Full Length Research Paper

Blind recognition of mixed and disturbed correlated ancient texts

Shahrouz Posht-Panah and Masoudreza Aghabozorgi Sahaf*

Signal Processing Research Group, Electrical and Computer Engineering Department, Yazd University
P. O. Box 89195-741, Yazd, Iran.

Accepted 3 August, 2012

Separation and recognition of ancient documents and texts that have been mixed and disturbed over centuries, is an interesting problem in image processing area and it has been investigated by many researchers. In recent years, independent component analysis (ICA) method has been used for solving this problem, but independence of sources is an essential assumption in ICA, whereas in some problems, sources are not independent. So in this paper, we have tried to propose a method for separation and recognition of mixed and disturbed correlated ancient documents and texts utilizing MULTiple Signal Classification (MUSIC) and Estimation of Signal Parameters via Rotational Invariance Technique (ESPRIT) algorithms from blind source separation (BSS) techniques. The good performance of this method has been investigated for real images.

Key words: Recognition of ancient texts, MULTiple Signal Classification (MUSIC) algorithm, Estimation of Signal Parameters via Rotational Invariance Technique (ESPRIT) algorithm, independent component analysis (ICA).

INTRODUCTION

Usually over centuries, some disturbances are created in ancient documents. For instance a text which has been written on a page several centuries ago is distributed by seeping ink from other side of page or scattering ink through the page, or losing the quality of the sheet and converting from whiteness to darkness. Also when the sheets are thin, and there are texts in two sides of them, sometimes these texts appear on each other and that cause disturbances on original texts. So far various methods and techniques have been investigated for solving these problems (Tonazzini and Bedini, 2000; Govindaraju and Srihari, 1991; Avi-Itzhak et al., 1995; Leedham et al., 2002; Tan et al., 2002). Recently, after introduction of blind source separation (BSS) techniques (Comon et al., 1991; Choi et al., 2005), image recognition and separation has been done by some methods based on these techniques where independent component analysis (ICA) is the most well-known and important of them (Comon, 1994; Hyvarinen et al., 2001). These methods have shown their superiority and better performance with respect to other methods by attention to assumptions and circumstances of problems (Tonazzini et al., 2003, 2004a, b, 2007).

BSS consists of separation, recognition and estimation of several sources that have been mixed in an unknown channel (mixing system) and only some combinations of them have been observed. The problem is called "blind" because no information is available about mixture. So far, the problem of BSS has been solved using various algorithms. The lack of prior information must be compensated by considering some special assumptions on sources or mixing system. The most popular BSS algorithms consider independence of sources, that is, Independent Component Analysis (ICA). The aim in these algorithms is to achieve a separation process that produces the most independent outputs (Comon, 1994; Hyvarinen et al., 2001; Hyvarinen and Oja, 2000). In another class of BSS problems, mixer system and receivers have a special structure which in using it sources estimation is achieved. One of the most important of these algorithms is MULTiple Signal Classification (MUSIC) algorithm that has been applied in array signal processing (Schmidt, 1986; Stoica and Nehorai, 1990). Extension of this algorithm for correlation sources is Estimation of Signal Parameters via Rotational...
Invariance Technique (ESPRIT) (Roy and Kailath, 1986; Ottersten et al., 1991). In the problem of this paper, experiments and their results show that correlation factor between sources has effect on the separation performance. Therefore, ICA algorithms do not have proper performance for mixed images which have correlation with each other, and this is revealed completely in simulation results.

In this paper the aim is to propose a solution for the problem based on MUSIC and ESPRIT algorithms for separating correlated mixed and disturbed ancient images.

This paper is organized as follows: Firstly, MUSIC and ESPRIT algorithms are briefly introduced; thereafter the proposed methodology is presented. This is then followed by the simulation results and concluding remarks.

MUSIC AND ESPRIT ALGORITHMS

MUSIC and ESPRIT algorithms belong to BSS class which operates based on special structure for mixing system, and therefore estimation and separation procedures achieve based on arrival angle of sources and obtaining mixture matrix.

MUSIC algorithm works based on assumption that there are \( m \) flat receivers (in the same plate) and they receive signals from \( d \) sources; hence combined output signal of \( k \)th receiver is as shown in the following relation (considering narrow band assumption):

\[
x_k(t) = \sum_{i=1}^{d} a_k(\theta_i) s_i [i - \tau_k(\theta_i)] \\
= \sum_{i=1}^{d} a_k(\theta_i) s_i(t) e^{-j \omega_o \tau_k(\theta_i)}
\]

where \( \tau_k(\theta_i) \) is delay between \( i \)th source and \( k \)th receiver, that is, \( i \)th input arrives on receiver array with angle \( \theta_i \). \( a_k(\theta_i) \) is corresponding to reaction (amplitude and phase) of \( k \)th elements of receivers in \( \omega_o \) (\( \omega_o \) is central frequency). Using vector form for outputs of \( m \) receivers, the model will take the following form:

\[
X(t) = \begin{bmatrix} a(\theta_1) s_1(t) \\ a(\theta_2) s_2(t) \\ \vdots \\ a(\theta_m) s_m(t) \end{bmatrix} = A S(t)
\]

\[
a(\theta) = \begin{bmatrix} a(\theta_1) e^{-j \omega_o \tau_1(\theta_1)} \\ a(\theta_2) e^{-j \omega_o \tau_2(\theta_2)} \\ \vdots \\ a(\theta_m) e^{-j \omega_o \tau_m(\theta_m)} \end{bmatrix}^T
\]

Now, by considering \( A(\theta) = [a(\theta_1), ..., a(\theta_m)] \) as channel matrix (mixer system), \( S(t) = [s_1(t), ..., s_d(t)]^T \) as sources vector and \( X(t) \) as observations vector, the model (2) is rewritten as:

\[
X(t) = A(\theta) S(t)
\]

where it coincides with linear model of BSS problem but parametric structure based on \( \theta \). ESPRIT algorithm is the same as MUSIC algorithm; though the receivers have symmetric couple form such that if \( r_1, r_2, r_3 \) are three receivers then \( r'_1, r'_2, r'_3 \) are symmetry of receivers with constant distance \( \Delta \). So total array can be considered as two sub-array \( Z_X, Z_Y \) that have constant distance \( \Delta \). Therefore received signals in \( k \)th couple of symmetric receivers are:

\[
x_k(t) = \sum_{i=1}^{d} s_i(t) a_k(\theta_i) \\
y_k(t) = \sum_{i=1}^{d} s_i(t) e^{j \omega_o \Delta \sin \frac{\theta_i}{c}} a_k(\theta_i)
\]

that can be written as vector form for any sub-array:

\[
X(t) = A S(t) \\
y(t) = A \Phi S(t)
\]

where \( S(t) \) is sources vector \( (d \times 1) \), \( X(t) \) is vector of observed signals in \( Z_X \) sub-array receivers and \( y(t) \) is vector of observed signals in \( Z_Y \) sub-array receivers. \( \Phi \) is a diagonal matrix \( (d \times d) \) that consists phase delays between symmetry receivers:

\[
\Phi = \text{diag} \{ e^{j \gamma_1}, ..., e^{j \gamma_d} \}
\]

where \( \gamma_i = \omega_o \Delta \sin \frac{\theta_i}{c} \). Total output vector \( Z(t) \) has following form:

\[
Z(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} = \tilde{A} S(t)
\]

\[
\tilde{A} = \begin{bmatrix} A \\ A \Phi \end{bmatrix}
\]

In ESPRIT, using two symmetric arrays causes ability of separating correlated sources (Roy and Kailath, 1986; Lindgren and Van Der Veen, 1996).

PROBLEM FORMULATION AND SOLUTION

Here, at first the problem, that is, recognition of disturbed and with a special property that mixing matrix \( A(\theta) \) has mixed texts images in consequence passing time, is formulated as a BSS problem and then proposed method for solving the problem is expressed using MUSIC and ESPRIT algorithms.

Generally, we can assume that an ancient text image is a combination of three images: Front of page (original text), back of page (text of back of page) and background of page (distance between back and front of page). In BSS structure these three images mean three different sources that have been combined in
According to an important principle of BSS technique, that is, the equality of the number of observations and sources, we pass the observation from three channels of colors (Red, Blue and Green channels) till we achieve three new observations, therefore we will have the same number of observations as the number of sources. Considering \( S_i, i = 1,2,3 \) as original sources (front, back and background of page) and \( \{x_1, x_2, x_3\} \) as obtained observations by passing of existent observation (image of a text \( x \)) from three channels of main colors (RGB), we will have the model below for the problem:

\[
\begin{bmatrix}
    x_1 \\
    x_2 \\
    x_3
\end{bmatrix} =
\begin{bmatrix}
    \tau_1 & \tau_2 & \tau_3 \\
    g_1 & g_2 & g_3 \\
    b_1 & b_2 & b_3
\end{bmatrix}\begin{bmatrix}
    s_1 \\
    s_2 \\
    s_3
\end{bmatrix}
\]  

(8)

Model (8) is consistent with linear BSS model \( X(t) = A.S(t) \) where matrix

\[
A = \begin{bmatrix}
    \tau_1 & \tau_2 & \tau_3 \\
    g_1 & g_2 & g_3 \\
    b_1 & b_2 & b_3
\end{bmatrix}
\]

shows characteristics of mixer (channel).

As was mentioned in MUSIC and ESPRIT algorithms, estimation of the sources was based on arrival angles of sources to receivers and special parametric structure for channel matrix. However, it is obvious that for the problem in this paper arrival angles of sources are irrelevant and meaningless, therefore in this situation there is a question that "What do arrival angles of sources to receiver mean in separation and recognition of ancient text images and how are they applicable?" First we will discuss this subject for MUSIC algorithm because it has simply conditions and then it will be extended for ESPRIT algorithm.

According to the relations in "MUSIC and ESPRIT algorithms" especially Equation (3) that shows the model of source combination and formation of observations by channel matrix

\[
A(\theta) = [a(\theta_1), \ldots, a(\theta_3)]
\]

(based on \( \theta_j \)), and with attention to the first part of this “Problem formulation and solution” presentation especially Equation (8) that based on it a passed image from channels of three main colors can be expressed as combination of three observations by mixing matrix \( 3 \times 3 \), and for coinciding these two subjects three meaningful angles \( \theta_1, \theta_2, \theta_3 \) must be obtained till the mixer matrix (channel) can be expressed based on them.

Because we can consider that the sources (texts) in this paper are combined together with a shift relation to each other, we interpret these shifts as three angles in the model (Figure 1) where

\( \theta_1 \) is shift between first and second sources, \( \theta_2 \) is shift between second and third sources and \( \theta_3 \) is shift between first and third sources. Therefore we will have a set of parameters \( a_1(\theta_1) \) and \( a_1(\theta_2) \) and \( a_1(\theta_3) \) for Red channel and so for other channels (Green and Blue) such that the total channel can be expressed in the following form:

\[
A(\theta) = \begin{bmatrix}
    a_1(\theta_1) & a_1(\theta_2) & a_1(\theta_3) \\
    a_2(\theta_1) & a_2(\theta_2) & a_2(\theta_3) \\
    a_3(\theta_1) & a_3(\theta_2) & a_3(\theta_3)
\end{bmatrix}
\]

(9)

The problem in ESPRIT algorithm is more complex than MUSIC algorithm because we have to work with three couple receivers; it means that in this algorithm we need six observations. Therefore we must obtain images from front and back of an ancient text so we have two images of it. Then we will pass any image from three channels of main colors; thus from any image will be obtained three observations and consequently we will have six observations that are relevant with the text. So, observations vector is a \( 6 \times 1 \) vector

\[
[z(t)] = [x(t)][v(t)]
\]

mixing matrix is a non-quadrate matrix \( 6 \times 3 \), and we can write channel matrix based on relation (7) as

\[
\bar{A} = [A]\Phi
\]

and by considering:

\[
A = [a(\theta_1) \quad a(\theta_2) \quad a(\theta_3)]
\]

(10)

we can rewrite it as:

\[
\bar{A} = \begin{bmatrix}
    a_1(\theta_1) & a_1(\theta_2) & a_1(\theta_3) \\
    a_2(\theta_1) & a_2(\theta_2) & a_2(\theta_3) \\
    a_3(\theta_1) & a_3(\theta_2) & a_3(\theta_3)
\end{bmatrix}
\]

(11)

where \( a_{ij}, i = 1,2,3 \) are elements of matrix \( \Phi \).

Therefore for the problem, we will achieve desired structure that coincides with ESPRIT algorithm and by solving it based on solution techniques of ESPRIT algorithm, we can separate sources that this separation means restoration of original images from disturbed images.
RESULTS AND DISCUSSION

Here, simulation results of applying proposed method on some disturbed ancient documents and texts images (http://www.columbia.edu/cu/libraries/indiv/rare/image; The National Iranian Library and Central Documents) are investigated. Also the results are compared with the results of applying ICA algorithm for correlated images.

First color image that has been experimented, is a manuscript from the 13th century A.D (Figure 2). Outputs of applying ICA algorithm have been shown in Figure 3. As seen, separation does not have acceptable performance. This image has been experimented using MUSIC algorithm that outputs are in Figure 4a to c. Figure 4d image is back of page, Figure 4e is main text and Figure 4f is background. As can be seen the MUSIC algorithm has done better separation. For simulation of ESPRIT algorithm, we need two symmetric images. In Figure 5a and b are two images of front and back of an ancient Persian manuscript (The National Iranian Library and Central Documents) that has been disturbed by seeping ink from other side of pages especially between line 3 and line 6. The results of applying ESPRIT algorithm are shown in Figure 6a to c. It is seen that separation performance is good and original text has separated obviously.

It must be noticed that blurring in some resulting images is an intrinsic result of applying BSS algorithms on a text image. As mentioned in “Problem Formulation and Solution”, we consider that an ancient text image is combination of three images: front of page, back of page and background of page. So after separation, resulting images are corresponding to these three parts of text image and because one of these images is corresponding to background of page, it will be a blurred image.

Conclusion

In this paper, a method was proposed for separating and recognizing correlated ancient documents and text images that have been mixed together and disturbed over time. This method was based on MUSIC and ESPRIT algorithms from blind source separation techniques. At first, the problem was modeled as a BSS problem and then by introducing an arrival angle interpretation which was needed in MUSIC and ESPRIT structures, these algorithms were applied to the problem. Simulation results showed effectiveness and good performance of the proposed method.
Figure 4. Outputs of applying MUSIC to Figure 2.
Figure 5. Two images of front (a) and back (b) of an ancient Persian manuscript (It is seen that some disturbances has occurred on text by seeping ink from other side of pages especially between lines 3 and 6).
REFERENCES


Figure 6. Outputs of applying proposed ESPRIT method (It is seen that original texts (a), (b) has separated obviously).