

INDIAN INSTITUTE OF TECHNOLOGY, BOMBAY

Speech Enhancement using Wiener Filter

by

Pranav Sankhe

Supervisor: Prof. Debraj Chakraborty

November 2018

Contents

1	Introduction	1
1.1	Task Description	1
1.2	Methods Investigated	1

Chapter 1

Introduction

1.1 Task Description

The task of speech enhancement by suppression of acoustic noise has been explored here. The problem of speech enhancement has been heavily researched on owing to its immense importance in plethora of audio processing tasks like speech recognition and compression, restoration of analog audio recordings, etc. In particular, the following work focuses on enhancement of speech recordings and the transfer of the techniques from speech to music domain sure exists but is not trivial. This transfer has not been explicitly discussed in this paper but analysis of audio suppression in music signals is an important and interesting research in its own right.

If we do not assume any particular nature whatsoever of the noise we aim to suppress in a speech signal, the task of denoising becomes an extremely challenging. In this paper, the noise has been assumed to be additive and stationary i.e. white noise (broadband noise, like tape hiss), colored noise, and different kinds of narrowband noises.

1.2 Methods Investigated

Specifically, two methods of noise suppression have been explored here namely, Spectral Subtraction and Wiener Filtering. The implementation of both of these methods and their peculiarities have been discussed in the subsequent sections. It is observed that the techniques considered are not equally good at removing noises of different nature.

The choice of the technique significantly impacts the performance given a specific application and a specific noise structure. The paper concludes with the advantages and the disadvantages of both methods and suggest some plausible improvisations that can potentially improve the noise supression.

Chapter 2

Speech Enhancement by Spectral Subtraction

2.1 Theory

Spectral Subtraction is a widely used algorithm for speech enhancement due to its simplicity of implementation.

Given our assumption of the additive property of noise and the linearity property of Fourier transform, we have the following:

$$Y(e^{j\omega}) = X(e^{j\omega}) + N(e^{j\omega})$$