<table>
<thead>
<tr>
<th><strong>Title</strong></th>
<th>A matrix-pencil approach to blind separation of non-white sources in white noise</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Author(s)</strong></td>
<td>Chang, C; Ding, Z; Yau, SF; Chan, FHY</td>
</tr>
<tr>
<td><strong>Citation</strong></td>
<td>IEEE International Conference on Acoustics, Speech and Signal Processing Proceedings, Seattle, WA, USA, 12-15 May 1998, v. 4, p. 2485-2488</td>
</tr>
<tr>
<td><strong>Issued Date</strong></td>
<td>1998</td>
</tr>
<tr>
<td><strong>URL</strong></td>
<td><a href="http://hdl.handle.net/10722/46052">http://hdl.handle.net/10722/46052</a></td>
</tr>
<tr>
<td><strong>Rights</strong></td>
<td>This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License.; ©1998 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.</td>
</tr>
</tbody>
</table>
ABSTRACT

The problem of blind source separation in additive white noise is an important problem in speech, array and acoustic signal processing. In general this problem requires the use of higher order statistics of the received signals. Nonetheless, many signal sources such as speech with distinct, non-white power spectral densities, second order statistics of the received signal mixture can be exploited for signal separation. While previous approaches often assume that additive noise is absent or that the noise correlation matrix is known, we propose a simple and yet effective signal extraction method for signal source separation under unknown white noise. This new and unbiased signal extractor is derived from the matrix pencil formed between output auto-correlation matrices at different delays. Simulation examples are presented.

2. PROBLEM FORMULATION

A memoryless mixture of multiple signals is modeled by the following equation

\[ \tilde{x}(n) = A \tilde{s}(n) + \tilde{w}(n), \]

where \( \tilde{x}(n) \) is a vector of source signals of dimension \( N \) and \( \tilde{s}(n) \) is the received signal vector of dimension \( M \). In (1), \( \tilde{w}(n) \) is the additive white noise vector and \( A \) is a \( M \times N \) memoryless mixing matrix.

In order for all sources to be separable, \( A \) must have full column rank, requiring \( M \geq N \). When \( A \) is not full rank, then signals can only be separated as classes \([a] \). In this paper, we only consider systems in which the sources are fully separable.

Our objective is to find a signal extracting matrix

\[ B = [\tilde{b}_1 \quad \tilde{b}_2 \quad \ldots \quad \tilde{b}_N] \]

such that

\[ B^T \tilde{x}(n) = B^T A \tilde{s}(n) + B^T \tilde{w}(n) = E \tilde{s}(n) + B^T \tilde{w}(n) \]

where

\[ E = B^T A \]

is a permutation matrix which only has one nonzero element in each row and column.

We assume that the source signals

\[ \tilde{s}(n) = [s_1(n) \quad s_2(n) \quad \ldots \quad s_N(n)]^T \]

are uncorrelated of one another but are not white. Hence,

\[ R_{s_i}(k) \triangleq E\{s_i(n)s_j(n-k)^T\} = 0, \quad \text{for all } k \]

and

\[ R_s(k) \triangleq E\{\tilde{s}(n)\tilde{s}(n-k)^T\} \neq 0, \quad \text{for some } k \neq 0. \]
3. SIGNAL SEPARABILITY BASED ON SECOND ORDER STATISTICS

Let \( r_i(0) = E\{s_i(n)\} \). When all signals are non-zero, then denote

\[
\Lambda(k) = R_s(k) = \text{diag}(r_1(k), r_2(k), \ldots, r_N(k))
\]

is full rank. Without loss of generality, we can assume that \( R_s(0) = I \).

Given white noise with known correlation matrix \( \Lambda_n \), we have

\[
R_s(0) = A R_s(0) A^T + \Lambda_n = A A^T + \Lambda_n.
\]

Thus, we can use a whitening matrix \( W \) such that

\[
I = W(R_s(0) - \Lambda_n)W^T = W A A^T W^T,
\]

which implies that \( U = WA \) is unitary.

We note that the assumption of known noise correlation matrix is typically impractical. In the special case when the noise correlation matrix is of the form \( \sigma^2 I \) and \( M > N \), then \( \sigma^2 \) can be estimated as the smallest eigen-value of \( R_s(0) \).

Denote the whitened data vector as \( \tilde{z}(n) = W z(n) \), with auto-correlation \( R_s(k) = E\{\tilde{z}(n)\tilde{z}^T(n-k)\} \). Then we have

\[
R_s(k) = W A R_s(k) A^T W^T = U R_s(k) U^T. \tag{10}
\]

Hence columns of \( U \) are eigen-vectors of \( R_s(k) \).

If \( R_s(k) \) has identical eigen-values, then their eigen-vectors are not unique. Hence we cannot always find the eigen-vector of \( R_s(k) \) as solutions. However, we can separate groups of signals by exploiting eigen-vectors corresponding to different eigen-values. In a total blind situation, let \( R_s(k) \) to have only \( L \) distinct eigen-values, then we can have \( L \) subclasses of signals that are orthogonal and hence \( L \) subclasses of signals can be extracted by

\[
U_i^T \tilde{z}(n), \quad i = 1, 2, \ldots, L. \tag{11}
\]

The dimension of \( U_i \) is equal to the multiplicity of the corresponding eigen-value.

Once we extract a new signal vector with lower dimension

\[
\tilde{z}_i(n) = U_i^T \tilde{z}(n), \tag{12}
\]

its components can be extracted based on \( R_s(k) \) using identical steps. To ensure that this procedure can eventually extract all the sources, it is necessary that the sources have different power spectral densities. In [6], an algorithm which is exponentially convergent is proposed.

The algorithm discussed above assumes that the correlation matrix of the noise is either known or can be estimated when \( M > N \). However, when the noise correlation matrix is either unknown or cannot be estimated, the methods will not apply. In view of this difficulty, we present in the next section an alternative algorithm which is insensitive to additive white noise.

4. A NEW MATRIX PENCIL SEPARATOR

We now present a method that does not rely on the information in \( R_s(0) \) which is corrupted by white noise. Instead, we shall use information contained in

\[
R_s(k) = A \Lambda(k) A^T, \quad k \neq 0
\]

that is insensitive to additive white noises.

The algorithm is based on the following generalized eigen-value problem. If we solve for

\[
R_s(k_1) \tilde{v} = \lambda R_s(k_2) \tilde{v}, \tag{13}
\]

effectively the equation can be rewritten as

\[
A [\Lambda(k_1) - \lambda \Lambda(k_2)] \tilde{v} = 0. \tag{14}
\]

One trivial solution to the eigen-vector problem is

\[
\tilde{v}_0 = 0, \tag{15}
\]

Note that the trivial solution will exist if \( M > N \). Since

\[
\tilde{v}_0^T \tilde{z}(n) = \tilde{v}_0^T A \tilde{z}(n) + \tilde{v}_0^T \tilde{w}(n) = \tilde{v}_0^T \tilde{w}(n), \tag{16}
\]

\( \tilde{v}_0 \) annihilates all signals in \( \tilde{z}(n) \). In fact, if the noise correlation matrix is \( \sigma^2 I \), the trivial solution \( \tilde{v}_0 \) can be used to estimate \( \sigma^2 \) since

\[
\tilde{v}_0^T R_s(k_0) \tilde{v}_0 = \sigma^2. \tag{17}
\]

Since \( A \) has full column rank, any other non-trivial solution requires eigen-value to satisfy

\[
r_i(k_1) - \lambda r_i(k_2) = 0 \quad \text{or} \quad \lambda_i = \frac{r_i(k_1)}{r_i(k_2)}, \tag{18}
\]

whose eigen-vector must satisfy

\[
A^T \tilde{v}_i = \alpha_i \tilde{e}_i. \tag{19}
\]

Therefore, if the ratio \( \frac{r_i(k_1)}{r_i(k_2)} \) is unique, then \( \tilde{v}_i \) is simply the \( i \)-th signal extractor to yield

\[
\tilde{v}_i^T \tilde{z}(n) = \tilde{v}_i^T A \tilde{z}(n) + \tilde{v}_i^T \tilde{w}(n) = \alpha_i s_i(n) + \tilde{v}_i^T \tilde{w}(n). \tag{20}
\]

In general, if the ratio \( \frac{r_i(k_1)}{r_i(k_2)} \) is not unique and \( \frac{r_i(k_1)}{r_i(k_2)} = \gamma, \quad \forall i \in I = \{i_1, i_2, \ldots, i_m\} \), then let \( V_i \) consists of all eigen-vectors for eigen-value \( \gamma \)

\[
V_i = [\tilde{v}_{i_1}, \tilde{v}_{i_2}, \ldots, \tilde{v}_{i_m}], \tag{22}
\]

We have

\[
A^T V_i = [\tilde{e}_{i_1}, \tilde{e}_{i_2}, \ldots, \tilde{e}_{i_m}] U_i \tag{23}
\]

in which \( U_i \) is an \( i_m \times i_m \) unitary matrix. Hence \( V_i \) can be used to extract the set of signals \{\( s_{i_k} \), \( i \in I \)\}

\[
\tilde{z}(n) \triangleq V_i^T \tilde{z}(n) = V_i^T A \tilde{z}(n) + V_i^T \tilde{w}(n) = U_i^T [s_{i_1}(n) \ s_{i_2}(n) \ \ldots \ s_{i_m}(n)]^T + V_i^T \tilde{z}(n). \tag{24}
\]

Signals within the set \( I \) can be extracted from \( \tilde{z}_i \) by applying the same principle on its auto-correlation matrices

\[
R_s(k) \triangleq V_i^T R_s(k) V_i = U_i^T A(k) U_i, \tag{25}
\]

for a pair of different delays, \( k_3 \) and \( k_4 \). Alternatively, since \( U_i \) is unitary, its estimate can be obtained directly as the eigen-vectors of \( R_s(k) \).

Here we note that the trivial solutions of (15) are not separable from the other general eigen vectors of \( (R_s(k_1), R_s(k_2)) \). However, this is not a problem. The extracted signals according to the trivial solutions are pure noises result in the smallest output power while signal extractors will produce stronger outputs. Thus we can always discard the trivial solution as the \( M - N \) vectors that result in the smallest product \( \tilde{v}_0^T R_s(0) \tilde{v}_0 \).
5. SIMULATION

Before numerical simulations are presented. It is helpful to analyze the best possible results one can obtain in signal separation when \( A \) is known a priori. Defining the pseudoinverse of \( A \) as \( A^\dagger \), the signal estimate should be

\[
\hat{s}(n) = A^\dagger \hat{x}(n) = \hat{s}(n) + A^\dagger \hat{e}(n).
\]

This represents the best possible result when noise is unknown. Experimental results will be compared with the ideal situation.

When \( A \) is unknown, blind separation performance will naturally be inferior due to two reasons:

- The data length may be too short to accurately estimate the correlation matrices \( R_k(\ell) \).
- The underlying sources may not be fully uncorrelated with one another. This violation of our basic assumptions will also degrade the performance.

In the simulation, we consider a system with four sources and five outputs. Two sets of signals are tested. The first set of signals are four random color Gaussian signals. Each signal is generated by filtering a white Gaussian noise using a 20-tap FIR filter with impulse response

\[
h(n, \sigma) = \exp\left(-\frac{n+1}{10\sigma}\right) \quad n = 0, 1, \cdots, 19
\]

where \( \sigma = 1, 0.4, 0.3, 0.2 \) respectively. The second set of signals are four speech signals sampled at 5500Hz. The mixing matrix used in both examples is

\[
A = \begin{bmatrix} 0.2380 & 0.2887 & -0.7120 & 0.4914 \\ 0.3397 & -0.7494 & -0.1157 & 0.2097 \\ 0.6107 & 0.4959 & 0.2661 & 0.2504 \\ 0.3558 & 0.2644 & -0.4216 & -0.6640 \\ -0.5731 & -0.1983 & -0.4807 & 0.4593 \end{bmatrix}.
\]

The columns of \( A \) are normalized so that the total power of each source in the mixtures is equal. Simulations were conducted for 3 different signal lengths of 1000, 4000, and 16000, with individual input signal to noise ratio (SNR) at \(-10, 0, 10, 30\), and \(40\)dB. For each situation, we run 100 Monte Carlo simulations to get the statistical performance.

The individual input SNR is now identical for all sources since we normalized columns of \( A \). This allows a fair comparison. To measure the statistical performance of our method, we use the output signal to interference and noise ratio (SINR) as the performance index. The delay pair used in the matrix pencil method is \((k_1 = 1, k_2 = 3)\). The output SINR according to the individual input SNR are given in Figure (1) and Figure (2) for the two examples of signal separation, respectively. The output SINR is averaged over 100 Monte Carlo simulations. Let \( L \) be the length of data samples. In each figure, circle 'o' denotes cases for \( L = 1000 \), x-mark 'x' for \( L = 4000 \), and plus '+' for \( L = 16000 \). Finally to compare with the upper bound of output SINR, solid lines are given for ideal cases in which the mixing matrix \( A \) is known.

It can be seen that the matrix pencil method works well for real world speech signals. Comparing results from different data lengths, we find that the performance at \( L=1000 \) is very close to what is achieved at length \( L=16000 \). This result demonstrates that the proposed method converges very fast since the estimation of second order statistics does not require many data samples.

For the specific examples, the output SINR saturates at about 20dB. This is because in high SNR case, the performance is dominated by the level of cross interference which is determined by the cross correlation between sources and the mixing matrix. Clearly for the same mixing matrix \( A \), different signal sources with different cross correlation also result in different output SINR.

6. CONCLUSION

In this paper we present a new second order statistical method for blind source separation. It is based on the general eigen decomposition of the matrix pencil formed between output auto-correlation matrices at different delays. The method has several attractive features compared with higher order statistical methods. First, it relies only on second order statistics of the received signals and hence requires fewer data samples to converge. Second, it can separate Gaussian sources. Compared with other known methods based on second order statistics, it does not require zero additive noise or that the noise correlation matrix is known, and it is unbiased. Numerical simulations show that the method works well for real world speech signals.

ACKNOWLEDGMENT

The work of Chang and Chan was supported by in part by grants from Research Grants Council of Hong Kong, grant numbers HKU553/96M and HKU7036/97E. The work of Ding was supported in part by the US Army Research Office and by the visiting appointment at HKUST. The work of Yau was supported in part by grants from Research Grants Council of Hong Kong, grant number HKUST776/96E.

REFERENCES

Figure 1. Performance for randomly generated sources

Figure 2. Performance for real world speech sources