JID: NEUCOM

ARTICLE IN PRESS

Neurocomputing 000 (2017) 1-11



Contents lists available at ScienceDirect

Neurocomputing



NEUROCOM PUTING

journal homepage: www.elsevier.com/locate/neucom

Effective Kalman filtering algorithm for distributed multichannel speech enhancement $\!\!\!\!\!^{\bigstar}$

Jingxian Tu^a, Youshen Xia^{b,*}

^a Laboratory of Complex System Simulation & Intelligent Computing, School of Information and Electronic Engineering, Wuzhou University, Wuzhou, China ^b Department of Software Engineering, College of Mathematics and Computer Science, Fuzhou University, Fuzhou, China

ARTICLE INFO

Article history: Received 11 May 2016 Revised 11 April 2017 Accepted 21 May 2017 Available online xxx

Communicated by Dr. K Chan

Keywords: Colored noise reduction Kalman filtering Multichannel speech enhancement Time domain

ABSTRACT

Kalman filtering is known as an effective speech enhancement technique. Many Kalman filtering algorithms for single channel speech enhancement were developed in past decades. However, the Kalman filtering algorithm for multichannel speech enhancement is very less. This paper proposes a Kalman filtering algorithm for distributed multichannel speech enhancement in the time domain under colored noise environment. Compared with conventional algorithms for distributed multichannel speech enhancement, the proposed algorithm has lower computational complexity and requires less computational resources. Simulation results show that the proposed algorithm is superior to the conventional algorithms for distributed multichannel speech enhancement in achieving higher noise reduction, less signal distortion and more speech intelligibility. Moreover, the proposed algorithm has a faster speed than several multichannel speech enhancement algorithms.

© 2017 Elsevier B.V. All rights reserved.

ment, the multichannel speech enhancement algorithms outperform the single channel speech enhancement algorithms. Multi-

channel speech enhancement algorithms mainly include the clas-

sic beamforming algorithm [22,23], the multichannel wiener al-

gorithm [24,25], the multichannel subspace algorithm [26–28],

the multichannel minimum distortion algorithm [29], and multi-

channel statistical estimation algorithm [30-32]. These multichan-

nel speech enhancement algorithms can reduce background noise

with less speech distortion when increasing the number of micro-

phones. The multichannel wiener algorithm reduces the noise by

assuming the noise is stationary and minimizing the mean square

error [24,25]. It can obtain the good performance in the station-

ary noise case and but the poor performance in the nonstationary

noise case. The multichannel subspace algorithm first decomposes

1. Introduction

Since received signals are usually corrupted by background noise, noise reduction is required in signal processing such as speech communication, speech recognition, speaker identification, etc. Over the past decades, many of speech enhancement algorithms for noise reduction were presented. According to the number of channel available, speech enhancement algorithms can be classified as single channel algorithms [1–21] and multichannel speech enhancement algorithms [22–34]. In term of the microphone configuration, multichannel microphone can be classified as dual, array, and distributed microphones. This paper focuses on distributed multichannel speech enhancement.

Multichannel speech enhancement techniques have been attractively studied as the multi-microphone system was developed. Many multichannel speech enhancement algorithms were presented. Since the multichannel microphone can utilize much more information to improve the performance of speech enhance-

* Corresponding author.

E-mail addresses: 568205127@qq.com (J. Tu), ysxia@fzu.edu.cn, ysxia2001@yahoo.com (Y. Xia).

the noisy space into the pure noise space and speech-plus-noise space
 the noisy space into the pure noise space and speech-plus-noise space
 space, and then reduces the noise in the speech-plus-noise space
 by minimizing the mean square error [26–28]. The multichannel
 subspace algorithm can obtain the better performance than the
 wiener algorithm, especially in the nonstationary noise case. Multichannel statistical estimation algorithm reduces the background
 noise by assuming the Fourier coefficients of the clean speech signal and noise signal obey certain probability distribution and giving certain optimization strategy, such as, minimizing the mean square error, maximizing posterior probability and so on [30–32].
 Kalman filtering is known as an effective speech enhancement technique, where speech signal is usually modeled as autoregressive (AR) process and represented in the state-space domain and

http://dx.doi.org/10.1016/j.neucom.2017.05.048 0925-2312/© 2017 Elsevier B.V. All rights reserved.

Please cite this article as: J. Tu, Y. Xia, Effective Kalman filtering algorithm for distributed multichannel speech enhancement, Neurocomputing (2017), http://dx.doi.org/10.1016/j.neucom.2017.05.048

^{*} This work is supported by the National Natural Science Foundation of China under Grant No. 61179037, and in part by Scientific Research Fund of Education Department of Guangxi Zhuang Autonomous Region under Grant No. 2013YB223 and the Construction Fund of Master's Degree Grant Unit under Gui Degree [2013] No. 4.

2

ARTICLE IN PRESS

J. Tu, Y. Xia/Neurocomputing 000 (2017) 1-11

the speech signal is then recovered from Kalman filter. In comparison with other speech enhancement algorithms, the Kalman filtering algorithm has low computational complexity without the assumption of stationarity signals. So, the Kalman filter has been of great interest in speech enhancement. Many single channel speech enhancement algorithms based on Kalman filter have been proposed. Among them, the earliest Kalman filtering algorithm for speech enhancement in white noise was proposed by Palial and Basu [5]. The Kalman filtering algorithm for speech enhancement in colored noise were proposed by Gibson et al. [6] and Ning et al. [11]. Other existing Kalman filtering algorithms for speech enhancement were presented by improving the accuracy of AR parameters of the Kalman filter [7,8,12-16]. On the other hand, the Kalman filtering-based multichannel speech enhancement algorithm is very less. Only one Kalman filtering-based frequency domain algorithm for multichannel speech enhancement, called the LPC-based multichannel speech enhancement algorithm, was found in a conference paper [33]. In general, the frequency domain algorithm requires more computational sources than the time domain algorithm.

In this paper, we propose a Kalman filtering algorithm for distributed multichannel speech enhancement in the time domain under colored noise environment. The proposed algorithm is the first multichannel speech enhancement algorithm based on Kalman filtering in the time domain. Compared with traditional multichannel speech enhancement algorithms, the proposed algorithm has lower computational complexity and requires less computational resources. Therefore, It is easily used in practical applications. Simulation results show that the proposed algorithm is superior to several conventional multichannel algorithms for distributed multichannel speech enhancement in achieving higher noise reduction and lower signal distortion. Moreover, the proposed algorithm has a faster speed than several multichannel speech enhancement algorithms.

This paper is organized as follows. Section 2 introduces multichannel model and proposes a Kalman filtering-based multichannel speech enhancement algorithm under color noise environment. Section 3 describes how to estimate parameters of the Kalman filter. Section 4 presents the performance evaluation, and Section 5 gives the conclusion.

2. The model and the proposed multichannel algorithm

2.1. Distributed multichannel model

We are concerned with a distributed microphone system that can accurately time align the M noisy speeches [31,32]. The distributed multichannel microphone model is described as

$$y_i(n) = c_i s(n) + v_i(n), \ i = 1, 2, \dots, M$$
 (1)

where *M* is the number of channels, $y_i(n)$ and $v_i(n)$ are the noisy speech and background noise in the *n*th sample and channel *i*, s(n) is the true source signal, and $c_i \in [0, 1]$ are time invariant attenuation factors. In a special case that M = 1 and $c_1 = 1$, the distributed multichannel model then becomes a well-known single channel model. Our goal is to estimate speech signal s(n) from *M* noisy signal observations $\{y_i(n)\}_{i=1}^M$.

Traditional distributed multichannel speech enhancement algorithms mainly include the multichannel wiener algorithm, the multichannel subspace algorithm, the multichannel minimum distortion algorithm, and multichannel statistical estimation-based algorithm. Recently, a Kalman filter-based frequency domain algorithm, called the LPC-based multichannel speech enhancement algorithm, was presented in a conference paper [33]. In general, the frequency domain algorithm requires more computational sources than the time domain algorithm.

2.2. Proposed multichannel algorithm

In this subsection, we propose a Kalman filtering algorithm for distributed multichannel speech enhancement in the time domain in colored noise cases.

Let the speech signal s(n) be modeled as the *p*th-order AR process:

$$s(n) = \sum_{i=1}^{p} a_i s(n-i) + u(n)$$
(2)

where a_i is the *i*th AR speech model parameter and u(n) is driving white noise with variance being $\sigma_u^2(n)$. For our discussion, (2) is expressed as in vector form:

$$\mathbf{s}(n) = \mathbf{F}\mathbf{s}(n-1) + \mathbf{u}(n) \tag{3}$$

where $\mathbf{s}(n) = [s(n-p+1), \dots, s(n)]^T$ is a $p \times 1$ vector, $\mathbf{u} = [0, \dots, 0, u(n)]^T$ is a $p \times 1$ vector, and **F** is a $p \times p$ matrix defined as:

$$\mathbf{F} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ a_p & a_{p-1} & a_{p-2} & \cdots & a_1 \end{pmatrix}.$$

Consider that the speech signal of each channel is corrupted by additive colored noise. Let the *i*th channel noise $v_i(n)$ be modeled as *q*th-order AR process:

$$\nu_i(n) = \sum_{j=1}^{q} b_{ij} \nu_i(n-j) + w_i(n), i = 1, \dots, M$$
(4)

where b_{ij} is the *j*th AR noise model parameter of the *i*th channel and $w_i(n)$ is white Gaussian noise at the *i*th channel with zero mean and variance being $\sigma_{w_i}^2(n)$. (4) can be written as the vector form:

$$\mathbf{v}_i(n) = \mathbf{G}_i \mathbf{v}_i(n-1) + \mathbf{w}_i(n) \tag{5}$$

where $\mathbf{v}_i(n) = [v_i(n-q+1), \dots, v_i(n)]^T$ is the $q \times 1$ vector, $\mathbf{w}_i(n) = [0, \dots, 0, w_i(n)]^T$ is the $q \times 1$ vector, and \mathbf{G}_i is the $q \times q$ matrix given by

$$\mathbf{G}_{i} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ b_{iq} & b_{i(q-1)} & b_{i(q-2)} & \cdots & b_{i1} \end{pmatrix}$$

The noise driving correlation matrix of the *i*th channel $\mathbf{W}_i(n) = E[\mathbf{w}_i(n)\mathbf{w}_i(n)^T]$ is expressed as:

$$\mathbf{W}_{i}(n) = \left(\begin{array}{ccccc} 0 & \cdots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 \\ 0 & \cdots & 0 & \sigma_{W_{i}}^{2}(n) \end{array}\right).$$

Let $\mathbf{x}(n) = [\mathbf{s}^T(n), \mathbf{v}_1^T(n), \dots, \mathbf{v}_M^T(n)]^T$ and let $\mathbf{w}(n) = [\mathbf{u}^T(n), \mathbf{w}_1^T(n), \dots, \mathbf{w}_M^T(n)]^T$. Then the state equations for this system can be written as:

$$\mathbf{x}(n) = \mathbf{G}\mathbf{x}(n-1) + \mathbf{w}(n) \tag{6}$$

$$\mathbf{y}(n) = \mathbf{L}\mathbf{x}(n) \tag{7}$$

where

$$\mathbf{G} = \left(\begin{array}{ccccc} \mathbf{F} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{G}_1 & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{G}_M \end{array} \right),$$

Please cite this article as: J. Tu, Y. Xia, Effective Kalman filtering algorithm for distributed multichannel speech enhancement, Neurocomputing (2017), http://dx.doi.org/10.1016/j.neucom.2017.05.048 Download English Version:

https://daneshyari.com/en/article/6864805

Download Persian Version:

https://daneshyari.com/article/6864805

Daneshyari.com