

Title	等化-キャンセル理論にもとづいた両耳聴音源方向推定に関する研究
Author(s)	Chau, Thanh Duc
Citation	
Issue Date	2014-09
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/12288
Rights	
Description	Supervisor:赤木 正人, 情報科学研究科, 博士

Abstract

Simulating the human auditory system to deal with problems in sound signal processing is an interesting topic that has attracted a huge number of research recently. One of the important problems is binaural sound source localization (SSL), which plays a crucial role in binaural speech enhancement, binaural source separation and humanoid robot. Although previous research has achieved many impressive results in sound localization, the problem of binaural SSL in the presence of noise and reverberation has not been completely solved. This thesis aims at an effective SSL method based on the human hearing mechanism, which is able to work on binaural systems in practical noisy reverberant environments.

Binaural SSL is an important task in binaural signal processing field as it provides the location of sound source, commonly the direction of arrival (DOA) of the target sound. In the past decades, a large number of DOA estimation methods have been introduced, in which each one differs from others by the way of exploiting two main localization cues: the interaural time difference (ITD) and the interaural level difference (ILD). The well-known conventional GCC-PHAT method is based on only ITD and does not account well for noise. Therefore, there has been many research showing that it is not effective for binaural SSL. Azimuth-dependent models of binaural cues, such as joint estimations of ITD and ILD and DOA classification, have been presented. Although these research showed relatively good results by combining both ITD and ILD, their applicability in adverse noisy reverberant environments is still limited since there has been lack of methods accounting for the effect of interference signals efficiently. Methods directly based on head-related transfer functions (HRTFs) have also been studied, such as the inverse HRTF filtering and the cross-channels HRTFs. However, these methods highly depend on the HRTFs and suffer from reverberation because the HRTFs vary largely along the reverberation levels.

In psychoacoustic research field, binaural hearing has been studied for more than a century and several theoretical models of binaural processing have been developed. Among them, the equalization-cancellation (EC) model of Durlach has received a significant attention as its description is consistent with the human perception on binaural data. The EC model was originally proposed to explain the phenomenon of binaural masking-level differences (BMLDs) in binaural detection. Due to its well performance on BMLD prediction, the EC model was further extended to selective hearing in the 'cocktail party' scenarios. This suggested that the EC model has great potential for sound localization and segregation in the presence of multiple interference signals.

Inspired by the EC model, this thesis investigates a binaural DOA estimation method based on the EC mechanism. The principle idea is that the EC procedures are first utilized to eliminate the sound signal component at each interest direction; the direction of sound source is then determined as the direction at which the residual energy is minimal. In order to make this idea applicable in practice, two approaches are proposed to accommodate it with the problem of SSL under the effect of noise and reverberation, resulting in two improved algorithms namely Adaptive EC-BEAM and Weighted EC-BEAM. The Adaptive EC-BEAM algorithm improves SSL performance by adapting the EC model to the level of reverberation in room, using the direct-to-reverberant energy ratio (DRR). The Weighted EC-BEAM algorithm deals with the problem in a contradict way, in which two weighted functions are applied to reduce the negative effects from the observed signals, without modifying the localization model. Improvement of the suggested algorithms is verified by experimental results in various noisy reverberant conditions.

The proposed Weighted EC-BEAM algorithm is then selected to apply in two binaural applications, speech enhancement and source separation, as its assumption is easier to be satisfied in practice. In the first application, the

proposed method is employed to localize the meaningful sound signals for an intelligent speech enhancement system, which is able to extract and present the meaningful signals together with the target speech. The second application applies the proposed SSL method to estimate the DOAs of all sound sources before extraction (separation), resulting in a new blind source separation method. Experimental results showed that the Weighted EC-BEAM localized the desired sound sources correctly in both applications, from which the effectiveness of the proposed SSL method is confirmed

Keywords: Binaural sound localization, binaural hearing, Equalization-Cancellation model, noisy reverberant environments, humanoid robot