Chapter 6

Sound Source Localization: Conventional Methods and Intensity Vector Direction Exploitation

Banu Günel
University of Surrey, UK

Hüseyin Hacıhabiboğlu
King’s College London, UK

ABSTRACT

Automatic sound source localization has recently gained interest due to its various applications that range from surveillance to hearing aids, and teleconferencing to human computer interaction. Automatic sound source localization may refer to the process of determining only the direction of a sound source, which is known as the direction-of-arrival estimation, or also its distance in order to obtain its coordinates. Various methods have previously been proposed for this purpose. Many of these methods use the time and level differences between the signals captured by each element of a microphone array. An overview of these conventional array processing methods is given and the factors that affect their performance are discussed. The limitations of these methods affecting real-time implementation are highlighted. An emerging source localization method based on acoustic intensity is explained. A theoretical evaluation of different microphone array geometries is given. Two well-known problems, localization of multiple sources and localization of acoustic reflections, are addressed.

DOl: 10.4018/978-1-61520-919-4.ch006

Copyright © 2011, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
Sound Source Localization

ing aids (Desloge, Rabinowitz, & Zurek, 1997; Welker, Greenberg, Desloge, & Zurek, 1997; Kates, 1998; Widrow, 2001) and human-computer interaction (HCI).

Sound source localization is impossible using a single sensor. Considering a fixed sensor structure, it is a natural conclusion that there should be at least two sensors. Various methods using an array of microphones have previously been proposed for sound source localization. These include, but are not limited to, steered response power (SRP) localization, high-resolution spectral-estimation, and time-delay-of-arrival (TDOA) estimation. There are also different variations and combinations of these methods. However, all of these methods use the time and level differences between the signals captured by each sensor in a microphone array.

Steered response power localizers carry out beamforming at all directions by digitally steering the array and look for the direction that maximizes the signal power. A simple version of this type of locator is known as the steered-beamforming based localizer.

High-resolution spectral estimation based localizers, such as the Multiple Signal Classification (MUSIC) algorithm, compute a spatio-spectral correlation matrix using the signals recorded by each microphone and decompose it into signal and noise subspaces. A search is then carried out in these subspaces to detect the possible direction of arrivals.

TDOA-based localizers use the time delays between pairs of microphones and the array position information. For a plane wave, the observed time delay between a pair of microphones is constant on a hyperboloid. A direction estimate can be obtained from several microphone pairs by finding the optimal solution of this system of hyperboloids. TDOA estimation is usually made by the generalized cross-correlation (GCC) method. GCC involves a weighting operation to improve the performance of the TDOA based localizers under reverberant conditions.

In addition to these conventional methods, there are other approaches proposed for sound source localization, such as biologically inspired methods that mimic binaural hearing mechanism as well as methods that mimic cognitive aspects of hearing by using artificial intelligence. The auditory system provides an efficient means of sound source localization. The mammalian auditory systems consist of two ears and the central processing in the brain to determine the DOA exploiting the time and level differences between the sounds arriving at the two ears. Positions of ears can be changed by head/body movements to allow optimal and adaptive localization when combined with visual cues. In addition, the auditory cognitive system uses other cues, such as the loudness, and direct-to-reverberant sound energy ratio in order to determine the source distance.

There are several real-life requirements that should be met by a successful sound source localization system. Most important of these requirements are good localization accuracy, high speed, low cost, small array size, small number of channels for ease of data interfacing and 3D symmetry of operation. Conventional array processing methods do not satisfy most of these requirements due to real-life considerations. For example, cosmetic constraints may limit the number of microphones for hearing aid applications.

This chapter provides an overview of conventional microphone array processing methods including the recent developments based on these methods. The factors that affect the performance of sound source localization algorithms are classified and discussed. The limitations of conventional localizers are highlighted within the context of real-life design considerations. An emerging method based on sound intensity, which was proposed as a solution to these limitations, is explained. Special microphone array geometries required for calculating the intensity vector directions and the directional accuracy achieved by these geometries are discussed in detail. Solutions to some well-known practical problems, more