Applied Acoustics 131 (2018) 186–191

Contents lists available at ScienceDirect

Applied Acoustics

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

A dual-channel noise reduction algorithm based on the coherence function and the bionic wavelet

Wahbi Nabi*, Mohamed Ben Nasr, Noureddine Aloui, Adnane Cherif

APEES Laboratory, Science Faculty of Tunis, University of Tunis El-Manar, Tunis, Tunisia

ARTICLE INFO

Keywords:
Noise reduction
Dual-channel
Coherence function
Bionic wavelet
Kalman filter

ABSTRACT

A dual-microphone noise reduction method is proposed in this paper. It is based on the coherence function and the bionic wavelet transform using Kalman filter. This filter is applied to the Bionic wavelet coefficients obtained through the application of the BWT to the input signal. The adopted method can treat noisy signals using two closely spaced microphones in different noise scenarios and outperforms other dual-channel speech enhancement methods referring to the perceptual evaluation of speech quality, the spectrogram analysis and the time domain waveforms.

1. Introduction

Cellular phones users suffer from low speech quality in noisy environments. Within this framework, several speech enhancement algorithms have been suggested in order to improve speech quality and intelligibility. Particularly, single channel noise reduction technique has been widely used during the last five decades due to its simple implementation [1,2]. Hence, many single microphone speech enhancement methods have been proposed such as: The spectral subtraction [3], the Kalman filter [4,5] and the Wiener filter [6,7]. However, the performance of these algorithms degrades when the speech is corrupted by the non-stationary noise because of some difficulties in noise estimation [2]. The dual-microphone noise reduction algorithms offer a solution for mobile phones when speech is degraded by a stationary or a non-stationary noise, since it has good performances in term of speech quality and intelligibility. Also, these algorithms are easy to implement and do not require a large computational complexity compared to other multi-channel speech enhancement techniques.

Many dual-channel speech enhancement algorithms have been suggested. In Ref. [8], the authors proposed a dual-microphone speech enhancement method using a complex spectrum plane. This method was dedicated to Hands-free Communication System and gave better results than the original complex spectrum circle centroid method. Kallel et al. [9] presented a dual-channel speech enhancement algorithm based on the Cross Power Spectral Density (CPSD) for bilateral cochlear implant. This approach uses an improved minimum tracking (IMT) technique for noise CPSD prediction. In Ref. [10], authors depicted two dual-microphone noise reduction algorithms based on spectral subtraction. These two algorithms exploit the power spectral density (PSD) and the CPSD of the two recorded signals. Moreover, an improved based coherence function algorithm was proposed in Ref. [11]. This method is applied with two closely microphones and does not require noise statistic prediction. Furthermore, Kim and Kim have proposed in Ref. [12] a noise variance prediction based on dual-microphone phase difference for noise reduction. This method is based on the signal-to-noise ratio (SNR) estimation and has proved its effectiveness in different noise conditions. Finally, Nabi et al. presented in Ref. [20] a dual-microphone noise reduction algorithm for mobile communications exploiting Kalman filter. This algorithm is based on the coherence function and the Kalman filter and can deal with so closely distance between the two microphones.

Related to works cited previously, a dual-channel noise reduction algorithm based on the coherence function is presented in this paper. The proposed method has the capacity to deal with two closely spaced microphones and does not require noise statistics estimation as algorithms proposed in Refs. [11,20]. Moreover, the proposed algorithm is specified by the hybridization of the coherence based-algorithm and the bionic wavelet transform in order to improve the speech quality. This paper is organized as follows. Section 2 presents the original coherence function and the estimate coherence function. Section 3 depicts the proposed bionic wavelet algorithm using Kalman filter applied in our work and the proposed dual-channel speech enhancement method. Section 4 presents the evaluation of the suggested algorithm compared to other methods in terms of the perceptual evaluation of speech quality (PESQ), the time domain waveforms and the spectrogram analysis. Section 5 concludes this paper.

* Corresponding author.
E-mail address: wahbi.nabi@gmail.com (W. Nabi).

https://doi.org/10.1016/j.apacoust.2017.11.003
Received 13 September 2017; Received in revised form 27 October 2017; Accepted 3 November 2017

2. The coherence function

2.1. The original coherence function

This section introduces the coherence function of the two recorded signals. Assuming that the two microphones are placed in noisy environment where the noise and speech signals are separated. After delay compensation, the two input signals are expressed as follows:

$$x_1(m) = s_1(m) + n_1(m) \tag{1}$$

$$x_2(m) = s_2(m) + n_2(m) \tag{2}$$

where $s_1(m)$ and $s_2(m)$ are the speech signals at each microphone, $n_1(m)$ and $n_2(m)$ are the noise signals at the first and the second microphone respectively.

The discrete Fourier transform (DFT) of the two last equations can be written as:

$$X_1(n,k) = S_1(\omega,k) + N_1(\omega,k) \tag{3}$$

$$X_2(n,k) = S_2(\omega,k) + N_2(\omega,k) \tag{4}$$

where k presents the frame index and $\omega = 2\pi d/L$ is the angular frequency where L denotes the frame length in samples.

The coherence function is a criterion used to estimate the speech signal through the use of two microphones [13,14]. The coherence function between two recorded signals can be computed as:

$$F_{x1x2}(\omega,k) = \frac{P_{x1x2}(\omega,k)}{\sqrt{P_{x1}(\omega,k)P_{x2}(\omega,k)}} \tag{5}$$

where P_{x1x2} presents the cross power spectral density (CPSD) of the two input signals, P_{x1} and P_{x2} are the power spectral density (PSD) at each microphone.

The magnitude of the coherence function is used to gather information about the presence of the noise or the speech. Thus, the speech is dominant when the magnitude tends to zero and the noise is dominant when the magnitude tends to one.

In this work, we use a configuration of two closely spaced microphones. So, interfering signals are highly correlated specially for lower frequencies [22]. In this case, when the noise is diffused, the coherence function is real valued and can be expressed analytically as follows:

$$F_{x1x2} = \mathrm{sinc}\left(\frac{\omega f_s d}{s}\right) \tag{6}$$

where f_s is the sampling frequency and s is the sound velocity.

Fig. 1 presents the design of the two omnidirectional microphones used in our work, where:

- The two microphones are spaced of 2 cm.
- The speech and noise source are located at 0° and α° azimuths respectively.
- The two sources and the microphones are spaced by 1.2 m.

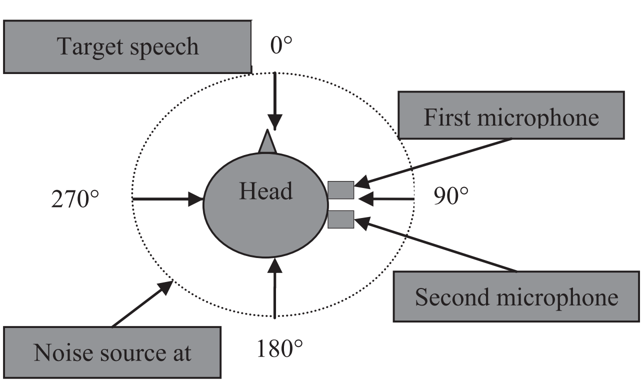

Fig. 1. Design of the two omnidirectional microphones with target sound sources.

Based on the last design, the coherence function can be expressed as follows:

$$F_{x1x2}(\omega) = e^{j\omega f_s (d/s)\cos\alpha} \tag{7}$$

where α is the angle of incidence.

2.2. The estimate coherence function

Generally, the coherence function (CF) can be expressed from speech signal and interfering signals. So, it can be defined as follows [9]:

$$F_{x1x2} = F_{s1s2}\sqrt{\frac{P_{s1}}{P_{s1}+P_{n1}}}\sqrt{\frac{P_{s2}}{P_{s2}+P_{n2}}} + F_{n1n2}\sqrt{\frac{P_{n1}}{P_{s1}+P_{n1}}}\sqrt{\frac{P_{n2}}{P_{s2}+P_{n2}}} \tag{8}$$

where F_{s1s2} is the CF between s_1 and s_2, P_{si} is the PSD of s_i $(i = 1,2)$, P_{ni} is the PSD of n_i (i = 1,2) and F_{n1n2} is the CF between n_1 and n_2.

Noting that the signal-to-noise-ratio (SNR) at the i^{th} microphone can be expressed as follows:

$$SNR_i = \frac{P_{si}}{P_{ni}} \qquad (i = 1,2) \tag{9}$$

Thus, the estimate CF defined in [11] can be expressed as:

$$\widetilde{F}_{x1x2} = [\cos(\omega\tau) + j\sin(\omega\tau)]\frac{\widetilde{SNR}}{1+\widetilde{SNR}} + [\cos(\omega\tau\cos\alpha) + j\sin(\omega\tau\cos\alpha)]\frac{1}{1+\widetilde{SNR}} \tag{10}$$

where $\tau = f_s(d/s)$ and $\widetilde{SNR} = SNR_1 = SNR_2$ because the two microphones are closely spaced and have SNR values almost identical.

3. The proposed dual-microphone speech enhancement algorithm

3.1. The proposed noise reduction algorithm based on BWT and Kalman filter

The bionic wavelet transform (BWT) is a wavelet transform that integrates the mechanism of the active cochlear [23], in order to obtain an adaptive time-frequency analysis and a system based on biology. Furthermore, the BWT depicts a good energy specification that presents a good speech separation in the coefficients. It was used in the speech enhancement field [15,16] and outperforms other single microphone algorithms such as: the Ephrahim and Malih filter [17], the spectral subtraction [3] and the discrete wavelet transform [18].

The Kalman filter is a well-known filter commonly used as a solution to the linear MMSE problems in stochastic system since it can perform an optimal estimation of the system state [4]. In fact, the Kalman filter is exploited in the proposed noise reduction algorithm as presented in Fig. 2. The idea of the proposed system is inspired from the Wiener filter application in the BWT for the noise reduction as depicted in Ref. [19]. From the figure, it is clear that the Kalman filter is applied to the Bionic wavelet coefficients given by the application of the BWT to the noisy signal. Then, the inverse of the Bionic wavelet transform was applied in order to obtain the enhanced signal.

In Fig. 2, $k_1, k_2, \ldots, k_n$ present the 18 bionic wavelet coefficients of the input signal and $\widetilde{k}_1, \widetilde{k}_2, \ldots, \widetilde{k}_n$ are the 18 filtered bionic wavelet coefficients obtained after Kalman filter application.

3.2. The proposed dual-microphone noise reduction algorithm

The proposed dual-microphone noise reduction algorithm is composed of two steps. The first step is the dual-channel speech enhancement method based on the coherence function proposed in Ref. [11]. In this step, we apply on $x_2(m)$ the delay compensation which is used by many dual-microphone noise reduction systems. Then, the two noisy signals are treated with a Hanning window in a 20-ms frames and a