Time Difference of Arrival Estimation of Speech Source in a Noisy and Reverberant Environment

Tsvi G. Dvorkind [a,*] and Sharon Gannot [b]

[a]*Faculty of Electrical Engineering, Technion, Technion City, 32000 Haifa, Israel*

[b]*School of Electrical Engineering, Bar-Ilan University, 52900 Ramat-Gan, Israel*

Abstract

Determining the spatial position of a speaker finds a growing interest in video conference scenario where automated camera steering and tracking are required. Speaker localization can be achieved with a dual step approach. In the preliminary stage microphone array is used to extract the *time difference of arrival* (TDOA) of the speech signal. These readings are then used by the second stage for the actual localization. In this work we present novel, frequency domain, approaches for TDOA calculation in a reverberant and noisy environment. Our methods are based on the speech quasi-stationarity property, and on the fact that the speech and the noise are uncorrelated. The mathematical derivations in this work are followed by an extensive experimental study which involves static and tracking scenarios.

Key words: Source localization, non-stationarity, decorrelation
PACS:

1 Introduction

Determining the spatial position of a speaker finds a growing interest in video conference scenario where automated camera steering and tracking are required. Microphone array, which is usually used for speech enhancement in a noisy environment [1], can be used for the task of speaker localization as well. The related algorithms can be divided into two groups: single and dual

---