

# Single Channel Speech Enhancement using a Complex Spectrum Method

A. M. Mutawa

**Abstract**—Speech enhancement plays an important role in speech communication systems. Speech signal enhancement in an additive noise environment in speech recognition and speaker verification system is still a challenging task. In speech enhancement process, the spectral analysis method has more advantageous than other methods due to its simplicity and effective localization of noise components in the signal. But, this method does not analyze the phase information for efficient speech enhancement. This present work proposes a modified phase spectrum compensation method for speech enhancement in a single channel environment that analyzes both the magnitude and phase spectrum of the speech signal. The performance of the proposed method is compared with that of three conventional methods (Spectral Subtraction (SSUB), Minimum Mean Square Error (MMSE) estimator, and Phase Spectrum Compensation (PSC)) through four different objective measures: Log spectral distance (LSD), log likelihood ratio (LLR), itakura-siata (IS) measure, and short-time objective intelligibility (STOI). The experimental results show that the three objective measures (LLR, LSD, and STOI) of the proposed method gives better results over the conventional methods in four different noise Signal to Noise Ratio (SNR).

**Keywords**—Complex spectrum, Noise reduction, Objective measures, Speech enhancement

## I. INTRODUCTION

**S**PEECH enhancement is a field of research that focuses on improving the quality of a speech signal under different environments, and it has become a very popular research topic in recent decades. Speech enhancement plays a key role in the speech communication process. Speech communication is one of the important modes of communication between humans and between human and machine. In real life, the users are expecting that the speech communication system should be so robust to work on any environment at any time [1]. Speech communication involves two important processes namely speaker verification and speech recognition. Interferences due to noise is significantly affecting the performance of any speech communication system and it affects the quality of the original speech signal by some ratio and it is measured using Signal to Noise Ratio (SNR). In specific, improving the speech quality of the signal under moderate to high noise (-5 dB to 15 dB) is highly challenging [1], [2]. There are several kinds of research in the literature over the past several decades

discusses the speech enhancement process in a noisy environment [3]–[5].

Besides, the type of environment or surrounding also plays a key role in speech enhancement process. Because, the signal can be acquired from a different environment such as additive noise, reverberation, filtering, and clipping, etc. Hence, it is highly important to estimate the original speech information from the noise effects in the speech signal processing for developing intelligent speech communication systems. Recent developments in speech communication devices such as speech assisting devices, speech communication systems (mobile phones), hearing aids, and cochlear implants are highly sensitive to noise information present in the signal, which must be carefully removed from the original signal for efficient sound reproduction. This present work mainly focused on analyzing the degradation due to additive noises of different SNR from moderate to high values (0 dB, 5 dB, 10 dB, and 15 dB).

The major interest on speech enhancement methodology is to suppress the effect of noise in the original speech signal to improve its quality. It is one of the common problems in either single channel (one microphone) system or multi-channel (more than one microphone) system. Most of the earlier work in the literature focused on investigating the speech enhancement process in single channel microphone system due to its size, cost and computational efficiency [3]–[6]. The most successful method of speech enhancement depends on two major factors: (i) how effectively the method localizes the noise component in the signal and (ii) how intelligently it reduces the effects of noise to enhance the speech signal. Most conventional speech-enhancement methods detect the unvoiced region in the signal as noise or vice-versa [6], [7]. Noise is mainly due to artefacts or environmental factors, and the spectral analysis method can effectively distinguish the unvoiced region from the signal. Furthermore, in place of reducing the effect of noises, the speech enhancement algorithms remove the original signal information. Thereby, the performance of any speech enhancement algorithms depends on, (a) parameter settings of the algorithms (b) the value of SNR (c) type of noise and its environment and (d) calculation of noise estimation [7]. Therefore, it is always challenging to design and develop an intelligent speech enhancement algorithm that is suitable for environments with different noise backgrounds. Hence, it is highly evident to develop an intelligent and adaptive speech enhancement method for efficient speech enhancement in real-life

This work was funded by Kuwait Foundation for the Advancement of Science (KFAS) grant number PR1718SM05

A. M. Mutawa is with the Computer Engineering Department, Kuwait University, 13060, +(965) 24987160, Kuwait (e-mail: dr.mutawa@ku.edu.kw).





















