CCIT Report #398
October 2002

Subspace Methods for Multi-Microphone Speech Dereverberation

Sharon Gannot

Department of Electrical Engineering, Technion — Israel Institute of Technology,
Technion City, Haifa 32000, Israel
E-mail: gannot@siglab.technion.ac.il;
Tel.: +972 4 8294756; Fax: +972 4 8323041.

Marc Moonen

Department of Electrical Engineering, ESAT-SISTA
K.U.Leuven,
Kasteelpark Arenberg 10,
3001 Leuven-Heverlee, Belgium
E-mail: marc.moonen@esat.kuleuven.ac.be

Abstract

A novel approach for multi–microphone speech dereverberation is presented. The method is based on the construction of the null subspace of the data matrix in the presence of colored noise, using the *generalized singular value decomposition* (GSVD) technique, or the *generalized eigenvalue decomposition* (GEVD) of the respective correlation matrices. The special Silvester structure of the filtering matrix, related to this subspace, is exploited for deriving a *total least squares* (TLS) estimate for the *acoustical transfer functions* (ATFs). Other, less robust but computationally more efficient methods are derived based on the same structure and on the QR decomposition (QRD). A preliminary study of the incorporation of the subspace method into a

This research work was partly carried out at the ESAT laboratory of the Katholieke Universiteit Leuven, in the frame of the Interuniversity Attraction Pole IUAP P4-02, *Modeling, Identification, Simulation and Control of Complex Systems*, the Concerted Research Action *Mathematical Engineering Techniques for Information and Communication Systems* (GOA-MEFISTO-666) of the Flemish Government and the IT-project *Multi-microphone Signal Enhancement Techniques for handsfree telephony and voice controlled systems (MUSETTE-2)* of the I.W.T., and was partially sponsored by Philips-ITCL.

subband framework proves to be efficient, although some problems remain open. Speech reconstruction is achieved by virtue of the *matched filter beamformer* (MFBF). An experimental study supports the potential of the proposed methods.

Keywords

Speech Dereverberation , Microphone Arrays, Subspace Methods, Subband Structures.

I. Introduction

In many speech communication applications, the recorded speech signal is subject to reflections on the room walls and other objects on its way from the source to the microphones. The resulting speech signal is then called reverberated. The quality of the speech signal might deteriorate severely and this can even cause a degradation in intelligibility. Subsequent processing of the speech signal, such as speech coding or automatic speech recognition might be rendered useless in the presence of reverberated speech. Although single–microphone dereverberation techniques do exist, the most successful methods for dereverberation are based on multi-microphone measurements.

Spatio-temporal methods, which are directly applied to the received signals, have been presented by Liu *et al.* [1] and by Sánchez–Bote *et al.* [2]. They consist of a spatial averaging of the minimum–phase component of the speech signal and cepstrum domain processing for manipulating the all–pass component of the speech signal. Other methods use the linear prediction residual signal to dereverberate the speech signal [3],[4].

Beamforming methods [5],[6] which use an estimate of the related ATFs can reduce the amount of reverberation, especially if some a priori knowledge of the acoustical transfer is given. The average ATFs of all the microphones proves to be efficient and quite robust to small speaker movements. However, if this information is not available, these methods can not eliminate the reverberation completely. Hence, we will avoid using the small movement assumption in this work, as it is not valid in many important applications.

Subspace methods appear to be the most promising methods for dereverberation. These methods consist of estimating the null subspace of the data matrix. These null subspace vectors are used to extract the ATFs (e.g. [7] and [8]). Of special interest is the EVAM algorithm presented by Gürelli and Nikias [9]. As the null subspace vectors are shown to be filtered versions of the actual ATFs, extraneous zeros should be eliminated. This is done by the “fractal” method which is essentially a