Title	A DOA Estimation Algorithm based on Equalization-Cancellation Theory and Its Applications
Author(s)	CHAU, Thanh Doc
Citation	
Issue Date	2010-09
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/9144
Rights	
Description	Supervisor:Masato Akagi, 情報科学研究科, 修士

A Binaural Sound Source Localization Approach Based on Equalization-Cancellation Theory and Its Applications

Chau Thanh Duc (0810205)

School of Information Science,
Japan Advanced Institute of Science and Technology

August 10, 2010

Keywords: sound source localization, equalization-cancellation, beamforming, speech enhancement, blind sound separation.

In this research, a binaural sound source localization (SSL) method based on Equalization-Cancellation (EC) theory and beamforming approach is proposed to estimates direction of arrival (DOA) of sound with only two microphones being robust under noisy environments and being adaptable to effect caused by recording system, such as head-related transfer function (HRTF). Secondly, the proposed method is applied to construct binaural applications in speech enhancement and blind source separation (BSS), resulting an intelligent speech enhancement system and a new directional blind source separation method.

SSL plays an important role in multi-channel signal processing systems. For example, in multi-channel noise reduction, the DOA information (steer vector) of the target signal is normally needed to compensate for the differences among different microphones. For blind source separation, in the case that the number of channels/microphones is less than the number of sound sources, blind beamforming technique is employed to extract individual sounds as an alternative method, in which sound localization is the key factor to detect the sound sources to be extracted.

Although the problem of sound localization has been studied for several decades, it is still a challenging problem in signal processing. In principle, approaches in this field localize sound source by simulating human and animals' perception mechanism to obtain localization information of sound. Due to the complexity of human auditory system, it is difficult to perfectively simulate these mechanisms. Therefore, microphone array is employed to obtain more location information for improving the accuracy of localization. This approach is, however, impractical for systems with small physical size, e.g. hearing aids. For some special systems on which the sounds received at two microphones are affected by the geometric shape, for example a robot head with the HRTF effect, localization performance is dramatically degraded if such effect is not considered. Moreover, in realistic environments where localization method is carried out, there are a lot of noises which may contribute significantly to the low performance of localization.

Regarding to human audition, location and direction of sound sources are specified based on the differences between sounds at two ears. These differences are known as binaural cues of signals. In signal processing, Cancellation is a common strategy for exploiting binaural cues. Several research have shown that the Cancellation process will be more effective if the Equalization is performed beforehand. These techniques are derived from a theory called Equalization-Cancellation (EC), which was firstly proposed Durlach and further improved by Culling and Summerfield. In EC model, when subject is presented with a binaural-masking stimulus, the auditory system attempts to eliminate the masking components by transforming the signal arriving at one ear relative to the signal at the other ear to make the masker components equalized (the E process). Then part of the signal in each ear is canceled by subtracting the signal in the other ear (the C process). Theoretically, the Cancellation process yields the binaural cues of signals, such as interaural-time difference (ITD) and interaural level difference (ILD), which are the most important cues for sound localization.

Inspired by EC theory, in this research, we propose a binaural sound source localization method, namely EC-BEAM, for DOA estimation by integrating EC model into beamforming strategy. Specifically, the Equalization process is applied to construct beamformer filters beforehand and these filters will be used in Cancellation process. To estimate DOA of sound, the proposed EC-BEAM steers null-beamformers to several predefined directions. For each null-beamformer at a specific direction, the Cancellation process is applied to cancel the signal coming from that direction, yielding remaining signals from other directions in beamformer output. It is noticed that if a beamformer steers to the biggest sound source, the energy of its output will be smallest. Finally, the DOA estimate is obtained via seeking the beamformer of which the output energy is minimum. The interpolation technique is further used to improve searching resolution without increase of computational expense. This enables the proposed EC-BEAM to localize sound source at even non-beamformed directions.

The effectiveness of the proposed EC-BEAM algorithm was firstly evaluated in term of localizing simulated data. Several experiments were carried out with data in several conditions and under two kinds of HRTF effect. Experimental results show that the proposed method with two microphones is able to localize accurately sound source and robust under noisy environments, especially, its performance is comparable under low noise condition. Moreover, the EC-BEAM achieves high results with both kinds of HRTF effect, which indicates that the proposed method potentially has adaptability to effects caused by geometric shape of recording systems. The superiority of EC-BEAM algorithm was further evaluated by comparing with the well-known GCC-PHAT algorithm. In the result, the estimation error of EC-BEAM algorithm is significantly lower than that of GCC-PHAT since in the original GCC-PHAT, there is no effect like HRTF was considered.

The effectiveness of EC-BEAM is also evaluated in term of implemented in binaural applications. Within this research, the problems of speech enhancement and blind source separation, which are closely related to sound localization, were considered.

Concerning speech enhancement, conventional approaches attempt to suppress the degradation factors of sound (e.g., noise and reverberation) and preserve target signal. A number of methods have been presented, in which the *two-stage binaural speech enhancement* (TS-BASE) model proposed by Li *et al.* was shown as high performance model. However, in daily communication, there are many important non-target signals need to be perceived by listener, for example, sound of telephone, sound of a call from someone. This problem has not been considered in the state-of-the-art of speech enhancement. In an effort to verify the applicability of EC-BEAM, we propose an intelligent speech enhancement model, namely iTS-BASE, based on EC-BEAM and TS-BASE by enhancing target signal and extracting meaningful signals at the same time. Specifically, there are two parallel processes in iTS-BASE: target signal enhancement and meaningful signal extraction. The first process is to enhance target signal by applying TS-BASE. The second process will detect extract meaningful sound by using EC-BEAM and extract the detected sound by employing TS-BASE again. Experimental results in some simple cases of meaningful sound show that in term of enhancing both target and meaningful signal, the iTS-BASE performed better TS-BASE.

Finally, the EC-BEAM was exploited to construct a directional blind source separation model to deal with this problem in the case the number of microphones is less than the number of sound sources. In this model, EC-BEAM is applied to estimate DOAs of multiple sources and TS-BASE is employed to extract (separate) detected sources. In the result, it is observed that the EC-BEAM performs localizing sound sources accurately and the extracted signals are quite similar to the original ones.

In summary, this thesis proposed a binaural sound source localization algorithm, namely EC-BEAM, based on equalization-cancellation theory and beamforming approach. The proposed method was evaluated on simulated data and applied in two binaural applications: speech enhancement and blind sound separation. Experimental result showed that, the proposed algorithm with two microphones can localize sound source accurately, be robust under noisy environments and has adaptability to some effects, such as HRTF. The two proposed intelligent speech enhancement and directional blind source separation showed that these systems are potential systems to be implemented in practice.