⊕ Help & FAQ   🌐

Home   Profiles   Research Units   **Research output**   Activities   Datasets   Equipment   …

Search...   🔍

Acoustic Source Localization in a Room Environment and at Moderate Distances

Pasi Pertilä

Research output: Book/Report › Doctoral thesis › Collection of Articles

901
Downloads
(Pure)

📖 **Overview**   ◎ Fingerprint

Abstract

The pressure changes of an acoustic wavefront are sensed with a microphone that acts as a transducer, converting sound pressure into voltage. The voltage is then converted into digital form with an analog to digital (AD) -converter to provide a discrete time quantized digital signal. This thesis discusses methods to estimate the location of a sound source from the signals of multiple microphones. Acoustic source localization (ASL) can be used to locate talkers, which is useful for speech communication systems such as teleconferencing and hearing aids. Active localization methods receive and send energy, whereas passive methods only receive energy. The discussed ASL methods are passive which makes them attractive for surveillance applications, such as localization of vehicles and monitoring of areas. This thesis focuses on ASL in a room environment and at moderate distances that are often present in outdoor applications. The frequency range of many commonly occurring sounds such as speech, vehicles, and jet aircraft is large. Time delay estimation (TDE) methods are suitable for estimating properties from such wideband signals. Since TDE methods have been extensively studied, the theory is attractive to apply in localization. Time difference of arrival (TDOA) -based methods estimate the source location from measured TDOA values between microphones. These methods are computationally attractive but deteriorate rapidly when the TDOA estimates are no longer directly related to the source position. In a room environment such conditions could be faced when reverberation or noise starts to dominate TDOA estimation. The combination of microphone pairwise TDE measurements is studied as a more robust localization solution. TDE measurements are combined into a spatial likelihood function (SLF) of source position. A sequential Bayesian method known as particle filtering (PF) is used to estimate the source position. The PF based localization accuracy increases when the variance of SLF decreases. Results from simulations and real-data show that multiplication (intersection operation) results in a SLF with smaller variance than the typically applied summation (union operation). The above localization methods assume that the source is located in the near-field of the microphone array, i.e., the source emitted wavefront curvature is observable. In the far-field, the source wavefront is assumed planar and localization is considered by using spatially separated direction observations. The direction of arrival (DOA) of a source emitted wavefront impinging on a microphone array is traditionally estimated by steering the array to a direction that maximizes the steered response power. Such estimates can be deteriorated by noise and reverberation. Therefore, talker localization is considered using DOA discrimination. The sound propagation delay from the source to the microphone array becomes significant at moderate distances. As a result, the directional observations from a moving sound source point behind the true source position. Omitting the propagation delay results in a biased location estimate of a moving or discontinuously emitting source. To solve this problem the propagation delay is proposed to be modeled in the estimation process. Motivated by the robustness of localization using the combination of TDE measurements, source localization by directly combining the TDE-based array steered responses is considered. This extends the near-field talker localization methods to far-field source localization. The presented propagation delay modeling is then proposed for the steered response localization. The improvement in localization accuracy by including the propagation delay is studied using a simulated moving sound source in the atmosphere. The presented indoor localization methods have been evaluated in the Classification of Events, Activities and Relationships (CLEAR) 2006 and CLEAR'07 technology evaluations. In the evaluations, the performance of the proposed ASL methods was evaluated by a third party from several hours of annotated data. The data was gathered from meetings held in multiple smart rooms. According to the obtained results from CLEAR'07 development dataset (166 min) presented in this thesis, 92 % of speech activity in a meeting situation was located within 17 cm accuracy.

Access to Document

View full fingerprint ›

Cite this

| APA | Author | BIBTEX | Harvard | Standard | RIS | Vancouver |

About web accessibility   Contact us