Signal Processing 94 (2014) 570–575

Contents lists available at ScienceDirect

Signal Processing

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

Fast communication

Widely linear general Kalman filter for stereophonic acoustic echo cancellation

Constantin Paleologu [a,*], Jacob Benesty [b], Silviu Ciochină [a]

[a] University Politehnica of Bucharest, 1-3, Iuliu Maniu Blvd., 061071 Bucharest, Romania
[b] INRS-EMT, University of Quebec, Montreal, QC, Canada H5A 1K6

ARTICLE INFO

Article history:
Received 8 April 2013
Received in revised form
24 June 2013
Accepted 2 August 2013
Available online 9 August 2013

Keywords:
Stereophonic acoustic echo cancellation (SAEC)
Widely linear (WL) model
Kalman filter
Adaptive filters
Recursive least-squares (RLS) algorithm

ABSTRACT

The stereophonic acoustic echo cancellation (SAEC) problem is usually modelled as a two-input/two-output system with real random variables. Recently, the SAEC scheme was recast as a single-input/single-output system with complex random variables, thanks to the widely linear (WL) model. In this paper, we motivate the use of a more general form of the Kalman filter with the WL model for SAEC. Simulation results indicate that this algorithm outperforms the recursive least-squares (RLS) algorithm, which is usually considered as the benchmark for SAEC.

1. Introduction

Stereophonic acoustic echo cancellation (SAEC) is a very challenging system identification problem [1]. Usually, an SAEC system consists of four adaptive filters aiming at identifying four echo paths from two loudspeakers to two microphones. The main difficulty comes from the fact that the loudspeaker (input) signals are linearly related, which results in the so-called nonuniqueness problem [2]. This issue can be addressed by manipulating the signals transmitted to the near-end room, e.g., using a preprocessor on the loudspeaker signals to make them less coherent [3], but without affecting much the stereo perception and the sound quality.

The adaptive filters used in SAEC should exploit the cross-correlation between the channels [4]. In this context, the most interesting solutions belong to the recursive least-squares (RLS) family. Due to their convergence features, these algorithms were preferred in many real-world applications [5,6].

Recently, we proposed a different approach for the SAEC problem [1,7], by using the widely linear (WL) model [8]. Basically, the classical two-input/two-output system with real random variables was recast as a single-input/single-output system with complex random variables. As a consequence, the four real-valued acoustic impulse responses are converted to one complex-valued impulse response. One advantage of this approach is that instead of handling two (real) output signals separately, we only handle one (complex) output signal, which is convenient for the main challenges of SAEC.

In this paper, we derive a general Kalman filter (GKF) with the WL model for SAEC, namely the WL-GKF. The term "general" refers to a different approach we propose, i.e., a block of time samples is considered at each iteration, instead of one time sample (as in the conventional approach). The main motivation behind this work is the appealing performance of the Kalman filter for echo cancellation [9–11]. Also, the WL complex Kalman filters

* Corresponding author. Tel.: +40 21 402 4634.
E-mail addresses: pale@comm.pub.ro (C. Paleologu), benesty@emt.inrs.ca (J. Benesty), silviu@comm.pub.ro (S. Ciochină).

0165-1684/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.sigpro.2013.08.001

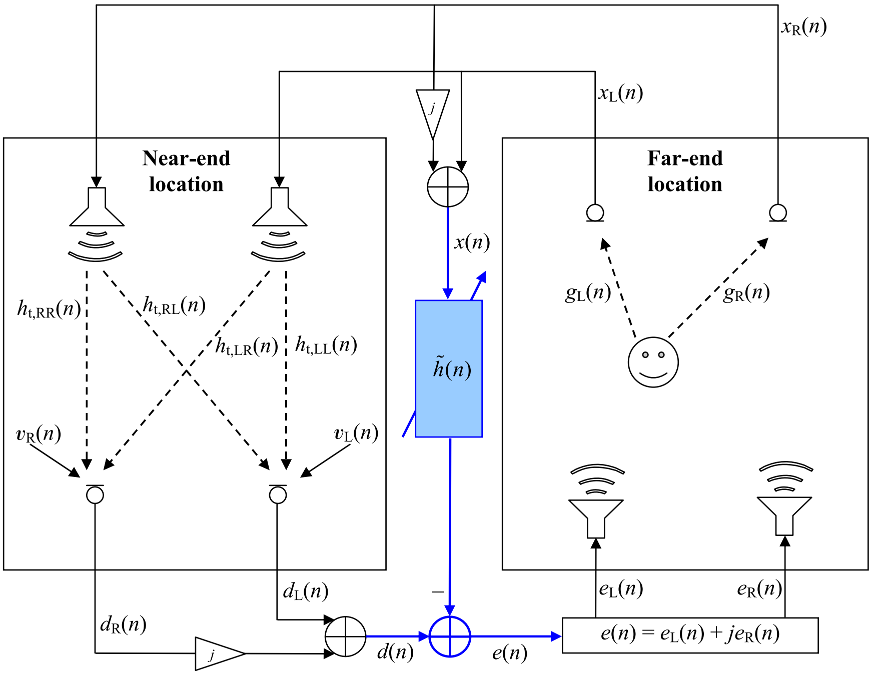

Fig. 1. The WL model for SAEC.

[12,13] were found to be attractive for many applications. The proposed algorithm has inherited some similarities with the WL augmented complex Kalman filter presented in [13]. However, the WL-GKF is derived based on a state variable model suitable for the SAEC problem. The proposed WL-GKF joins the advantages of the WL model (as described before) and the good features of the GKF [11]. Simulation results indicate that the developed algorithm outperforms the RLS counterpart. Consequently, it could represent an attractive alternative in SAEC.

2. The WL model for SAEC

In this section, we briefly review the WL model for SAEC (Fig. 1) [1,7]. Let us denote the two input (or loudspeaker) signals by $x_L(n)$ and $x_R(n)$ (i.e., “left” and “right”), and the two output (or microphone signals) by $d_L(n)$ and $d_R(n)$, where n is the time index. Therefore, the microphone signals are obtained as

$$d_L(n) = y_L(n) + v_L(n), \tag{1}$$

$$d_R(n) = y_R(n) + v_R(n), \tag{2}$$

where $y_L(n)$ and $y_R(n)$ denote the stereo echo signals, and $v_L(n)$ and $v_R(n)$ are the near-end signals (i.e., noise or a combination of noise and near-end speech). The echo signals can be modelled as [2,3]

$$y_L(n) = \mathbf{h}_{t,LL}^T \mathbf{x}_L(n) + \mathbf{h}_{t,RL}^T \mathbf{x}_R(n), \tag{3}$$

$$y_R(n) = \mathbf{h}_{t,LR}^T \mathbf{x}_L(n) + \mathbf{h}_{t,RR}^T \mathbf{x}_R(n), \tag{4}$$

where $\mathbf{h}_{t,LL}, \mathbf{h}_{t,RL}, \mathbf{h}_{t,LR}$, and $\mathbf{h}_{t,RR}$ are L-dimensional vectors of the loudspeaker-to-microphone acoustic impulse responses (the subscript t stands for “true”), the superscript T denotes transposition, and

$$\mathbf{x}_L(n) = [x_L(n)\ \ x_L(n-1)\ \ \cdots\ \ x_L(n-L+1)]^T$$
$$\mathbf{x}_R(n) = [x_R(n)\ \ x_R(n-1)\ \ \cdots\ \ x_R(n-L+1)]^T$$

comprise the L most recent loudspeaker signal samples. In this context, the main goal is to estimate the four acoustic impulse responses, $\mathbf{h}_{t,LL}, \mathbf{h}_{t,RL}, \mathbf{h}_{t,LR}$, and $\mathbf{h}_{t,RR}$, from the microphone signals $d_L(n)$ and $d_R(n)$.

Next, let us form the complex random variable (CRV):

$$d(n) = d_L(n) + jd_R(n) = y(n) + v(n), \tag{5}$$

where $j = \sqrt{-1}$, $y(n) = y_L(n) + jy_R(n)$, and $v(n) = v_L(n) + jv_R(n)$. Also, let us define the complex random vector:

$$\mathbf{x}(n) = \mathbf{x}_L(n) + j\mathbf{x}_R(n). \tag{6}$$

Consequently, the (complex) echo signal can be obtained as

$$y(n) = \mathbf{h}_t^H \mathbf{x}(n) + \mathbf{h}_t'^H \mathbf{x}^*(n), \tag{7}$$

where the superscripts H and $*$ denote conjugate transpose and conjugate, respectively, and

$$\mathbf{h}_t = \mathbf{h}_{t,1} + j\mathbf{h}_{t,2}, \tag{8}$$

$$\mathbf{h}'_t = \mathbf{h}'_{t,1} + j\mathbf{h}'_{t,2}, \tag{9}$$

with $\mathbf{h}_{t,1} = (\mathbf{h}_{t,LL} + \mathbf{h}_{t,RR})/2$, $\mathbf{h}_{t,2} = (\mathbf{h}_{t,RL} - \mathbf{h}_{t,LR})/2$, $\mathbf{h}'_{t,1} = (\mathbf{h}_{t,LL} - \mathbf{h}_{t,RR})/2$, and $\mathbf{h}'_{t,2} = -(\mathbf{h}_{t,RL} + \mathbf{h}_{t,LR})/2$. Using the previous notation, we can express (7) as

$$y(n) = \tilde{\mathbf{h}}_t^H \tilde{\mathbf{x}}(n), \tag{10}$$