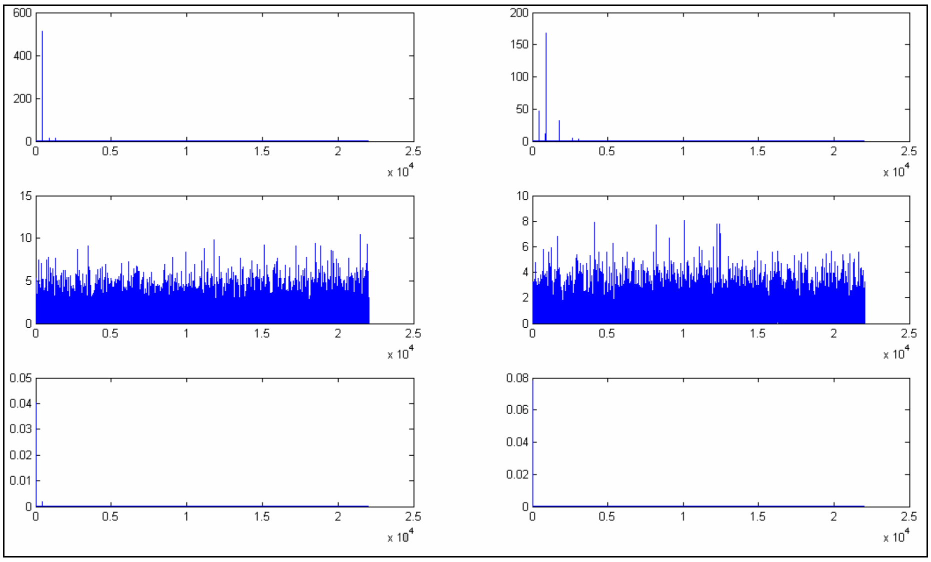

Figure 4.34 Power spectrums of spectrums of the source, mixture and recovered signals

The power spectrums of the source, mixture and recovered signals have been represented in Figure 4.34 The correlation coefficients for the simulation of square root nonlinear mixing type with polynomial kernel have much good as indicated in Table 4.15 But the power spectrums of the recovered signals have not had the source spectrum components any more.

Table 4.14 The Standard deviation of correlation coefficients between original source signals and recovered signals for square root nonlinear mixing and sigmoid kernel functions

	Sigmoid Parameters p_1	p_2	Mean of Correlation Coefficients' Matrix		Standard Deviation of Correlation Coefficients' Matrix		Mean of the Number of Feature Vectors	Standard Deviation of the Number of Feature Vectors
The square root nonlinear mixing type	1	2	0.0487	0.0021	0.0481	0.0019	3	0
			0.1391	0.0080	0.1022	0.0030		
	1	1	0.0650	0.0052	0.0457	0.0065	3	0
			0.0745	0.0074	0.0721	0.0069		
	2	5	0.0293	0.0028	0.0401	0.0027	3	0
			0.0635	0.0052	0.0188	0.0031		
	5	10	0.1073	0.1131	0.0684	0.1198	3	1.1547
			0.2504	0.0797	0.1729	0.1177		

In the previous table, the program has been simulated for the sigmoid kernel function having different parameters to separate the source signals from square root nonlinear mixtures. The results in the table have proven the insufficiency of the sigmoid kernel function for this nonlinear mixing type.

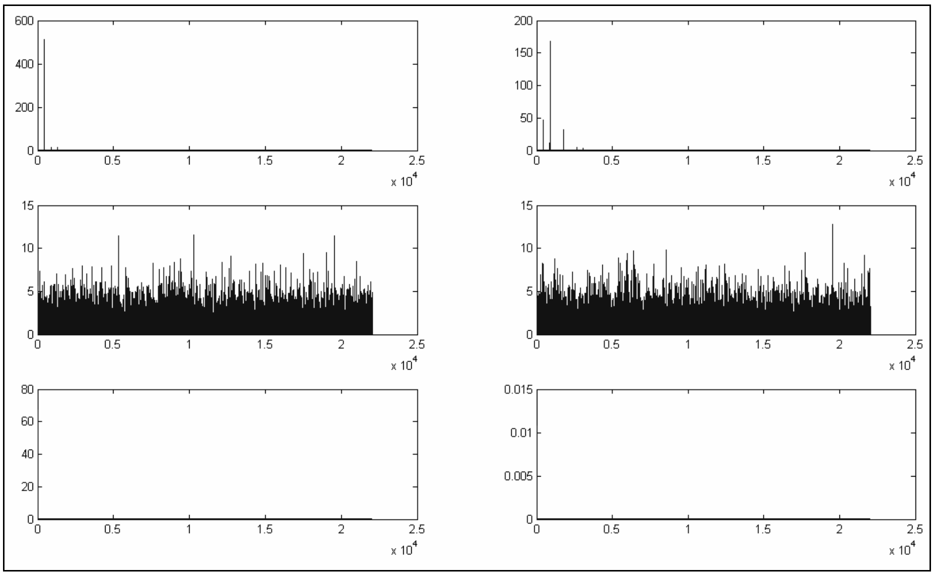

Figure 4.35 Power spectrums of spectrums of the source, mixture and recovered signals

The power spectrums of the source, mixture and recovered signals have been represented in the first, second and third rows of the previous figure. As seen from the power spectrum of the recovered signals, the method has not been able to separate the sources confirming the correlation coefficients in the previous table.

Table 4.15 The Standard deviation of correlation coefficients between original source signals and recovered signals for feedforward multilayer nonlinear mixing and sigmoid kernel function

	Polynomial Degree	Mean of Correlation Coefficients' Matrix		Standard Deviation of Correlation Coefficients' Matrix		Mean of the Number of Feature Vectors	Standard Deviation of the Number of Feature Vectors
The feedforward multilayer nonlinear mixing type	2	0.6557	0.0662	0.3657	0.0378	2.5000	0.5774
		0.5198	0.1224	0.2893	0.0912		
	3	0.7301	0.2225	0.1484	0.1761	2.7500	1.5000
		0.5181	0.2122	0.2019	0.1603		
	4	0.6957	0.1495	0.1238	0.0577	2	0
		0.5753	0.1741	0.1592	0.0565		
	5	0.4649	0.0245	0.3467	0.0426	2.7500	0.5000
		0.7295	0.0267	0.2540	0.0456		

Table 4.16 The mean of correlation coefficients between original source signals and recovered signals for feedforward multilayer perceptron nonlinear mixing and sigmoid kernel function

	Sigmoid Parameters p_1	p_2	Mean of Correlation Coefficients' Matrix		Standard Deviation of Correlation Coefficients' Matrix		Mean of the Number of Feature Vectors	Standard Deviation of the Number of Feature Vectors
The feedforward multilayer nonlinear mixing type	1	2	0.8570	0.1330	0.0909	0.0899	4.2500	2.0616
			0.3240	0.3535	0.1937	0.3781		
	1	1	0.7475	0.1658	0.0449	0.1331	3	0.8165
			0.5707	0.2743	0.0467	0.0971		
	2	5	0.8004	0.1075	0.0834	0.0247	2	0
			0.4443	0.1364	0.1526	0.0516		
	5	10	0.4981	0.1794	0.0711	0.0758	2.7500	0.9574
			0.7668	0.2248	0.0813	0.0193		

The correlation coefficients in Table 4.19 and in Table 4.20 have been collected from the simulations of the method for feedforward nonlinear mixing type. Both polynomial and sigmoid kernel have been simulated and the performance parameters have shown that the kernelized temporal predictability method has not been able to recover the sources from this type nonlinear mixing with the kernel functions used.

4.3.6 Separation of Nonlinear Mixture of Two Recorded Song Signals using Kernelized Temporal Predictability

In this section, two recorded song signals have been used as source signals for nonlinear blind source separation. They have been mixed by a nonlinear mixing transformation (mixing type) and the implemented kernelized temporal predictability method has been used to recover the source signals. The results have been observed for different kernel functions and kernel degrees. The figures of the source, mixture and recovered signals depicted below have been resulted for the polynomial kernel of degree 2.

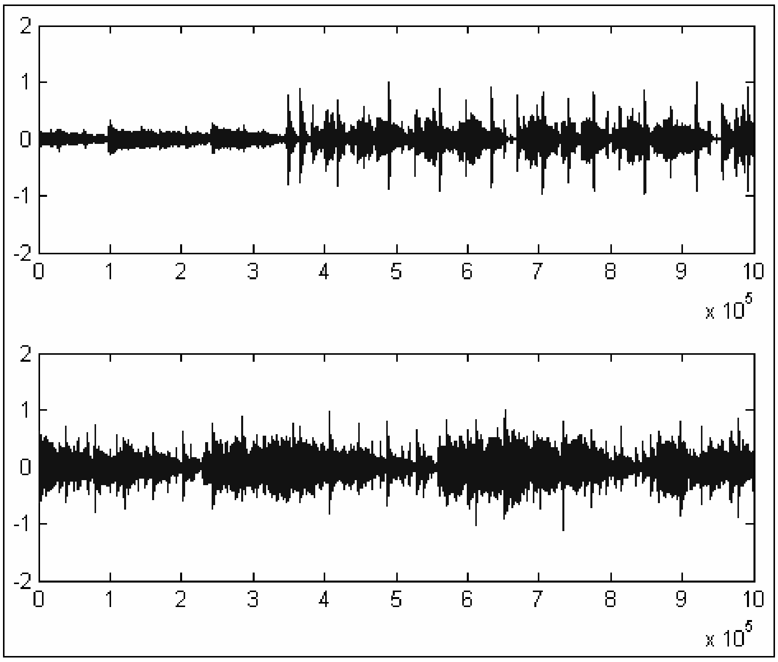

Figure 4.36 Two recorded song signals as source signals

Two recorded song signals as source signals have been shown in the figure above and the nonlinear mixture signals that have been created by a nonlinear function have been represented in the next figure.

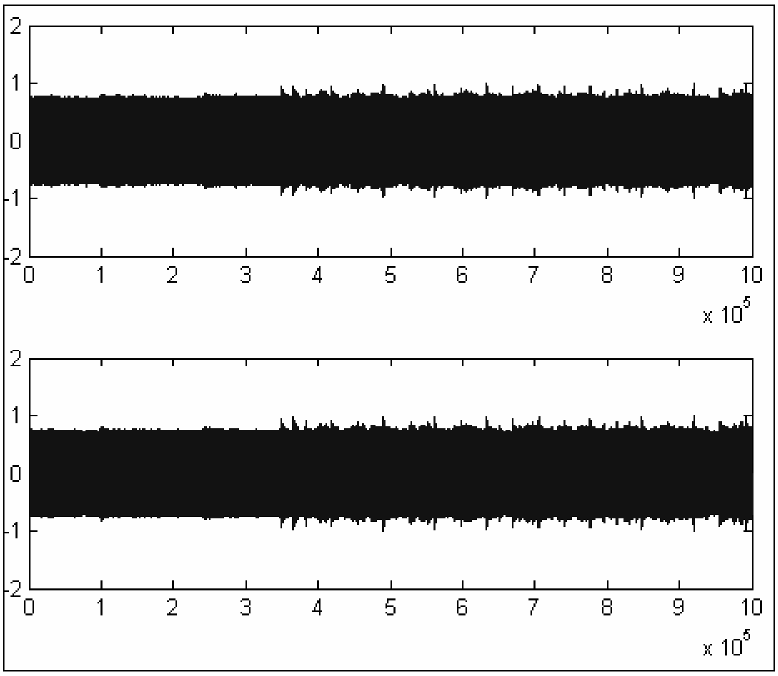

Figure 4.37 The nonlinear mixture signals of two recorded song signals

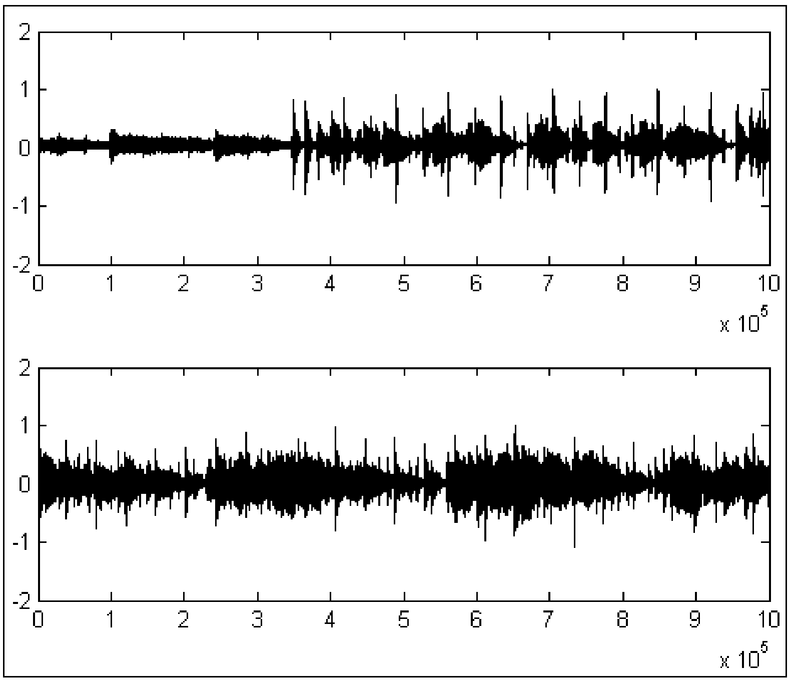

Figure 4.38 The recovered signals from nonlinear mixture signals

The kernelized temporal predictability algorithm has been simulated for the polynomial kernel function having different degrees to see the relation between the performance, kernel degree and the number of feature vectors. A table that contains the mean and standard deviation of correlation coefficients and number of feature vectors has been represented in the following table. The program has been simulated for one hundred times for each value of kernel degree.

As seen from the statistical parameters of the simulations that have been implemented to simulate the kernelized temporal predictability method performance for the nonlinear mixture signals of the sound data types, the algorithm has recovered the source signals successfully depending on the chosen kernel type and kernel parameters.

As a general approach it can be indicated that the performance of the blind source separation of nonlinear mixture signals depends on the source signal own structure, used kernel type and degree. Each source signals own structure is compatible with a feature number range. For the polynomial kernel function degree two gives nearly 6 or 7 feature vector and as seen from the results of the simulation this is compatible with the source signals. If the degree of kernel function has been increased, the performance of the simulations has got worse.

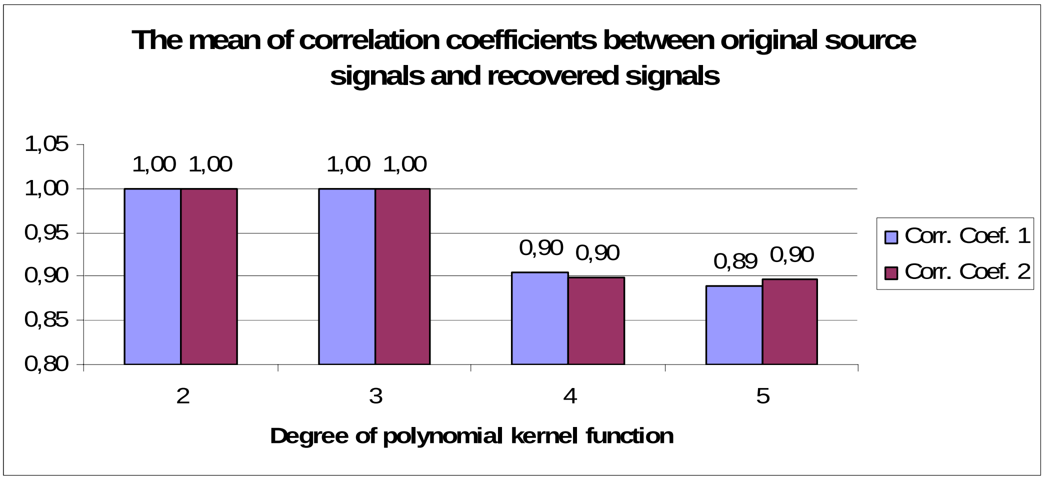

Figure 4.39 The mean of ccorrelation coefficients between original source signals and recovered signals for the square root nonlinear mixing

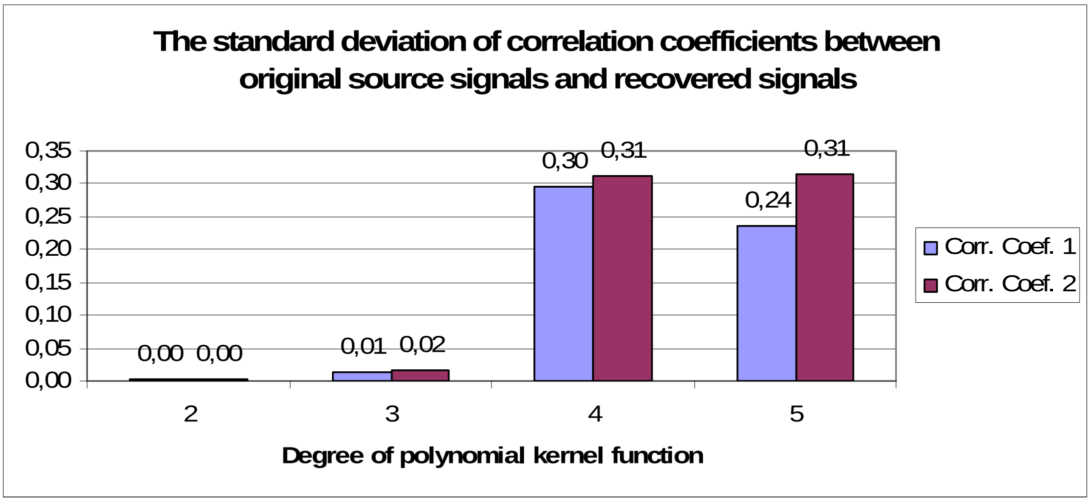

Figure 4.40 The Standard deviation of ccorrelation coefficients between original source signals and recovered signals for the square root nonlinear mixing

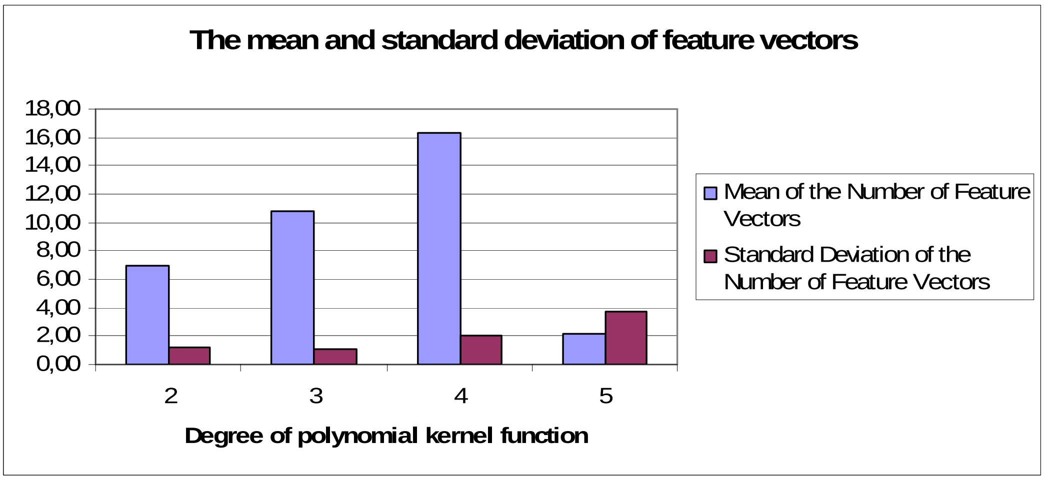

Figure 4.41 The mean and standard deviation of the number of the feature vectors

The power spectrum of the source, mixture and recovered signals of the simulation that has used polynomial kernel with degree 2 and source data mixed by square root kernel function have been represented in Figure 4.42 first, second and third row, respectively. As seen the each recovered signal power spectrum has been almost the same with its own source power spectrum.

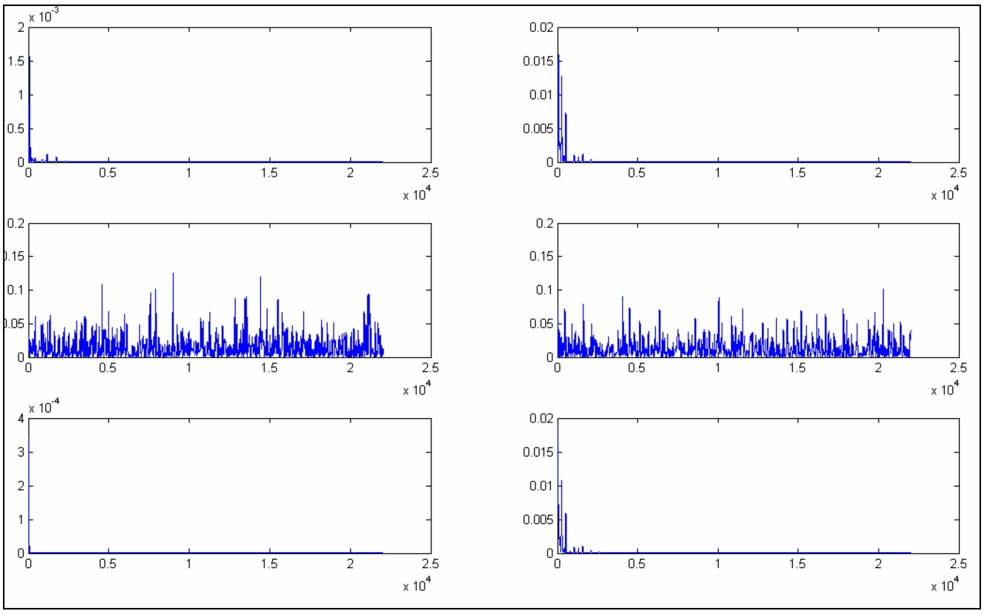

Figure 4.42 Power spectrums of spectrums of the source, mixture and recovered signals

Table 4.17 The mean of correlation coefficients between original source signals and recovered signals for feedforward multilayer perceptron nonlinear mixing and sigmoid kernel function

	Sigmoid Parameters p_1	p_2	Mean of Correlation Coefficients' Matrix		Standard Deviation of Correlation Coefficients' Matrix		Mean of the Number of Feature Vectors	Standard Deviation of the Number of Feature Vectors
The square root nonlinear mixing	1	2	0.0235	0.0006	0.0253	0.0002	3	0
			0.2125	0.0053	0.0049	0.0035		
	1	1	0.0376	0.0071	0.0147	0.0071	3	0
			0.1036	0.0048	0.0481	0.0043		
	2	5	0.0153	0.0010	0.0181	0.0009	3	0
			0.0924	0.0045	0.0612	0.0075		
	5	10	0.1774	0.0132	0.1345	0.0178	3.7500	0.9574
			0.2887	0.0152	0.1674	0.0253		

The sigmoid kernel function has not been able to separate the source signals from the nonlinear mixtures. As seen the number of feature vectors extracted by the algorithm has been decreased by comparing with the polynomial kernel.

Table 4.18 The Standard deviation of correlation coefficients between original source signals and recovered signals for feedforward multilayer perceptron nonlinear mixing and polynomial kernel functions

	Polynomial Degree	Mean of Correlation Coefficients' Matrix		Standard Deviation of Correlation Coefficients' Matrix		Mean of the Number of Feature Vectors	Standard Deviation of the Number of Feature Vectors
The feedforward multilayer nonlinear mixing type	2	0.6221	0.3377	0.3607	0.4209	6	0.8165
		0.3656	0.5447	0.3887	0.4516		
	3	0.6351	0.5238	0.3932	0.5369	10.2500	0.9574
		0.0885	0.5090	0.0338	0.5526		
	4	0.9615	0.2850	0.0237	0.3836	12.7500	2.8723
		0.0803	0.6324	0.0741	0.4068		
	5	0.6059	0.3511	0.4483	0.4633	16.2500	0.9574
		0.0602	0.5370	0.0578	0.5310		

Table 4.19 The mean of correlation coefficients between original source signals and recovered signals for feedforward multilayer perceptron nonlinear mixing and sigmoid kernel function

	Sigmoid Parameters		Mean of Correlation Coefficients' Matrix		Standard Deviation of Correlation Coefficients' Matrix		Mean of the Number of Feature Vectors	Standard Deviation of the Number of Feature Vectors
	p_1	p_2						
The feedforward multilayer nonlinear mixing type	1	2	0.2861	0.0438	0.3677	0.0115	3	0
			0.3952	0.1132	0.3277	0.0382		
	1	1	0.6165	0.0998	0.2385	0.0718	2.2500	0.5000
			0.6498	0.2804	0.2970	0.1253		
	2	5	0.3735	0.1782	0.2583	0.1655	2	0
			0.7848	0.3582	0.1487	0.2473		
	5	10	0.3883	0.0717	0.3504	0.0346	3	0.8165
			0.6891	0.3444	0.1071	0.3233		

The source signals mixed with feedforward nonlinear multilayer perceptron function have not been able to been separated by the kernelized temporal predictability algorithm for both polynomial and sigmoid kernel function. The following power spectrum has been plotted from the results of the simulation that has used the polynomial kernel with degree 2 for the feedforward multilayer perceptron nonlinear mixtures of sources.

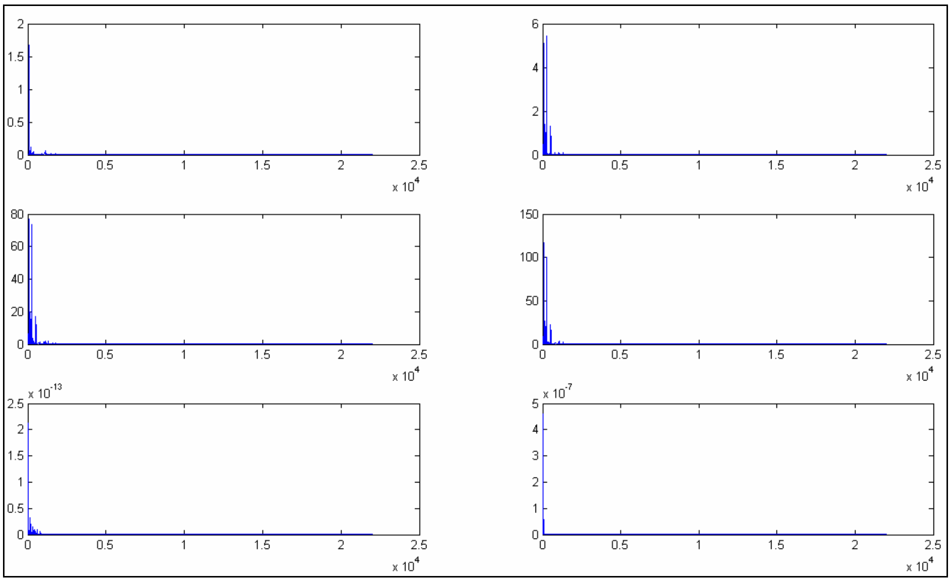

Figure 4.43 Power spectrums of spectrums of the source, mixture and recovered signals

The power spectrums of the source, mixture and recovered signals of the simulations have been represented for all simulations in Chapter 4 to show the spectral components of the signals and observe the changes caused by the linear and nonlinear mixing. The source signals own power spectrum have differ each other as seen. The speech signals, musical notes and songs have very different spectral components distribution. Power spectrums of the three source signals have been represented in order of speech, musical note and song in the figure next.

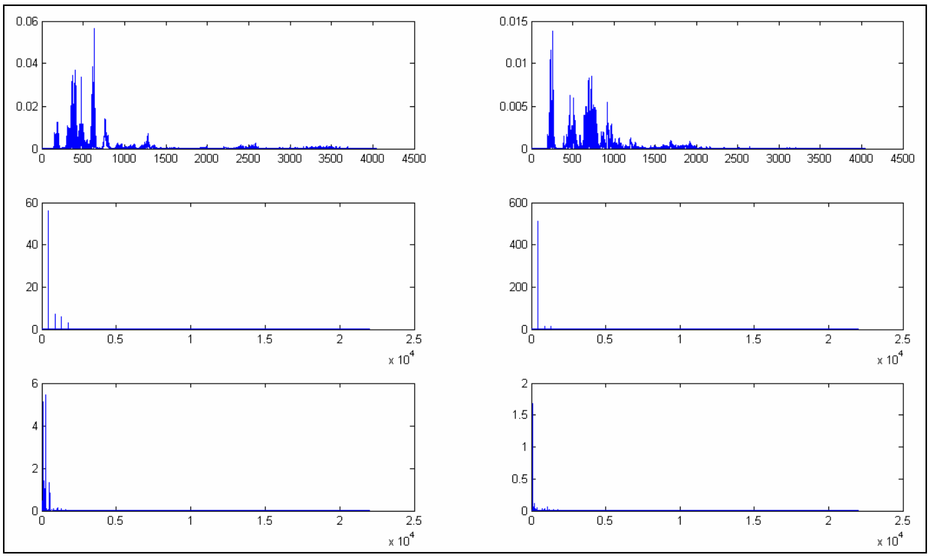

Figure 4.44 Power spectrums of spectrums of the source signals

CHAPTER FIVE
SEPARATION OF IMAGE MIXTUES

5.1 Introduction

In this study, different types of image data files (like face and writing; face and object image) have been used as source signals. The motivation of applying source separation to image mixtures comes from the recovery of the corrupted writings or images. In literature, there are several works deals with the image recovery problem with different kinds of methods in (Liang, Ahmadi & Shridbar 1997) (Almeida &Almeida, 2004) (Kasturi, Bow, El-Masri, , Shah, Gattiker & Mokate, 1990). Also, some simulations for face recognition applications have been implemented. However, the simulations have not resulted well since the sources seem not independent form each other. This application may have better results if some specifications like nose location and the other parameters, which are case in most of the face recognition applications, are given.

A second consideration for the linear and nonlinear blind source separation of image data has been thought as the image data mass size. The mass size means how many pixels of the image as a square are taken as one mass and converted to the row vector and these row vectors of the masses are combined to get a total row vector for an image data file. For the following examples in this section, the different mass sizes have been chosen as they have been image data size modules.

5.2 Separation of Linear Mixture Images Using Temporal Predictability Algorithm

In this section, different image files have been used as source signals to indicate that the performance of the algorithm is independent of the image data used particularly.

5.2.1 Motorcycle and Plane Images Separation using Temporal Predictability

The motorcycle and plane images have been taken from the face recognition database website. The image sizes have been set to 250 pixels as row and 450 pixels for columns and the implemented program has been simulated for the resolution mass sizes of 1, 5, 10, 25 and 50. The program has been simulated for one hundred times for each mass size. The statistical parameters of the mean and standard deviation of correlation coefficients have been collected in Figure 5.4 and Figure 5.5 to represent the linear blind source separation performance and mass size relation.

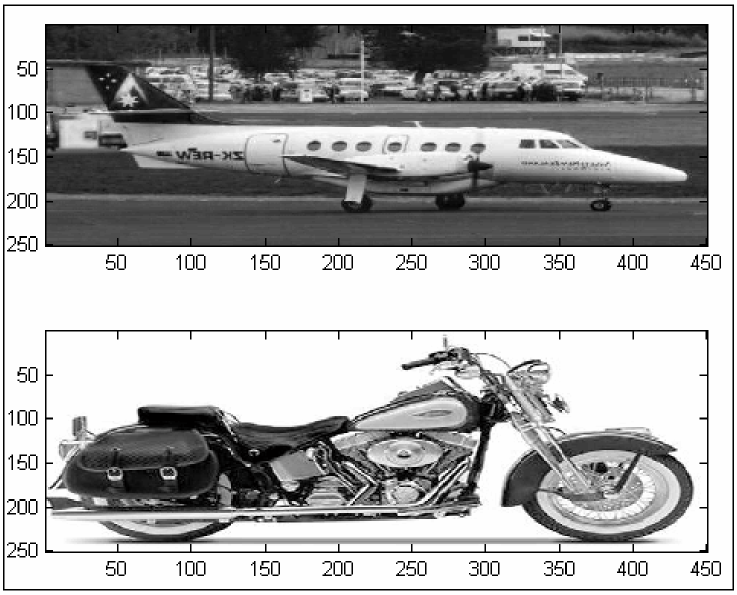

Figure 5.9 The source image data files

The source image files have been represented in Figure 5.1 and the signals have been mixed linearly by a random mixing matrix A. The linear mixture signals of source images have been shown in Figure 5.2.

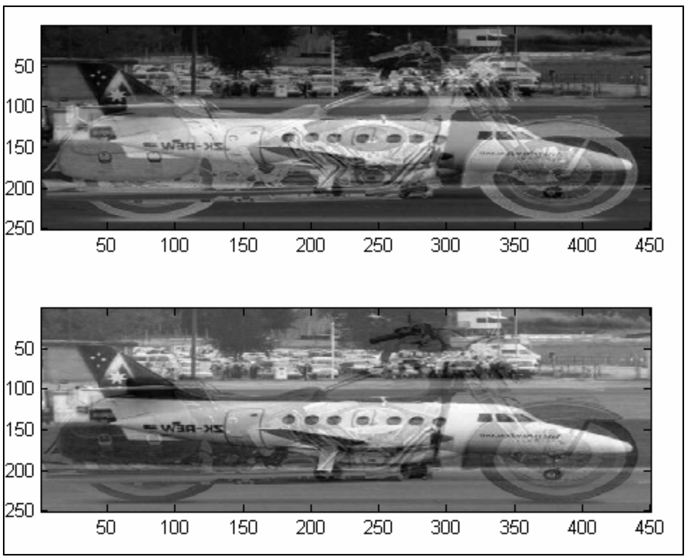

Figure 5.10 The linear mixture signals of source image files

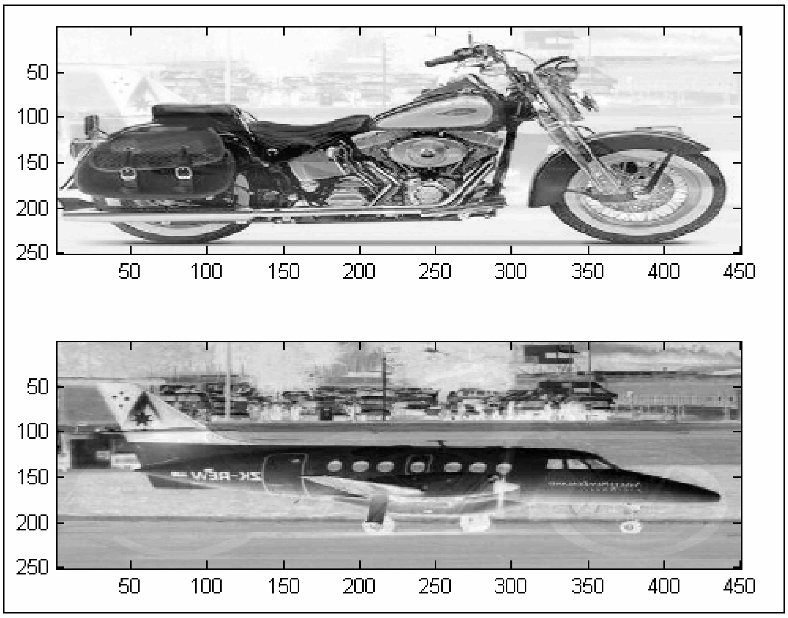

Figure 5.11 The recovered images from linear mixture signals

The recovered images have been represented in Figure 5.3. As seen from the recovered images, the backplane of the images have been affected each other. The statistical performance parameters of mean and standard deviation of correlation coefficients have been given in the following two figures, respectively.

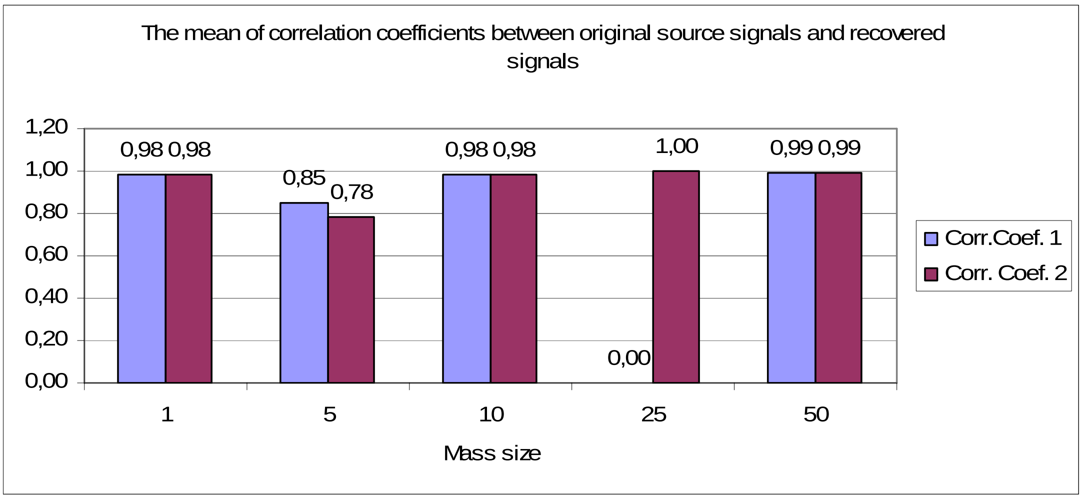

Figure 5.12 The mean of correlation coefficients between original source signals and recovered signals

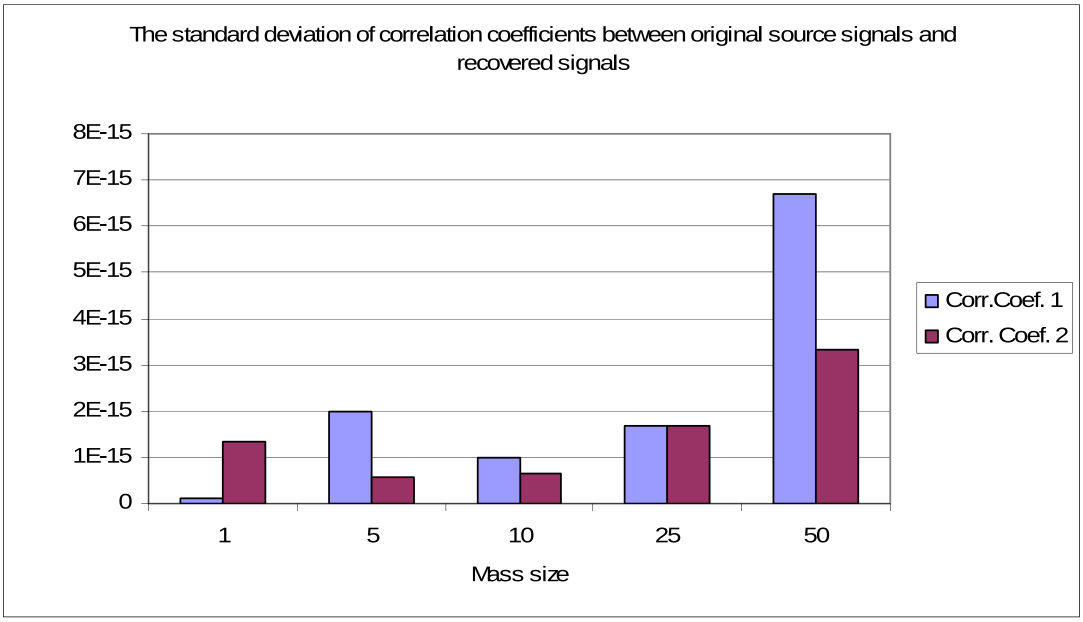

Figure 5.13 The standard deviation of correlation coefficients between original source signals and recovered signals

According to the simulation results in Figure 5.4, it can be exactly stated that the mass size has affected the mean value of the correlation coefficients. The mass sizes of 25 and 50 have given almost the same and the best results. The standard deviation values of the correlation coefficients have been very small for all mass size values. This fact has proven the temporal predictability algorithm's stable performance for linear blind source separation of image data.

5.2.2 Building and Bridge Images Separation using Temporal Predictability

In the previous example, the temporal predictability method has been simulated for the plane and motorcycle images having the size of 250x450. A different set of images having the size of 320x420 has been simulated for one hundred times to get the statistical performance parameters of mean and standard deviation of correlation coefficients.

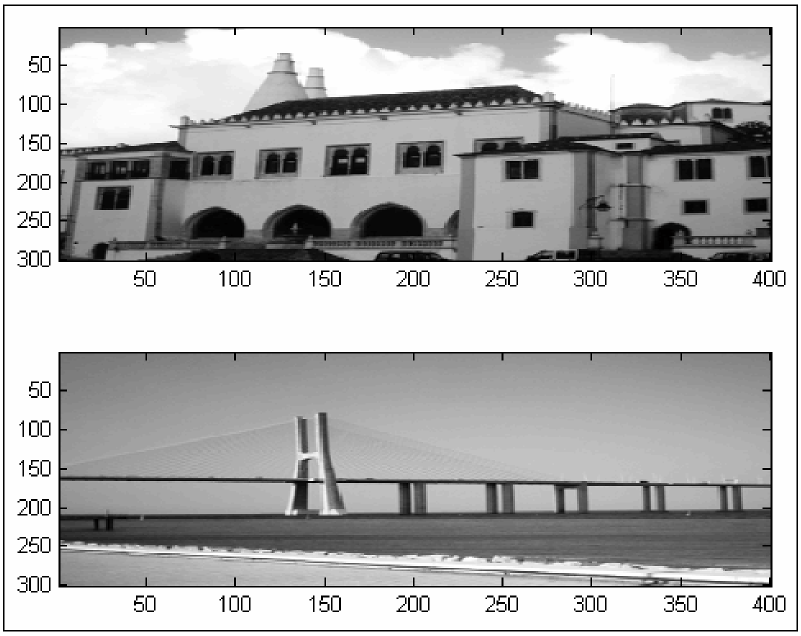

Figure 5.14 The source image data files

The source images have been represented in Figure 5.6. A building and a bridge images that have been used as source signals and the source signals have been mixed linearly by a random mixing matrix A in Matlab. The mixtures of source images have been shown in Figure 5.7.

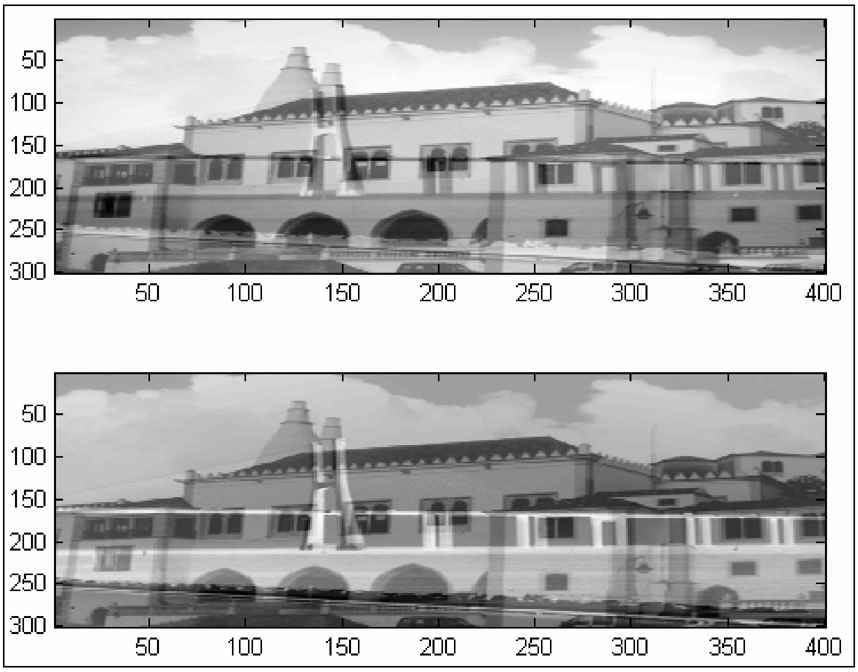

Figure 5.15 The linear mixture signals of original source image files

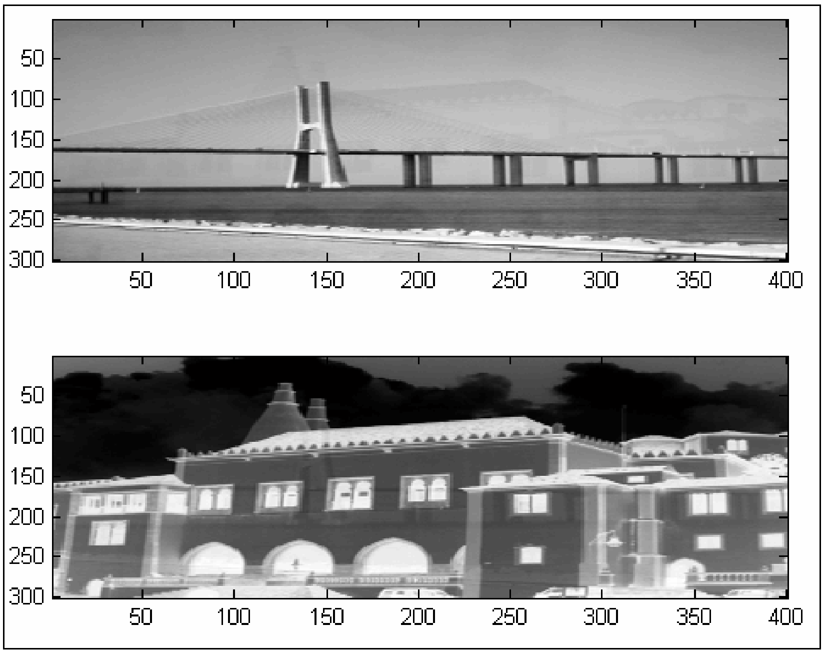

Figure 5.16 The recovered images from linear mixture signals

The recovered images have been represented in Figure 5.16. The linear blind source separation of image mixture signals have been achieved as seen from the figure and from the statistical performance parameters of the simulation the following two figures.

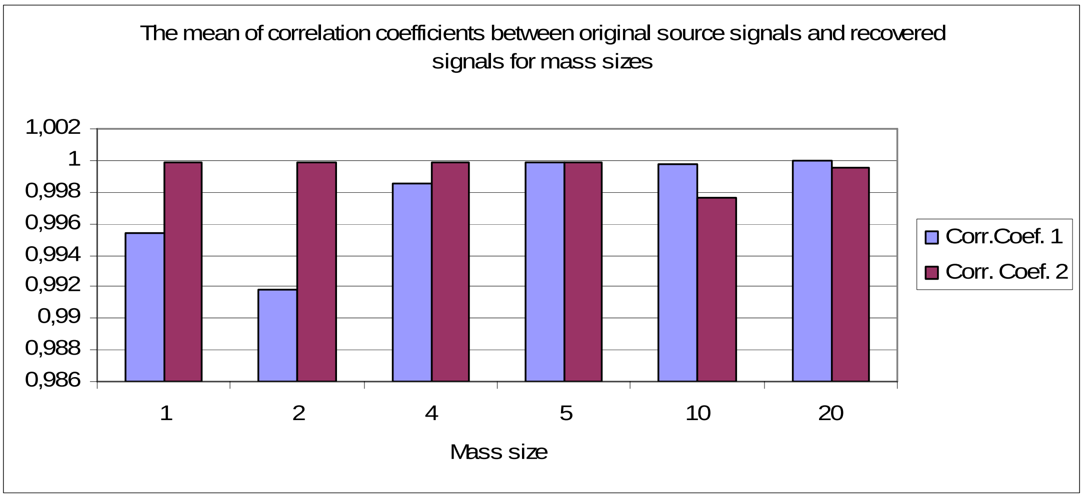

Figure 5.17 The mean of correlation coefficients between original source signals and recovered signals for mass sizes

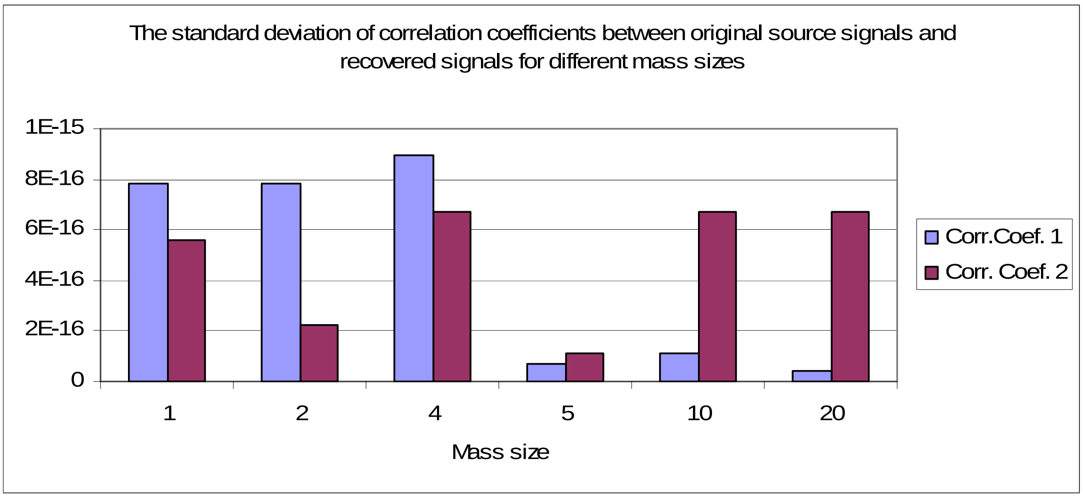

Figure 5.18 The standard deviation of correlation coefficients between original source signals and recovered signals for different mass sizes

The mean values of the correlation coefficients in Figure 5.9 have been almost the same for all mass size values. There is not much an apparent difference between the results of the corresponding mass sizes. The temporal predictability algorithm's performance for the linear blind source separation of images has represented an overall stability as seen from the standard deviation values in the figure above. The algorithm's stability has been the same for the motorcycle and plane images in the previous example.

The mean values of the correlation coefficients for two sets of images in example 1 and example 2 have drawn a different profile for mass size values. In example 1, the mean values of correlation coefficients have changed depending on the mass size apparently but in example 2 the values have not changed so much. The fact that the effect of image mass size value selected depends on the image own structure can be expressed according to the results.

5.3 Separation of Nonlinear Mixtures of Images Using the Kernelized Temporal Predictability Method

The kernelization of the temporal predictability method has been implemented in (Dominique Martinez and Alistair Bray 2003) have been used for nonlinear blind source separation of image data. Different types of image files have been used as source signals to test the algorithm performance.

The kernel type, the kernel function's parameters, number of feature vectors and the size of the data that the feature vector selection algorithm extracts the feature vectors have effected the correlation between source and recovered images. The program based on the kernelization of temporal predictability method has been simulated for three sets of images for one hundred times to get the statistical performance parameters. The mean value and the standard deviation of correlation coefficients have been collected for each data set in the following. The affect of kernel type, kernel degree and the number of feature vectors have been observed by these simulations.

5.3.3 Separation of Nonlinear Mixtures of a Face Image and an Object Image

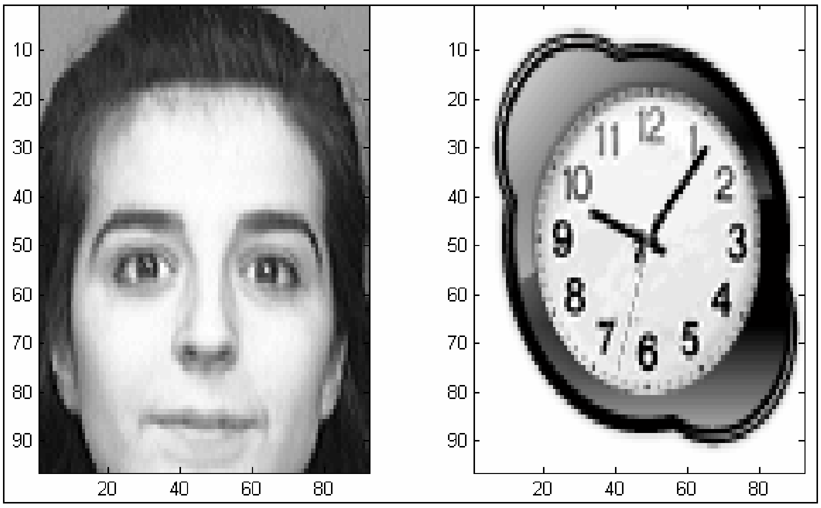

Figure 5.19 The source image data files

The source images have been shown in Figure 5.11. The size of the data that the feature vector selection algorithm has extracted the feature vectors is 5000. The size has been restricted by the memory size that the higher values have caused out of memory error. The polynomial kernel function has been used as kernel function for its different degrees. The program has been simulated for one hundred times for each kernel degree. The statistical values of mean and standard deviation of correlation coefficient have been collected in the following figures. The polynomial kernel function's degree has affected the number of feature vectors. The mean and standard deviation of the number of the feature vectors have been represented in Figure 5.12.

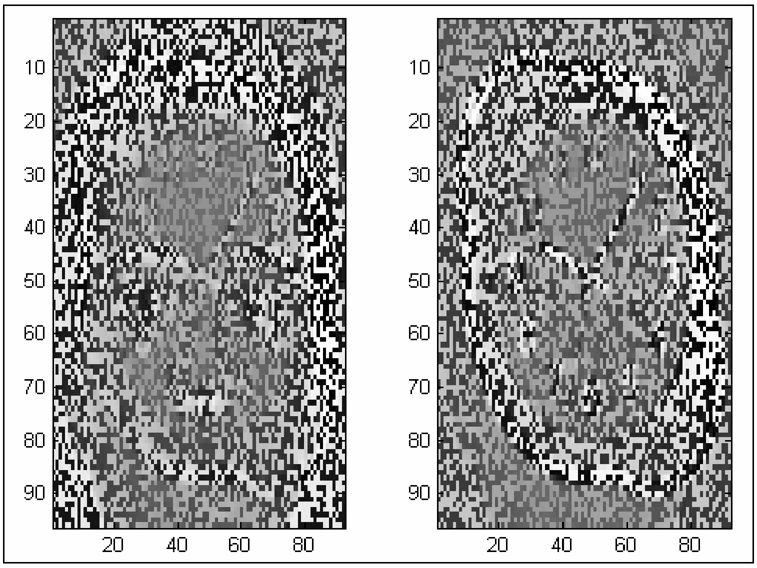

Figure 5.20 The linear mixture signals of original source image files

The source images have been transformed by a nonlinear transformation that has been defined by square root function and the nonlinear mixture of source images have been represented in Figure 5.12.

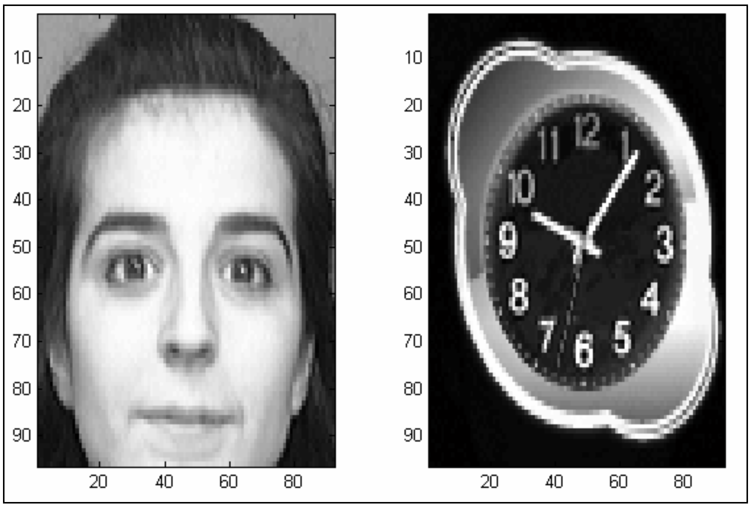

Figure 5.21 The recovered images from nonlinear mixture signals

The source images have been recovered from the nonlinear mixtures of images successfully as shown in Figure 5.13. The mean value of the correlation coefficients has been analyzed for both image mass size and polynomial kernel degree in the following figure.

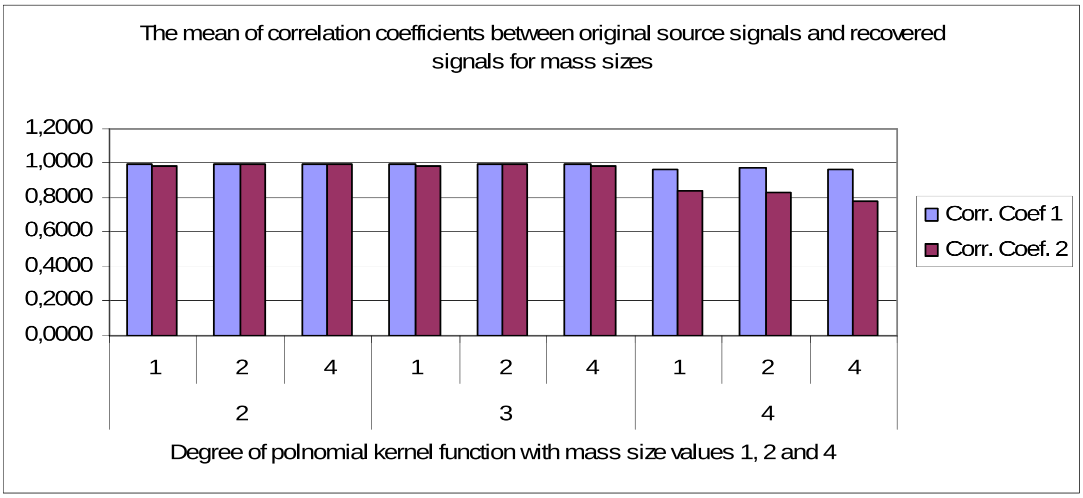

Figure 5.22 The mean of correlation coefficients between original source signals and recovered signals for mass sizes

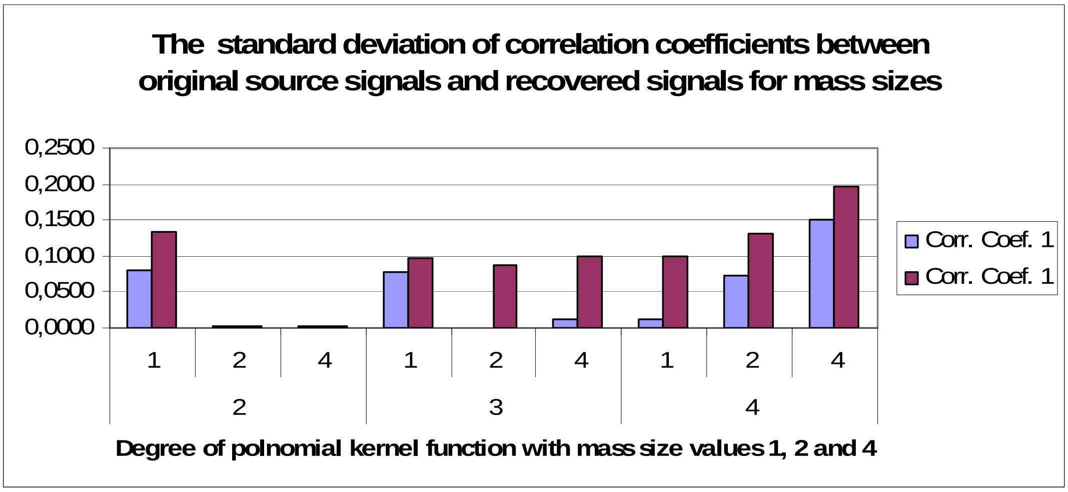

Figure 5.23 The standard deviation of correlation coefficients

The mean values of the correlation coefficients have changed depending on the kernel degree. It has the maximum value for kernel degree of 2 and 3. As seen from the results in Figure 5.15, the mass size has not much affected the correlation value for kernel degree 2 and 3. The values for the kernel degree of 4, the mass size has affected the correlation values much more apparently. Three different image sets have been used to determine whether the parameters have been dependent on the image data own structure.

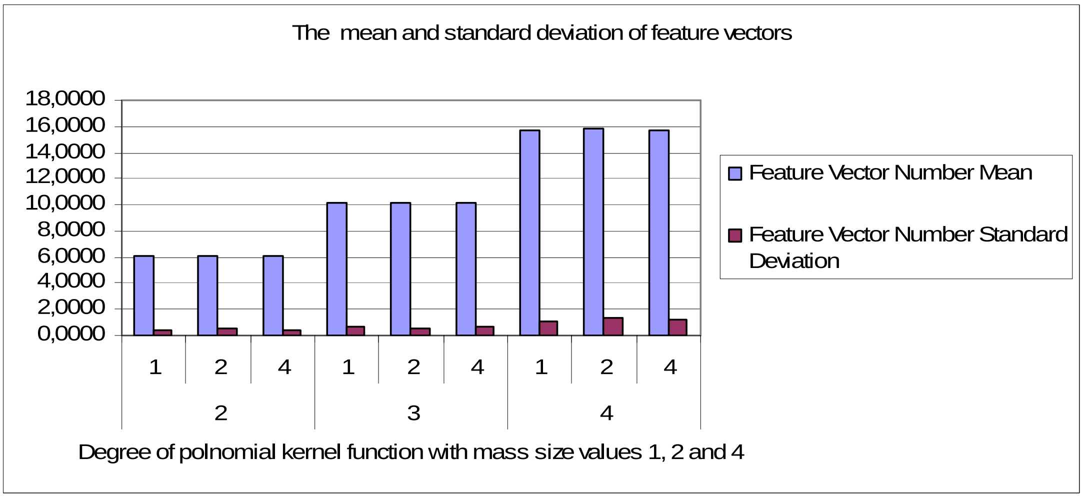

Figure 5.24 The mean and standard deviation of number of feature vectors

The mean and standard deviation of the number of the feature vectors have been collected in the table above. The results collected have indicated that the number of feature vector has depended on the polynomial kernel's degree but the image mass size have not affected the number of feature vectors apparently.

5.3.4 Separation of Nonlinear Mixtures of a Face Image and a Writing Image

The second image set is two images that are a face image and writing. These images have different structures. The kernelized temporal predictability method has used these signals as source signals. The program has been simulated for different kernel functions, polynomial kernel degrees and the number of feature vectors. The size of the mixture data that from which the feature vector selection algorithm extracts the feature vectors have been set to 5000, because upper values have caused out of memory error.

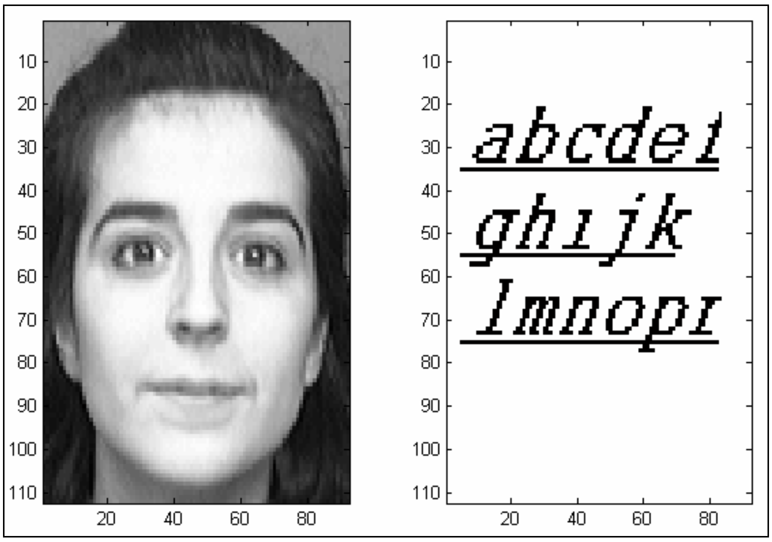

Figure 5.25 The source image data files

The source images have been represented in the figure above. The source images have been mixed by a nonlinear transform function defined as square root function. The nonlinear mixture images have been represented in the next figure.

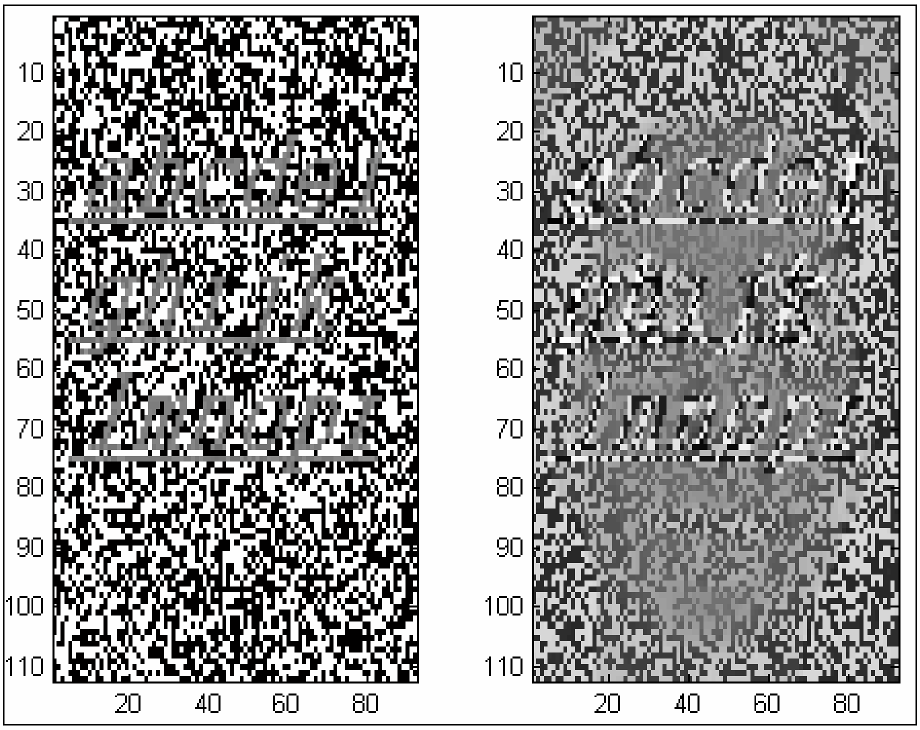

Figure 5.26 The linear mixture signals of original source image files

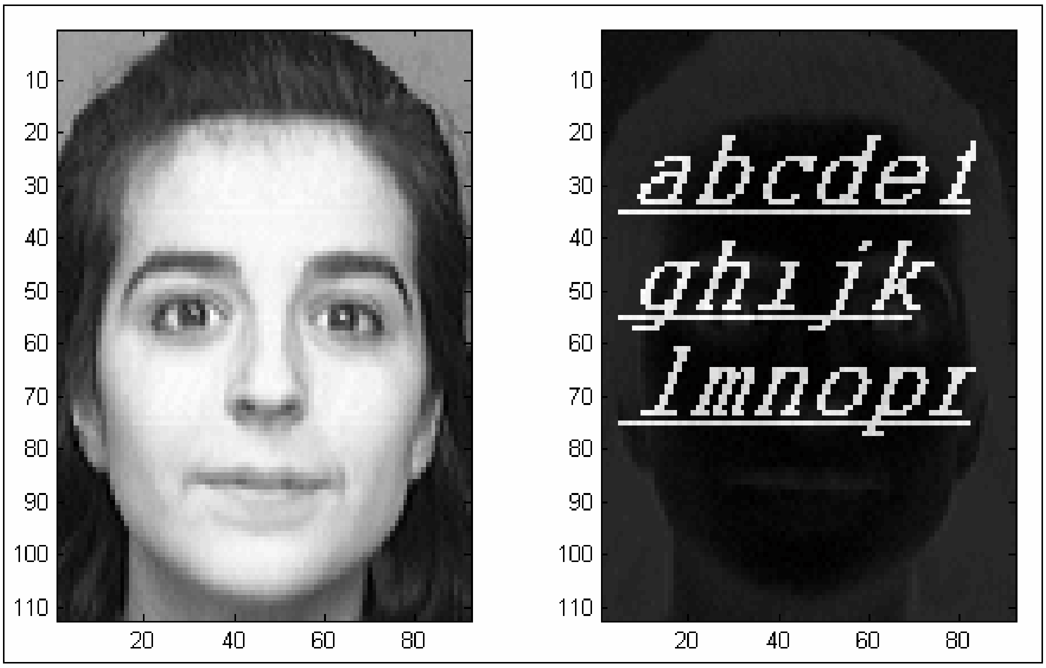

Figure 5.27 The recovered images from nonlinear mixture signals

The source images have been recovered from the nonlinear mixture images by the kernelized temporal predictability method. The polynomial kernel function with degree 2 has been used to get the recovered images in the figure above. The program has been simulated for the different image mass size values and kernel degrees to observe the relation.

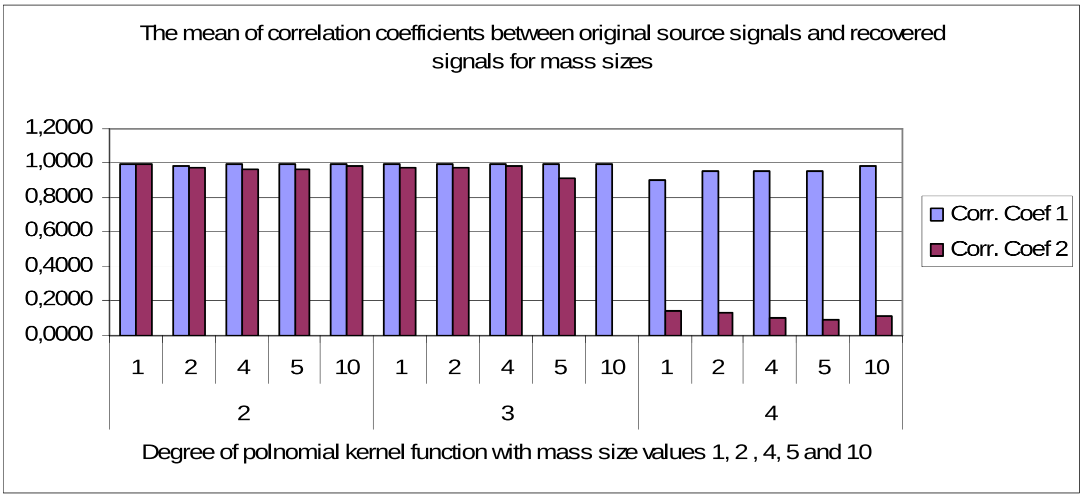

Figure 5.28 The mean of correlation coefficients between original source signals and recovered signals for mass sizes

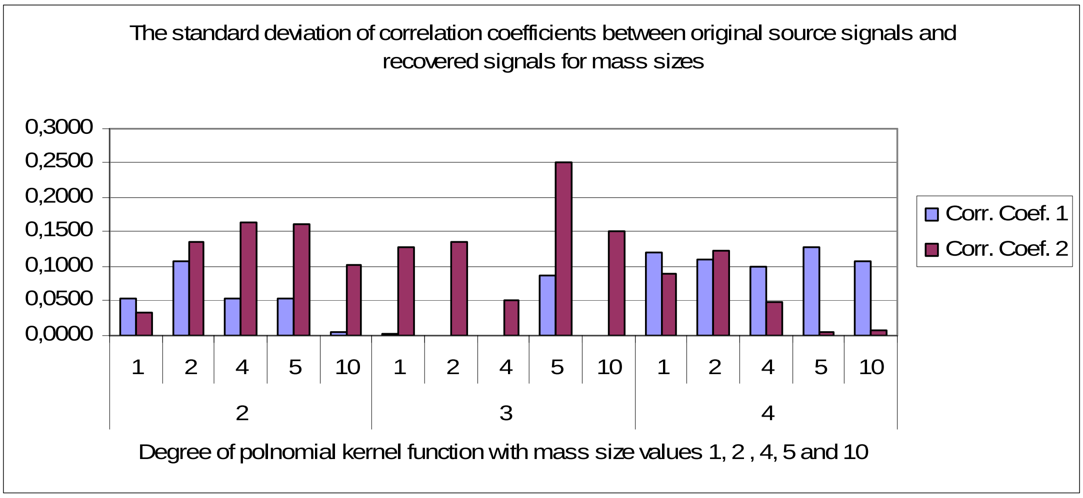

Figure 5.29 The standard deviation of correlation coefficients between original source signals and recovered signals for mass sizes

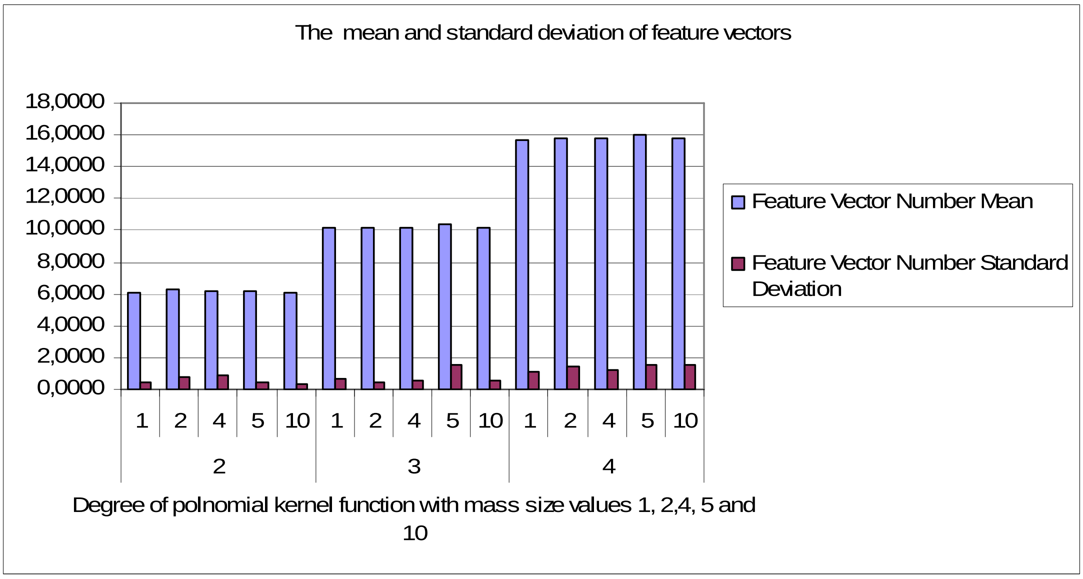

Figure 5.30 The mean and standard deviation of feature vectors

5.3.5 Separation of Nonlinear Mixtures of a Face Image and a Scribble

The face image and a scribble image have been used as source images in Figure 4.45. The data size that the feature vector selection algorithm extracts the feature vectors has been set to 5000 because of the out of memory error. The program has been simulated for polynomial kernel function having different degrees and for different mass sizes. So that mass size affect has been simulated for three different sets o images with this simulation.

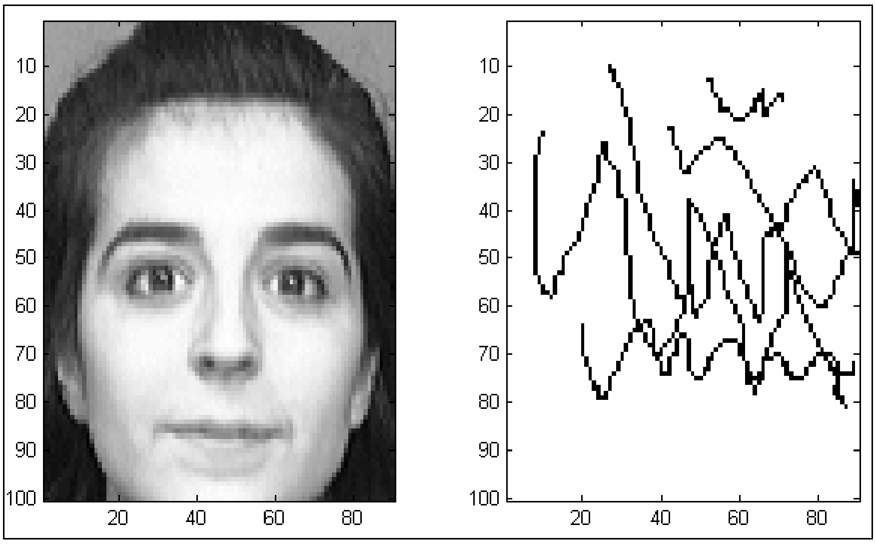

Figure 5.31 The source image data files

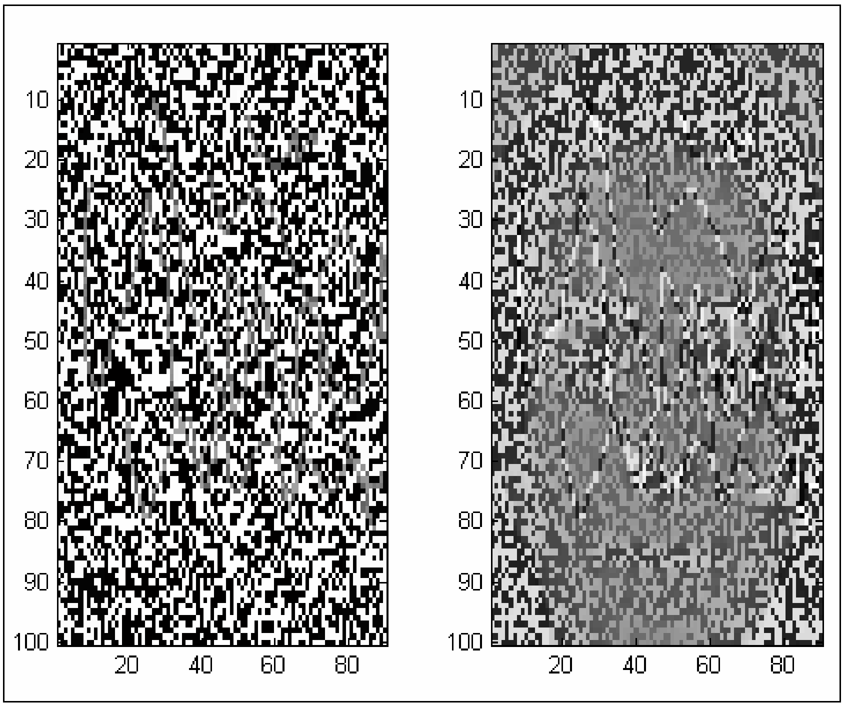

Figure 5.32 The linear mixture signals of original source image files

The source images have been mixed nonlinearly by a nonlinear function defined a square root function. The mixture images have been represented in the figure above.

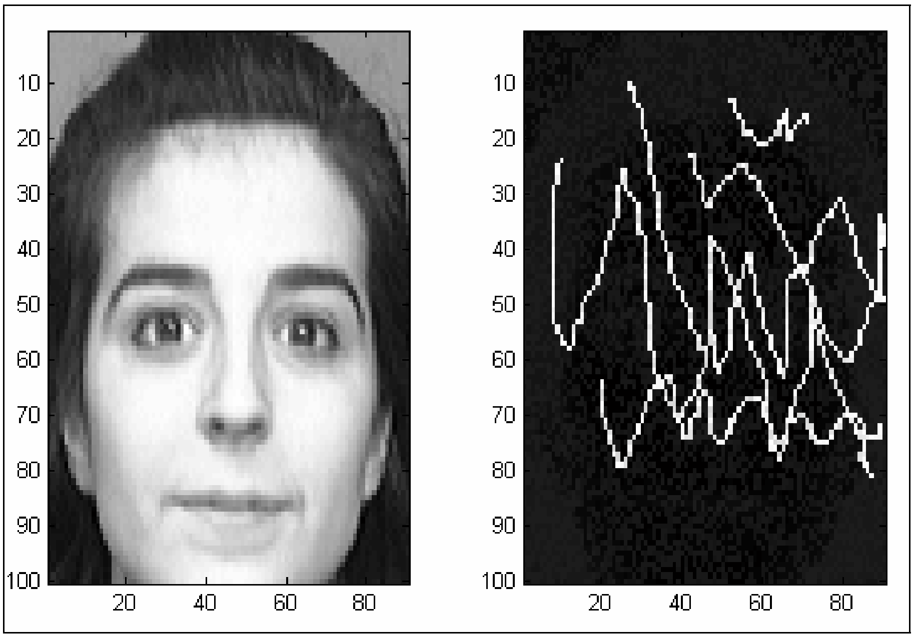

Figure 5.33 The recovered images from nonlinear mixture signals

The source images have been recovered successfully as seen in the figure above. The mean and the standard deviation of correlation coefficients have been given in the following two figures, respectively. The mass size affect is much more apparent for this simulation.

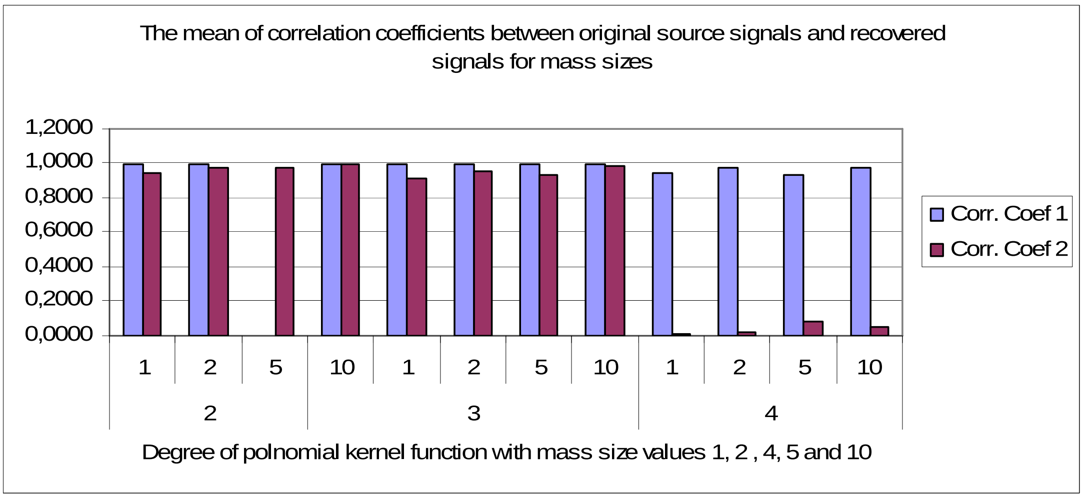

Figure 5.34 The mean of correlation coefficients between original source signals and recovered signals for mass sizes

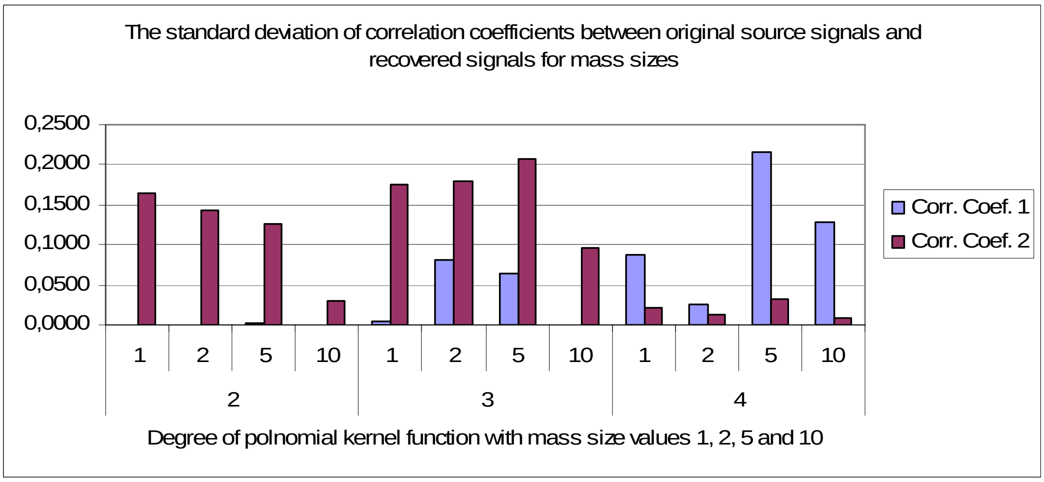

Figure 5.35 The standard deviation of correlation coefficients between original source signals and recovered signals for mass sizes

The correlation coefficients have been much better for the polynomial degree 2 and 3. The mass size corresponding values for each polynomial degree have shown that the mass size has affected the correlation coefficients directly for this simulation. The mass size of 10 have resulted best performance for polynomial degree of 2 and 3. The program for polynomial kernel function of degree 4 has not been able to separate the nonlinear mixture of images.

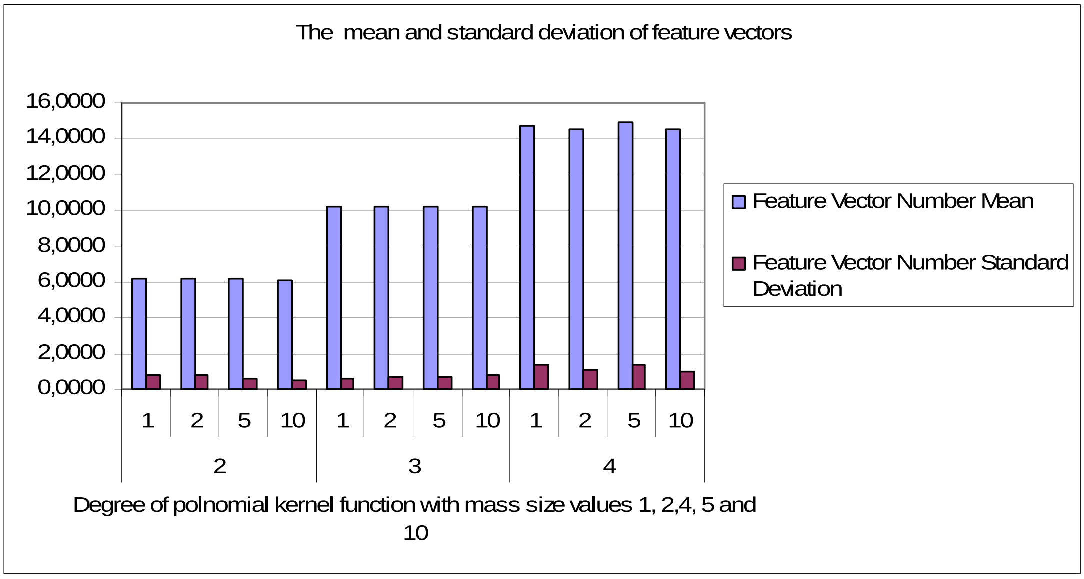

Figure 5.36 The mean and standard deviation of feature vectors

The mean number of feature vectors has been dependent on the degree of the polynomial kernel. As the kernel degree has increased, the number of feature vector has increased.

CHAPTER SIX
CONCLUSIONS

6.1 Conclusions

In this thesis, the aim was to investigate the performance of the linear and nonlinear blind source separation using the linear and nonlinear temporal predictability methods, respectively.

Linear BSS of sound data types:

The results for linear BSS of sound can be summarized as follows: FastICA method has been simulated for the linear BSS and nonlinear BSS. FastICA has only been able to separate the linear mixtures of source signals. The type of FastICA nonlinearity function has affected the performance of the simulation by depending on the sound data type. Although there is not an apparent difference between the results for the speech and song signals, the performance for the musical notes has been affected by the selection the nonlinearity of the FastICA. FastICA could recover the source signals for only using "skew" nonlinearity function.

The temporal predictability method has been simulated for the linear mixtures of speech, musical note and song signals. The algorithm has demonstrated a similar performance for all data types. The mean of correlation coefficients between source and recovered signals were adequately approximate to unity and the standard deviation values were about to 10^{-13} showing a successful separation.

The kernelized temporal predictability method has been used to solve the nonlinear BSS of sound data mixtures. For nonlinear BSS sound data applications, two types of nonlinear mixing functions that were square root and feedforward multilayer perceptron have been used. The two kernel functions have been used for the kernelized temporal method to solve the nonlinear BSS for both nonlinear mixing types and the results of the simulations have been observed for each combination. The kernelized temporal predictability method has been able to recover the source signals from the nonlinear mixture of square root with using the polynomial kernel function for all sound data types. One point that must be indicated is that the lowest performance has been observed for the musical note signals. This can be caused by the characteristics of the power spectrum. The musical note signals have low frequency nature with a few harmonics and the signals which are separated have similar frequency content. However, in the speech and song cases the spectrums of the signals are broader and are not too similar to each other.

The degree of polynomial kernel function has affected the number of feature vectors extracted. As the degree of kernel has increased, then the number of feature vectors has increased. The performance of the simulations has decreased as the number of feature vectors has increased depending on the kernel degree. In general the best result has been obtained for the degree 2. This result has been expected since the kernel function has the reverse effect of the square root kernel.

The method has not worked for sigmoid kernel function and for the nonlinear mixture of feedforward multilayer perceptron function showing the importance of the kernel selection for the given nonlinearity type.

The BSS algorithms which are applied to sound data mixtures have also been extended to separate image mixtures as a contribution of this thesis. The results are summarized briefly:

FastICA and the temporal predictability method have been used to recover images from the linear mixtures as in the case for sound data mixtures.

Two sets of images have been as source images. FastICA has been able to recover the source images from linear mixtures, successfully. For the first image set, the nonlinearity function of FastICA has not affected the results. But for the second image set, the results have changed with the nonlinearity function a little.

The temporal predictability method has been simulated for the different values of the mass size. Two sets of images have been used to simulate to observe the data dependency on the performance. The mass size has explicitly affected the performance of the simulation for the first image set. But for the second image set the mass size has not affected the performance. From this point, it can be concluded that the image type exactly affects the performance of the solution.

Three sets of images have been used to nonlinear BSS simulations. The kernelized temporal predictability method has been simulated for the different degree of polynomial kernel functions and mass size values. The value of mass size has not affected the performance for two images for polynomial kernel of degree 2. But as the degree of the kernel has increased, the performance has been decreased and the mass size effect on the performance has been more visible.

6.2 Advantages and Disadvantages

The temporal predictability method has given good results for linear BSS of sound and image sources. But for the nonlinear BSS, the kernelized temporal predictability method has only resulted well for polynomial kernel function. The performance of the other kernel functions has not noteworthy results. For image applications the value image size has affected the performance depending on the data type, so it has not seemed stable for data independency.

6.3 Future Directions

The application of kernelized temporal predictability method can be extended by investigating the sound specific parameters and using them as application constraints. The RF data like modulated signals can be separated for digital modulation techniques as receiver signal processing. The image applications can be worked in detail for OCR applications.

REFERENCES

Almeida, M.S.C. & Almeida, L.B. (2004). Separating nonlinear image mixtures using a physical model trained with ICA. *Institute of Telecommunications,* Lisbon, Portugal.

Aizerman M., Braverman, E. & Rozonoer, L. (1964. Theoretical foundations of the potential function method in pattern recognition learning. *Automation and Remote Control* **25**: 821-837.

Barlett, M.S., Movellan, J.R., Sejnowski, (November 2002). Face recognition by independent component analysis*. IEEE Transactions on Neural Networks,* Vol 13, No 6.

Cardoso, J.F. (October 1998). Blnd signal separation: statistical principles. Proceedings of the IEEE, Vo. 86, No. 10.

Choi, S., Cichocki, A., Park, H.M., Lee, S.Y. (January 2005). Blind source separation and independent component analysis: a review. *Neural Information Procesing-Letters and Reviews,* Vol.6, No.1.

Comon, P.(1994).Independent component analysis, A new concept?. *Special Issue on High-Order Statistics.*

Harmelling, S., Ziehe, A., Kawanae, M. An Muller, K.-R. (2002), Kernel fature spaces and nonlinear blind source separation*. In advances in Neural Information Processing Systems*, vol-14, Cambridge, MA, MIT Press.

Hyvarinen, A. (1999), Survey on independent component analysis. *Neural Computing Surveys 2,* 94–128, Helsinki University o Technology, Finland.

Hyvarinen, A., Karhunen, J., Oja, E., *Independent Components Analysis*. John Wiley&Sons.

Kasturi, R., Bow, S.T., El-Masri, W., Shah, J., Gattiker, Q. & Mokate, U.B. (October 1990). A system for interpretations of line drawings. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* Vol. 12, No. 10.

Liang, S., Ahmadi, M. & Shridbar M. (August 1997). Segmentation of handwrtten intererence Marks using the directional stroke planes and reformalized morpological appoach. *IEEE Transactions on Image Processing,* Vol.6, No.8.

Li, Y. Wen, P. & Powers, D. (December 11-17, 2003). Methods fort he blind signal separation problem. *IEEE Int. Conf. Neural Networks & Signal Processing,* Nanjing China.

Martinez, D., Bray, A. (2003). Nonlinear blind source separation using kernels. IEEE Transactions on Neural Networks Vol 14. No 1.

Müler, K.R. Mika, S. Rätsch, G. Tsuda, K. & Schölkopf, B. (March, 2001). An Introduction to Kernel-Based Learning Algorithms. IEEE Transactions on. Neural Networks, Vo. 12, No.2.

Para, L. (February, 2002). Tutorial on blind source separation and independent component analysis. *Adaptive Image& Signal Processing Group,* Sarnoff Corparation.

Stone, J.V. (2001), Blind source separation using temporal predictability. Neural Computation 13, 1559-1574, MIT.

Tonazzini, A., Gerace, I. & Cricco, F. (2004). Joint blind separation and restrotion of mixed degraded images for document analysis. *International conference on Image Processi*ng (ICIP).