Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 49 (2015) 244 – 252

Procedia Computer Science

ICAC3'15

Selective Frequency Enhancement of Speech Signal for Intelligibility Improvement in Presence of Near-end Noise

Premananda B.S.[a*], Uma B.V.[b]

[a,b]*R.V. College of Engineering, Bengaluru, India*

Abstract

In mobile phones, perceived speech signal deteriorates significantly in the presence of background noise as noise reaches directly at the listener's ears. There is an inherent need to increase the intelligibility of the received speech signal in noisy environments by incorporating speech enhancement algorithm at the receiverend. This paper focus on the enhancement of selective speech samples for the intelligibility improvement, when the near-end noise dominates. Audible speech samples are selected by consideringthreshold of hearing and auditory masking properties of the human ear. Intelligibility of enhanced speech signalis measured using Speech Intelligibility Index. Experimental results show the intelligibility improvement of the speech signal in lower SNR with the proposed approach to the unprocessed far-end speech signal. This approach is efficient in overcoming the degradation of speech signals in very noisy environments.

Peer-review under responsibility of organizing committee of the 4th International Conference on Advances in Computing, Communication and Control (ICAC3'15)

Keywords:Gain; Masking; Near-end noise; Speech enhancement; Speech intelligibility.

1. Introduction

Mobile devices are the most popular consumer devices in the current scenario and have made a huge impact on day to day life. In aquiet environment, less speech energy is required for the speakers to understand each other's conversation. At the receiving end, referredto as *near-end* in the literature, the listener may be in a noisy

* Corresponding author. Tel.: 08067178176.
E-mail address: premanandabs@gmail.com

Peer-review under responsibility of organizing committee of the 4th International Conference on Advances in Computing, Communication and Control (ICAC3'15)
doi:10.1016/j.procs.2015.04.250

environment. It makes hearing difficult, even though, the transmitting speech source is in a quiet environment because the near-end noise reaches the listener's ear directly. Listener experiences fatigue as the clarity of the speech signal is decreased. Increasing the volume of mobile phones to emphasize the far-end speech is a way to resolve the problem. The external volume control of the mobile phone cannot be used for this purpose as background noise changes dynamically.

The presence of noise masks the speech signal and makes it less intelligent or audible, effect is called masking. Itis of two types, one simultaneous masking where,one signal is masked when another signal is present (noise can also mask the speech signal) and other temporal masking whereinthe signal is masked by noise before and after the interval of occurrence of noise.Hence, the speech signal needs to be enhancedconsidering these situations in the purview also. The aim, of including masking effects in speech signal enhancement, is to suppress the non-audible spectral components of the speech signal.

The approaches proposed for far-end noise cancellation/reduction techniques discussed in the literature [15-18] arenot suitable as they focus on mitigating the near-end noise at the speakerend rather than at the receiverend. Several approaches to mitigate the near-end noise using speech enhancement is discussed by Bastian S. et al., [4-6] and Taal C.H. et al., [7, 8]. Near-end listening enhancement (NELE) algorithm by Bastian S. et al., in [4] maximizes the speech intelligibility index (SII) [14] and thus the speech intelligibility by selective frequency enhancing of the speech signal power. [6] investigates the listening enhancement under the constraint that the processed loudspeaker signal power is strictly equal to the power of the received signal.Two SII based NELE algorithms are compared in [7] to optimize the speech intelligibility in the presence of near-end noise and it focuses on a novel approach to linear filtering of speech prior to the degradation due to near-end noise. Taal C.H. et al., in [8] solved constrained optimization problems of [5] using a nonlinear approximation of the speech intelligibility that is correctat lower SNRs.

NELE by Premananda B.S. et al., in [2] increases speech signal energy above the noise energy and avoids listener fatigue. In [1,3] speech samples are given a relative weight using an absolute threshold hearing (ATH)but do not include the masking effect of signals. Approaches in [1-3] did not consider the audible speech samples, rather it involves enhancement of both audible and non-audible samples, resulting in waste of speech energy and mobile battery. NELE algorithm by Teddy S. et al., [11] provides an operative model of temporal masking that uses a fractional bark gammatone filter bank related to the changes in speech enhancement method.

This paper examines a novel speech enhancement method exploiting the psychoacoustic model of the human ear. The paper is organized as follows: section II describes NELE in FFT domain with implementation details. The loudness computation procedure is explained in section III. Dominant frequency estimation procedure is listed in section IV. Experimental results and conclusions are discussed in section V and VI respectively.

2. NELE in FFT Domain

An NELE approach in the frequency domain [10] (including threshold of hearing and auditory masking properties of the human ear)is proposed to enhance the speech signal in the presence of high near-end noise. Reduction of clarity or intelligibility of the speech signal due to the presence of near-end noise can be minimized by pre-processing the clean speech signal before being played in noisy environments or fed to the mobile speakers. The overall block diagram of the proposed approach is illustrated in Fig. 1. It is assumed that a clean speech signal with far-end noise removed (using noise-cancellation techniques) is available.