

Blekinge Institute of Technology  
Doctoral Dissertation Series No 2012:03  
ISBN 978-91-7295-226-3  
ISSN 1653-2090

# Direction of Arrival Estimation and Localization of Multiple Speech Sources in Enclosed Environments

Mikael Swartling

Akademisk avhandling

Som för avläggande av teknologie doktorsexamen vid Blekinge Tekniska Högskola kommer att offentligt försvaras i sal J1640, Karlskrona, fredagen den 20:e Januari, 2012, kl. 13:00.



**Blekinge Institute of Technology**  
Department of Electrical Engineering  
Blekinge Institute of Technology  
SE-371 79 Karlskrona, Sweden

## **Handledare**

Doc. Nedelko Grbić  
*Blekinge Tekniska Högskola, Karlskrona*

## **Fakultetsopponent**

Doc. Martin Stridh  
*Lunds Tekniska Högskola, Lund*

## **Betygsnämnd**

Prof. Börje Nilsson  
*Linnéuniversitetet, Växjö*  
Prof. Mikael Sternad  
*Uppsala Universitet, Uppsala*  
Dr. Niels Henrik Pontoppidan  
*Eriksholm Research Center, Oticon, Köpenhamn*

## **Abstract**

Speech communication is gaining in popularity in many different contexts as technology evolves. With the introduction of mobile electronic devices such as cell phones and laptops, and fixed electronic devices such as video and teleconferencing systems, more people are communicating which leads to an increasing demand for new services and better speech quality. Methods to enhance speech recorded by microphones often operate blindly without prior knowledge of the signals. With the addition of multiple microphones to allow for spatial filtering, many blind speech enhancement methods have to operate blindly also in the spatial domain. When attempting to improve the quality of spoken communication it is often necessary to be able to reliably determine the location of the speakers. A dedicated source localization method on top of the speech enhancement methods can assist the speech enhancement method by providing the spatial information about the sources.

This thesis addresses the problem of speech-source localization, with a focus on the problem of localization in the presence of multiple concurrent speech sources. The primary work consists of methods to estimate the direction of arrival of multiple concurrent speech sources from an array of sensors and a method to correct the ambiguities when estimating the spatial locations of multiple speech sources from multiple arrays of sensors. The thesis also improves the well-known SRP-based methods with higher-order statistics, and presents an analysis of how the SRP-PHAT performs when the sensor array geometry is not fully calibrated. The thesis is concluded by two envelope-domain-based methods for tonal pattern detection and tonal disturbance detection and cancelation which can be useful to further increase the usability of the proposed localization methods.

The main contribution of the thesis is a complete methodology to spatially locate multiple speech sources in enclosed environments. New methods and improvements to the combined solution are presented for the direction-of-arrival estimation, the location estimation and the location ambiguity correction, as well as a sensor array calibration sensitivity analysis.

## **Keywords**

Beamforming, Detection and Classification, Speech Enhancement, Source Localization