A Robust High Resolution Speaker DOA Estimation under Reverberant Environment

Yifan Guo , Y.X. Zou, Yongqing Wang*

ADSPLAB/ELIP, School of Electronic and Computer Engineering, Peking University, Shenzhen 518055

guoyifan@sz.pku.edu.cn, zouyx@pkusz.edu.cn*, wangyq@sz.pku.edu.cn

Abstract

Direction of arrival (DOA) estimation of the spatial speech source is a key technique in the audition system of the service robot. This paper investigates a robust high resolution speaker DOA estimation based on acoustic vector sensor (AVS) and spatial sparsity representation (SSR) theory of source. The approximate model of the inter-sensor data ratio (ISDR) of AVS in the time-frequency (TF) domain is derived with reverberation and noise, which determines the relationship between the AVS manifold vector and the ISDR. To obtain a robust speaker DOA estimation, the paper gets reliable high local signal-to-noise ratio (HLSNR) TF points by extracting the pitch of speech signal and fitting the curve. Then the SSR model of DOA estimation is formulated and the high DOA estimation accuracy is achieved. The experimental results under different reverberation and additive noise conditions show that the proposed DOA estimation method is able to achieve RMSE of below 0.5° when the SNR is from 5dB to 30dB. Moreover, the method is independent of the source frequencies and not sensitive to reverberation. Since AVS has a small size and few sensors, this DOA estimation approach will probably provide solutions for the speaker source DOA estimation of service robots in the natural home environment.

Index Terms: acoustic vector sensor; direction of arrival estimation; spatial sparse representation; inter-sensor data ratio; time-frequency sparsity

978-1-4799-4219-0/14/$31.00