Available online at www.sciencedirect.com

ScienceDirect

Speech Communication 58 (2014) 101–110

SPEECH COMMUNICATION

www.elsevier.com/locate/specom

A coherence-based noise reduction algorithm for binaural hearing aids

Nima Yousefian, Philipos C. Loizou [1], John H.L. Hansen *

Department of Electrical Engineering, University of Texas at Dallas, 800 West Campbell Road (EC33), Richardson, TX 75080, USA

Received 4 July 2013; received in revised form 25 October 2013; accepted 11 November 2013
Available online 23 November 2013

Abstract

In this study, we present a novel coherence-based noise reduction technique and show how it can be employed in binaural hearing aid instruments in order to suppress any potential noise present inside a realistic low reverberant environment. The technique is based on particular assumptions on the spatial properties of the target and undesired interfering signals and suppresses (coherent) interferences without prior statistical knowledge of the noise environment. The proposed algorithm is simple, easy to implement and has the advantage of high performance in coping with adverse signal conditions such as scenarios in which competing talkers are present. The technique was assessed by measurements with normal-hearing subjects and the processed outputs in each ear showed significant improvements in terms of speech intelligibility (measured by an adaptive speech reception threshold (SRT) sentence test) over the unprocessed signals (baseline). In a mildly reverberant room with $T_{60} = 200$, the average improvement in SRT obtained relative to the baseline was approximately 6.5 dB. In addition, the proposed algorithm was found to yield higher intelligibility and quality than those obtained by a well-established interaural time difference (ITD)-based speech enhancement algorithm. These attractive features make the proposed method a potential candidate for future use in commercial hearing aid and cochlear implant devices.

Keywords: Binaural signal processing; Coherence function; Noise reduction; SNR estimation; Speech intelligibility

1. Introduction

Noise reduction algorithms applicable to hearing aids have received growing interest recently. This is mainly due to the fact that reduced speech intelligibility under noisy conditions is one of the major complaints made by hearing aid and cochelar implant (CI) users. Over the past two decades, many noise reduction algorithms which utilizes binaural information to improve speech reception in noise and reverberation have been proposed. One main disadvantage of the techniques currently used in bilateral hearing aids is that they are not designed to preserve localization cues (Desloge et al., 1997). It has been shown that the ability to correctly localize sounds help users to exploit visual cues better and therefore lead to large improvements in speech intelligibility (Erber, 1975). In bilateral hearing aids each (left and right) monaural hearing aid relies on its own microphone inputs to generate an output for its corresponding ear and no information is shared between the two hearing aids. In such case, the interaural time difference (ITD) and interaural level difference (ILD), which are considered to be two important binaural cues for localization are not used. The same issue arises with techniques such as adaptive multi-microphone noise reduction techniques that generate an identical output signal for both ears. This limitation causes the binaural cues present in the input signals (including both speech and noise components) to disappear, and therefore, the listener ability for localization is reduced (Van den Bogaert et al., 2007).

It has been accepted that the human binaural system is very effective in separating sound sources even in complex scenarios like an environment where speakers talk simultaneously. For humans, the problem of understanding one talker in the presence of one or more competing speakers is called the cocktail party phenomenon. Over the last

* Corresponding author. Tel.: +1 972 883 2910; fax: +1 972 883 2710.
E-mail addresses: nimayou@utdallas.edu (N. Yousefian), john.hansen@utdallas.edu (J.H.L. Hansen).
[1] Deceased.

0167-6393/$ - see front matter © 2013 Published by Elsevier B.V.
http://dx.doi.org/10.1016/j.specom.2013.11.003

few decades, this problem has been mostly addressed in binaural noise reduction systems (see review in Bronkhorst (2000)). Based on these observations, numerous computational algorithms have been developed which utilize important binaural cues (i.e., ITD and ILD) in order to suppress noise and interference (e.g., Srinivasan et al. (2006), Aarabi and Shi (2004) and Yousefian et al. (2009)). In Harding et al. (2006), a combination of these cues has been employed for mask estimation. In the above-mentioned studies, a (binary or soft) mask is estimated to determine the time–frequency bins that the speech signal is predominant and must be passed without distortion. Although the masks employing the ITD have been shown to be effective for separating sound sources, their performance begins to degrade in more reverberant environments (Palomäki et al., 2004; Kim et al., 2011).

One of the earliest binaural noise reduction and dereverberation techniques, proposed in Allen et al. (1977), is to apply the coherence of input signals as a gain function. The premise behind the coherence function is that speech signals at the two microphones are highly correlated, while interfering signals are uncorrelated. Therefore, if the magnitude of the coherence function between the input signals in the two channels is one or close to one, the speech is dominant and must be passed through the filter. Although coherence-based methods work well when the noise components are uncorrelated, they are deficient when dealing with coherent noise. To address this issue, (Le Bouquin-Jeannés et al., 1997) modified the coherence filter in a way that the cross-power spectral density (CSD) of the noise signals in the two microphones is estimated and included in the original filter.

Another widely studied microphone array speech enhancement technique is multi-channel Wieiner filter (MWF), discussed in Doclo and Moonen (2002) and Spriet et al. (2004). The MWF is a minimum mean square error (MMSE) based technique, which provides an estimate of the desired speech from the signals captured by all microphones. Unlike the traditional beamformers (e.g., generalized sidelobe canceller (GSC) proposed in Griffiths and Jim (1982)), the MWF has the advantage of taking speech distortion into account, which is a useful tool for optimization purposes. It has been shown in Doclo et al. (2006) that MWF perfectly preserve the binaural cues for the speech component, but changes those of the noise component. In Van den Bogaert et al. (2007), the authors modified the cost function of the MWF by adding a interaural transfer function that gives the additional benefit of preserving the binaural cues of the noise component. It should be noted here that in all of the above studies, it is assumed that the noise statistics are known. In general, noise reduction methods that need a VAD, noise or speech presence probability (SPP) estimator (e.g., Yu and Hansen (2009)) to obtain prior knowledge of the signal statistics cannot deal with maskers with highly non-stationary nature (e.g. competing talker(s)).

In the present study, a novel coherence-based noise reduction technique, which is based on certain assumptions regarding the spatial properties of the desired and interfering signals is proposed. The technique utilizes the real and imaginary parts of the coherence function between the input signals as a criterion for mask estimation. Since the proposed algorithm does not require any knowledge of noise statistics for calculating the noise reduction filter, it can cope with highly non-stationary noise sources. Listening tests with normal-hearing subjects and objective quality assessments are conducted in various noise configurations which will show the superiority of the proposed method over the baseline and a well-known ITD-based two-microphone speech enhancement algorithm.

2. Proposed noise reduction technique

In this section, we start with the formal definition of the coherence function and show how it can be analytically modeled in a coherent noise field. Next, the proposed (coherence based) noise reduction filter calculation is discussed in detail. Finally, the implementation details of the proposed algorithm are provided.

2.1. Coherence function in a coherent noise field

There are two types of noise fields usually investigated in array processing studies. First, a diffuse noise environment which is characterized by uncorrelated noise signals of equal power propagating in all directions simultaneously. Second, a coherent noise field that is generated by a single well-defined directional noise source. In coherent noise fields, the noise signals captured by the microphone array are highly correlated. Let us consider the scenario in which a broadside array, consisting of two identical microphones mounted on either side of a head, is placed inside a low reverberant room. The configuration of the two microphones and sound sources is shown in Fig. 1. The target speech source is always at 0° azimuth. Both target speech and noise sources are at a distance of 1 m from the head center. In such a case, without modeling the reverberation

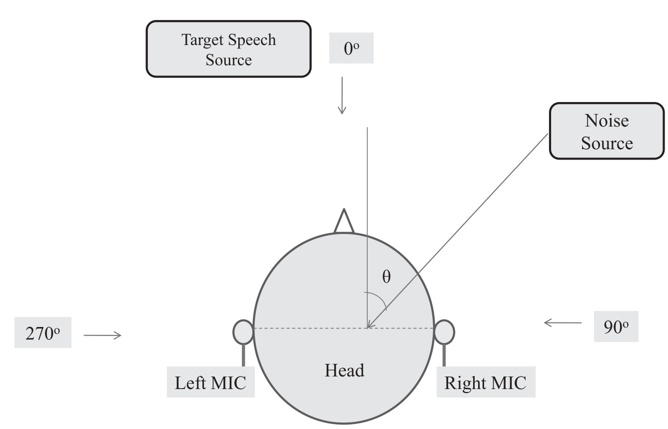

Fig. 1. Placement of the two microphones and sound sources.