HIGH QUALITY SIGNAL RECEPTION IN THE PRESENCE OF STATIONARY INTERFERENCE - A BLIND SIGNAL SEPARATION APPROACH

Wai Kuen Lai, Ting Wai Siu and Sze Fong Yau

Technical Services Division, The Government of Hong Kong SAR
27/F, Fairmont House, 8 Cotton Tree Drive, Hong Kong
The People's Republic of China

ABSTRACT

This paper considers the problem of interference rejection using sensor array with application to enhance signal reception. In contrast to conventional adaptive beamforming which requires knowledge of the array geometry and array response, we propose a two stage blind signal separation approach to achieve interference rejection. The proposed method is based on the practically viable assumptions that the desired signal is temporally non-stationary and the interference are temporally second-order stationary, and is applicable to arrays with uncalibrated response and geometry. Real-world experimental results are presented to demonstrate the performance of the proposed method.

1. INTRODUCTION

The problem of signal reception in the presence of directional interference arises in a variety of applications such as teleconferencing and hearing aids in sound reception. To enhance the quality of signal reception, sensor array has become the most commonly adopted strategy. Typically, a sensor array is used to form a directional beam pattern which can be steered to null out signals arriving from directions other than a direction of interest [1]. This is the familiar *adaptive beamforming* technique with its origin derived from applications in radar, sonar and communication systems [2, 3].

The effectiveness of a sensor array in rejecting the undesirable interference hinges on the accurate localization of the direction of the desirable source. Conventionally, the time delay (TD) of arrival estimates between direct signals received by sensors in the array are used to localize the desirable source. Commonly used estimation methods of TD are based on cross-correlation peak [4], cross-correlation of prewhitened signals [5] and cross power spectrum phase [6]. However, these methods require precise knowledge of the array configuration and often ignore additional complications such as multipaths.

In many practical situations, the interference are persistently stationary signals. A good example is acoustic noises from alternating current machines such as cooling fans or air compressors. On the other hand, the desired signal, for example speech signals, are rich in information and content, and hence are usually non-stationary. Based on this observation, we propose in this paper a subspace-based approach to separate, from the sensor array data, the desirable non-stationary signals from the undesirable stationary interferences, hence achieving high quality signal reception. The proposed method does not require knowledge of the signal array configuration and can accomodate possible multipaths.

The paper is organised as follows. In Section 2, we model the sensor array as a multi-input multi-output linear time-invariant system and describe some underlying assumptions. In Section 3, we present an algorithm to separate the interferences from the desired signal. In Section 4, an adaptive cancellation algorithm is described in which interferences are cancelled from the signals received by sensors. Real-world experiemental results are presented in Section 5. Section 6 gives the conclusion.

2. PROBLEM FORMULATION AND ASSUMPTIONS

We model a sensor array with M elements as a LTI discrete time system with M output signals $y_1[n], y_2[n], \ldots, y_M[n]$. We further assume that there are N input signals (sources) $a_1[n]$, $a_2[n]$, $\ldots$, $a_N[n]$ to this system. The relationship between the input and output signals can be represented by the equation

$$\boldsymbol{y}[n] = \sum_r \boldsymbol{h}[r]\boldsymbol{a}[n-r] \tag{1}$$

where the N channel inputs form the input vector

$$\boldsymbol{a}[n] = \begin{bmatrix} a_1[n] & a_2[n] & \ldots & a_N[n] \end{bmatrix}^T, \tag{2}$$

and the M channel outputs form the output vector

$$\boldsymbol{y}[n] = \begin{bmatrix} y_1[n] & y_2[n] & \ldots & y_M[n] \end{bmatrix}^T. \tag{3}$$

In the above equation, the channel impulse response of the MIMO system is modeled as

$$\boldsymbol{h}[r] = \begin{bmatrix} h_{11}[r] & h_{12}[r] & \dots & h_{1N}[r] \\ h_{21}[r] & h_{22}[r] & \dots & h_{2N}[r] \\ \vdots & \vdots & \ddots & \vdots \\ h_{M1}[r] & h_{M2}[r] & \dots & h_{MN}[r] \end{bmatrix}, \quad (4)$$

with entries $h_{ij}[r], i = 1, 2, \dots M, j = 1, 2, \dots N$. We assume that the impulse responses $h_{ij}[r]$ are FIR filters with tap length P, possibly non-minimum phase. In addition, we assume that the MIMO system $\boldsymbol{h}[r]$ is causally invertible by an FIR MIMO deconvolution filter.

Without loss of generality, we assume that the desired signal is $a_1[n]$, whereas the remaining signals, $a_i[n]$, $i = 2, \cdots, N$ are the interferences to be removed. In our discussion, we assume that the input signals $a_i[n]$ are mutually uncorrelated with possibly non-white power spectra. As a salient point of this paper, we assume that the signal $a_1[n]$ is a non-stationary process whereas the remaining signals are second-order stationary processes. To justify the assumption that $a_1[n]$ is non-stationary, we argue that the desired signal is usually rich in information and content, thus cannot be stationary. The assumption on the interference applies in many practical situations where these interferences are caused by machine or thermal noise.

To overcome the ambiguity that is inherent to the extraction of $a_1[n]$ [7], we further assume that $h_{11}[n] = \delta[n]$. Our goal is to design a linear MIMO deconvolution (filter) system with M inputs and single output such that for input $\boldsymbol{y}[n]$, its output signal is (possibly delayed and scaled version of) $a_1[n]$.

3. A SUBSPACE METHOD FOR SOURCE SEPARATION

Our algorithm will be divided into two stages. The first stage focuses on signal separation and the second deals with single signal deconvolution.

The relationship of $y_i[n]$ and $a_i[n]$ can be written in terms of multiple samples. Let

$$\boldsymbol{y}_i^{(n)} = \begin{bmatrix} y_i[n] & \dots & y_i[n-Q+1] \end{bmatrix}^T,$$

$$\boldsymbol{a}_j^{(n)} = \begin{bmatrix} a_j[n] & \dots & a_j[n-P-Q+1] \end{bmatrix}^T,$$

and

$$\mathcal{H}_{ij} = \begin{bmatrix} h_{ij}[0] & \dots & h_{ij}[P-1] & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & h_{ij}[0] & \dots & h_{ij}[P-1] \end{bmatrix},$$

then

$$\boldsymbol{y}_i^{(n)} = \sum_{j=1}^{N} \mathcal{H}_{ij} \boldsymbol{a}_j^{(n)}. \quad (5)$$

If we write

$$\boldsymbol{y}^{(n)} = \begin{bmatrix} \boldsymbol{y}_1^{(n)T} & \dots & \boldsymbol{y}_M^{(n)T} \end{bmatrix}^T,$$

$$\boldsymbol{a}^{(n)} = \begin{bmatrix} \boldsymbol{a}_1^{(n)T} & \dots & \boldsymbol{a}_N^{(n)T} \end{bmatrix}^T,$$

$$\mathcal{H} = \begin{bmatrix} \mathcal{H}_1 & \mathcal{H}_2 & \dots & \mathcal{H}_N \end{bmatrix}$$

and

$$\mathcal{H}_j = \begin{bmatrix} \mathcal{H}_{1j}^T & \mathcal{H}_{2j}^T & \dots & \mathcal{H}_{Mj}^T \end{bmatrix}^T,$$

then

$$\boldsymbol{y}^{(n)} = \mathcal{H}\boldsymbol{a}^{(n)} = \sum_{j=1}^{N} \mathcal{H}_j \boldsymbol{a}_j^{(n)}. \quad (6)$$

We assume that the system $h[r]$ is such that nonzero columns of $\mathcal{H}$ have full rank. This condition implies that there can be all zero columns in the matrix $\mathcal{H}$ because of difference in the delay spread of different channel responses.

We now form the autocorrelation matrix of $a_i[n]$ and $y_i[n]$. Since $a_i[n]$ are uncorrelated to $a_j[n]$, for $j \neq i$, $E\{\boldsymbol{a}^{(n)}\boldsymbol{a}^{(n-k)^T}\}$ is a block diagonal matrix with blocks $\boldsymbol{R}_{ii}[n,k] = E\{\boldsymbol{a}_i^{(n)}\boldsymbol{a}_i^{(n-k)^T}\}$, $i = 1, \dots N$, on its diagonal. We denote

$$\begin{aligned} \boldsymbol{P}[n,k] &= E\{\boldsymbol{y}^{(n)}\boldsymbol{y}^{(n-k)^T}\} & (7) \\ &= \mathcal{H}E\{\boldsymbol{a}^{(n)}\boldsymbol{a}^{(n-k)^T}\}\mathcal{H}^T & (8) \\ &= \mathcal{H}\text{diag}\{\boldsymbol{R}_{ii}[n,k]\}\mathcal{H}^T & (9) \end{aligned}$$

Note that $\boldsymbol{P}[n,k]$ is a $MQ \times MQ$ matrix.

Let $\boldsymbol{P}[n_1,k]$ and $\boldsymbol{P}[n_2,k]$ be autocorrelation matrices formed by (7) with $n_1 \neq n_2$. Consider the matrix

$$\begin{aligned} &\boldsymbol{Q}[n_1,n_2,k] \\ &\quad = \boldsymbol{P}[n_1,k] - \boldsymbol{P}[n_2,k] & (10) \\ &\quad = \mathcal{H}_1\{\boldsymbol{R}_{ii}[n_1,k] - \boldsymbol{R}_{ii}[n_2,k]\}\mathcal{H}_1^T & (11) \end{aligned}$$

where we have used the fact that for a fixed value of k, $\boldsymbol{R}_{ii}[n,k]$, $i = 2, \cdots, N$, are all independent of n because the corresponding $a_i[n]$'s are second-order stationary.

Since it has been assumed that non-zero columns $\mathcal{H}$ have full column rank, it follows that $\mathcal{H}_1$ is column rank deficient and there exists vectors $\boldsymbol{v}_l$, $l = 1, \cdots, L$ such that

$$\mathcal{H}_1^T \boldsymbol{v}_l = \boldsymbol{0}. \quad (12)$$

The vector $\boldsymbol{v}_l$ can be easily determined by first finding the eigen vectors of $\boldsymbol{Q}[n_1,n_2,k]$ corresponding to the zero eigen values, and then selecting those which are *not* orthogonal to $\boldsymbol{Q}[n_1,n_2,m]$ for $m \neq k$.

The vectors $\boldsymbol{v}_l$, $l = 1, \cdots, L$ are the desired solution of our problem. In fact, $\boldsymbol{v}_l$ can be used to construct a M-input 1-output FIR filter bank. When the input is $\boldsymbol{y}[n]$, the output signal will be a filtered mixture of $\boldsymbol{a}_i[n]$, $i = 2, \cdots, N$, but will *not* contain any part of $\mathbf{a}_1[n]$. More explicitly, for $l = 1, \cdots, L$, we have

$$z_l^{(n)} = v_l^T y^{(n)} \quad (13)$$

$$= \sum_{j=1}^{N} v_l^T \mathcal{H}_j a_j^{(n)} \quad (14)$$

$$= \sum_{j=2}^{N} u_{lj} a_j^{(n)} \quad (15)$$

where $u_{lj} = u_l^T \mathcal{H}_j$ and $u_{l1} = v_l^T \mathcal{H}_1 = \mathbf{0}$. We have thus separated the desired signal from the interference. Unfortunately the signals $z_l^{(n)}$ at hand are mixtures of interference. Thus, a second stage algorithm needs be used to obtain the desired signal. This will be discussed in the next section.

4. INTERFERENCE REJECTION USING REFERENCE SIGNAL

It is a common practice to extract a weak signal corrupted by strong additive noise by means of an adaptive interference canceler (AIC)[8]. The structure of a typical 2-input AIC is shown in Fig. 1. One of input signal, which is known as the primary signal, contains the desired signal corrupted by the interference. The second input signal is correlated to the interference and therefore acts as a reference signal. It is very important that the reference signal is uncorrelated to the desired source; otherwise, *power inversion* occurs in which the output SNR is limited by the reciprocal of the SNR in the reference signal.

From the discussion in Section 3, it can be seen that the signals $z_l^{(n)}$ can be used as multiple references for cancellation of the interference. An AIC structure with single primary source and multiple references is thus proposed in Fig. 2.

In Fig. 2, the reference signals $z_l^{(n)}$ are generated by the mulit-input multi-output filter according to equation (13). Without loss of generality, the signal $y_1[n]$ is choosen as

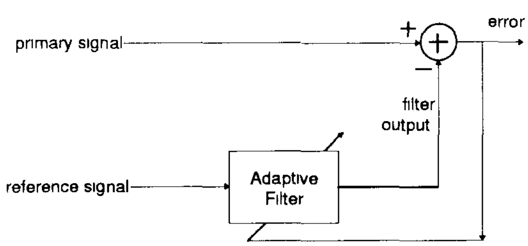

Figure 1: Block diagram for a conventional adaptive interference cancellation

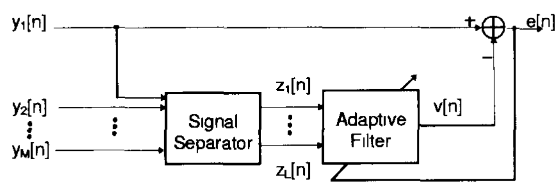

Figure 2: The proposed adaptive interference canceler

the primary signal. Similar to the conventional canceler, the reference signals $z_l[n]$ are processed by a multi-input single-output adaptive filter, whose output $v[n]$ is given by

$$v[n] = \sum_{l=0}^{L-1} \sum_{k=0}^{K-1} \hat{w}_l[k] z_l[n-k] \quad (16)$$

where the coefficients $\hat{w}_l[n]$ are the adjustable tap weights of the adaptive filter and K is the number of taps in the adaptive filter.

The output of the overall AIC is the error signal between the filter output $v[n]$ and $y_1[n]$, which is given by

$$e[n] = y_1[n] - v[n] \quad (17)$$

The system output is used to adjust the filter weights $\hat{w}_l[n]$ of the adaptive filter to minimize the mean-square value of $e[n]$. Under ideally conditions, i.e. perfect annihilation of the desired source $a_1[n]$ in $z_l^{(n)}$ and convergence of the adaptive filter, the AIC output will be

$$e[n] = h_{11}[n] \star a_1[n] = a_1[n], \quad (18)$$

using the assumption that $h_{11}[n] = \delta[n]$.

5. EXPERIMENT RESULT

We have performed experiments to test the proposed method. In our experiments, four audio microphones, arranged in a linear array, were used. The inter-microphone separation was set to 0.2m. The whole array is 1m away from the loud speakers. Two signals, namely, a background noise and a Cantonese speech, were broadcasted via two loud speakers which were 1m apart. The waveform of the original Cantonese speech and the interference are depicted in Fig. 3 respectively.

The output power level of the loud speakers were set so that the desired speech was barely percetible to the listener. The received signal from the microphones are shown in Fig. 4. It is obvious that the desired signal submerged completely in the interference.

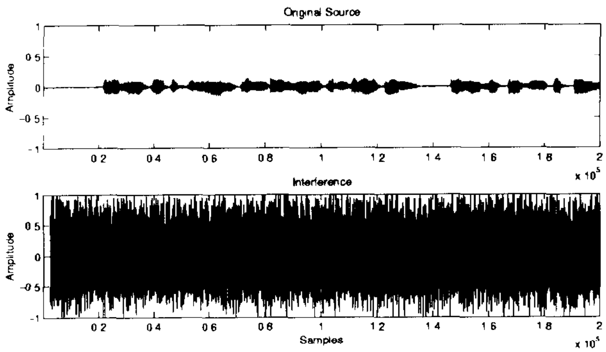

Figure 3: Waveform of the signal sources

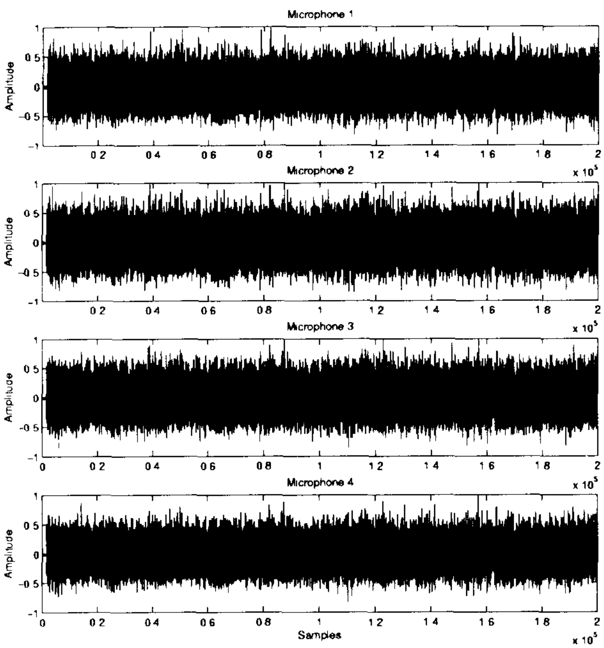

Figure 4: Signal waveform at the microphones

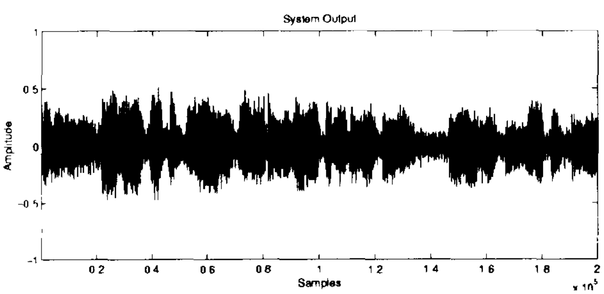

Figure 5: waveform at the system output

We first estimated the reference signal using the method described in Section 3. The reference signal vector was then passed to an adaptive filter whose gain constant was set to 0.03. For simplicity, LMS algorithm was used.

The AIC took about 40000 samples to reach its steady state. The output signal after the filter reached its steady state is shown in Fig. 5. In the output signal, the interference is greatly suppressed which can be seen by the emerging waveform. Subjective test reveals an impressive improvement in the audibility of the desired signal.

6. CONCLUSION

In this paper, we proposed a two-step method for sound reception in the presense of severe stationary interference. The method first estimated a mixture of the interfering signals blindly, which does not require any *a priority* information of the system, such as the array geometry and direction of arrival of the interference. This extracted signal is uncorrelated to the desired signal and is therefore an ideal reference signal for an adaptive interference canceler. Real-world experiments illustrate that the intelligibility of the desired speech can be greatly improved by applying our proposed method.

7. REFERENCES

[1] J. L. Flanagan, J. D. Johnston, R. Zahn and G. W. Elko, "Computer-steered microphone array for sound transduction in large rooms," *J Acoust Soc Am.*, vol. 78, no. 11, pp. 1508-1518, Nov. 1985.

[2] R. Monzingo, *Introduction to Adaptive Arrays*, Wiley, New York, NY, 1980.

[3] M. Soumekh, *Fourier Array Imaging*, Prentice Hall, Englewood Cliffs, NJ, 1994.

[4] H. F. Silverman and S. E. Kirtman, "A two-stage algorithm for determinating talker location from a linear microphone array data," *Computer Speech and Language*, vol. 6, pp. 129-152, 1992.

[5] I. Yamada and N. Hayashi, "Improvement of the performance of cross correlation method for identifying aircraft noise with prewhitening signals, " *J. Acoust. Soc. Japen*, vol. 13, no. 4, pp. 241-252, 1992.

[6] M. Omologo and P. Svaizer, "Acoustic event localization using a cross power spectrum phase based technique," *Proc. Int. Conf. on Acoust., Speech and Signal Proc.*, vol. 2, pp. 273-276, Adelaide, 1994.

[7] D. Chan, *Blind Signal Separation*, Ph.D. Thesis, The University of Cambridge, Jan. 1997.

[8] B. Widrow and S. D. Stearns, *Adaptive Signal Processing*, Pretnice Hall, EngleWood Cliffs, NJ, 1985.