Title	
Author(s)	,
Citation	
Issue Date	2014-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/12051
Rights	
Description	Supervisor: , ,

Study on DOA estimation method based on human ability of sound localization

Masaru Ando (1210004)

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

February 12, 2014

Human beings have the ability for sound localization in our living environments. For example, we can easily localize a direction of coming car by utilizing car noise. In general, human beings use binaural cues to localize the target sound. In addition, it is reported that human beings can also localize the target sound by using monaural cues. Solving human ability for sound localization is a very important issue to reveal the mechanism of our hearing. If we can apply human ability for sound localization to the engineering problems, a method of estimating direction of arrival (DOA) of the target sound using monaural cues can be realized as applications of single-channel signal processing.

It is well known that the main cues for sound localization by binaural hearing are interaural time difference (ITD), interaural level difference (ILD), and spectral information. They are included in head-related transfer functions (HRTF). HRTF is transfer function between sound source and eardrum position in each ear. In these cues, available monaural cues for the sound localization can be regarded as spectral cues in HRTF such as peeks and notches in monaural spectral envelope. However, it is unclear how the peeks and notches in monaural spectral envelope vary with DOA of the sound source so that these cues cannot be used to estimate DOA of the sound source directly in this time.

On the other hand, there are some studies on binaural modulation cues for sound localization. Thompson and Dau reported that ILD and ITD in the temporal envelope are also important cues of the sound localization. Furukawa reported that the auditory system has separate mechanisms for processing ITD and ITD for a modulation rate up to about 100 Hz. And the information about modulation direction is preserved at the early lebel of auditory processing, at which ITD and ILD are individually processed. These reports suggest that the monaural modulation spectrum (MMS) can be regarded as an important cue of the monaural sound localization.

There is a related work with regard to DOA by the MMS approach. Kliper proposed a DOA estimation method using monaural cues in amplitude modulation patterns, based on machine learning scheme. They used the MMS patterns of the signals observed at eardrum. However, they employed machine learning scheme for classification of the MMS patterns to estimate azimuth of the sound source directly. Therefore, their method cannot account for a mechanism of monaural DOA. In particular, from their approach it is still unclear how the MMS patterns can be used for monaural sound localization.

This study aims to find out important monaural cues for sound localization and to propose a method of estimating DOA using these cues that this study founds. In this paper, we thus investigate how the MMS of the observed signals vary with the azimuth to find out monaural cues of DOA estimation. this study then proposes a method of estimating monaural DOA based on a concept of the modulation transfer function. This study investigated how the MMSs of the observed signals vary with the azimuth from 0 to 355 degrees. Then, the front of the head was 0 degrees. This study thus analyzed the MMSs of the observed signals in all conditions by computer simulations. Two types of the amplitude modulated (AM) signals were used in these simulations as the sound source. One was AM with sinusoidal carrier (AM signal) and the other was AM with white noise carrier (AM noise).

As results, it was found that the shapes of MMS variations with azimuth were drawn in the arc as a function of azimuth at the ear side. In addition, it was observed that MMS variations of effect of head diffraction at the ear reverse side when carrier frequency of source signal was changed.

These results suggest that humans may use some kind of cues based on the tendency of variation change of the MMS. However, it is necessary to carry out a listening experiment.

Then, this study proposed a method of estimating DOA based on these results of MMS analysis. Procedures of the proposed method are as follows. The MMSs are approximated by second order polynomials (regression curves). Then, an inverse function is derived by using the regression curves. Finally, azimuth is estimated by substituting MMS values and regression coefficients into this inverse function.

Simulations were carried out to verify the effectiveness of the proposed method. AM signal and AM noise were used in these simulations. Two azimuths were estimated as the positive and negative values derived by the inverse function. In the case of positive value, estimates were correct in the back of ear position while, in the case of negative value, estimates were correct in the front of ear position. However in the case of each reverse, estimates were incorrect. These false estimates were the front-back confusion. These results indicated that the proposed method can correctly estimate DOA using the MMS except with the front-back confusion discrimination. Moreover, this study then proposed a method of discrimination front-back confusion. Simulations showed the effectiveness of the proposed method.