2Pa2-4

Proceedings of Symposium on Ultrasonic Electronics, Vol. 32 (2011) pp. 79-80
8-10 November, 2011

Direction of Arrival Estimation Based on Delayed-sum Method Using a Pair of Microphone in Reverberation Environment
反射環境下における遅延和アレイ法に基づいた 2 つのマイクロフォンによる音源方向推定

Keiichi Zempo[†], Tadashi Ebihara, and Koichi Mizutani (Univ. Tsukuba)
善甫 啓一[†]，海老原 格，水谷 孝一 (筑波大院・シス情工)

1. Introduction

Recently, the microphone arrays have been used for estimating direction of arrival (DOA) to enhance target sound signal from other signals. Though there are many methods to estimate DOA, we focus on DOA estimation with small number of microphones by time difference calculated from cross-correlation, from view point of decreasing the hardware costs and signal processing costs. Moreover, applying cross-power spectrum phase (CSP) analysis has been found effective to detect DOA precisely regardless of the shape of the power spectra of the signal[1]. However, CSP analysis sometimes does not work to estimate correct DOA, especially in high reverberation environment due to existence of reflected sounds.

The Purpose of this research is to estimate the correct DOA in reverberation environment (*e.g.* in room). Authors have proposed the method to separate the direct and reflected sound by using the arriving distinction of earliness, that is to say, direct sound always arrives earlier than others. The proposed the method was discussed in numerical experiments using ray theory[2] and also experiments in anechoic chamber[3]. In this paper, we estimate the correct DOA in reverberation environment such as at car parking space or in room.

2. DOA Estimation of Direct Sound

The proposed signal processing offers candidates of a direct signal corresponding to each peak of the correlation result, and correlation results of delayed-sum signals for each candidate to find out the earliest signal. In consequence, we would detect the direct signal, and the corresponding correct DOA would be estimated. The flow chart of these operations is shown in **Fig.1**.

In this study, we use two of microphones and define the signals arriving at these microphone, $s_1(t)$, $s_2(t)$, as,

$$s_1(t) = \phi(t) + \hat{h} * \phi(t-\tau_1) + n_1(t) \ ,$$
$$s_2(t) = \phi(t-\tau_\mathrm{d}) + \hat{h} * \phi(t-\tau_2) + n_2(t) \ , \quad (1)$$

where $\phi(t)$ denotes the direct sound arriving at microphone array, $\hat{h}$ is the operator to calculate the first reflected sound by direct sound, and $n_1(t)$ and $n_2(t)$ denote non-first reflected sounds, which are reverberations.

If the amplitude of non-first reflected sounds, $n_1(t)$ and $n_2(t)$, was small enough, the CSP function of reserved signals, $R_{\mathrm{CSP}}(u; s_1, s_2)$, will become,

$$R_{\mathrm{CSP}}(u; s_1, s_2) = |\hat{h}|\delta(u-\tau_2) + |\hat{h}|^2\delta(u-(\tau_2-\tau_1)) + \delta(u-\tau_\mathrm{d}) + |\hat{h}|\delta(u-(\tau_\mathrm{d}-\tau_1)), \quad (2)$$

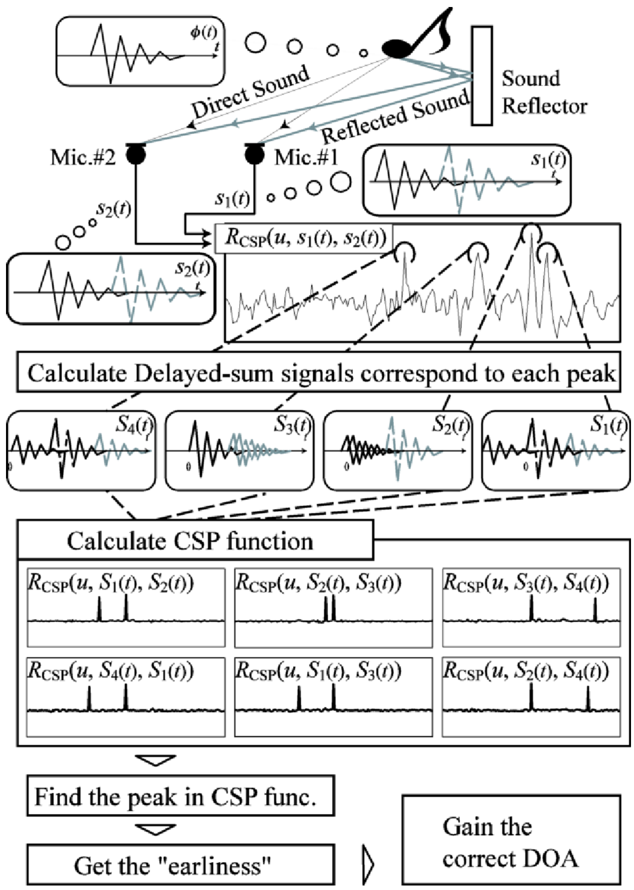

Fig. 1 Operation of DOA estimation in reverberation environment.

---

zempo@aclab.esys.tsukuba.ac.jp
{ebihara, mizutani}@iit.tsukuba.ac.jp

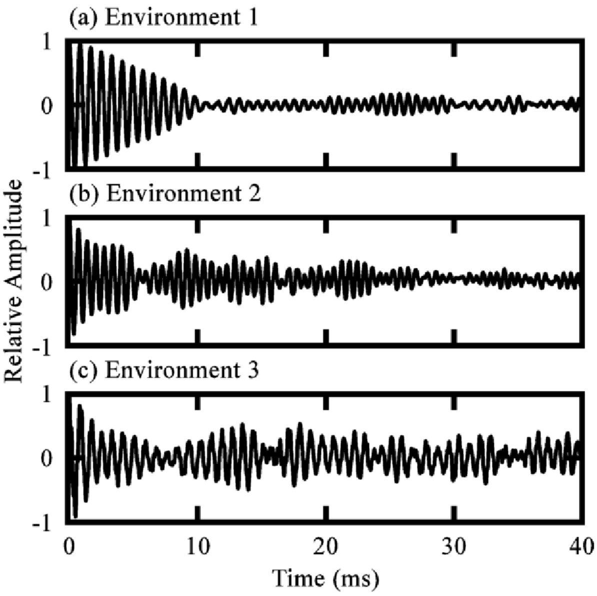

Fig. 2 Impulse response of each environment

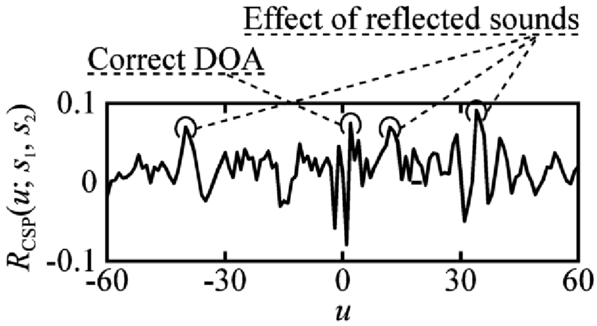

Fig. 3 CSP function which makes wrong estimation

where $\delta(u)$ is delta function, τ_d denotes the time difference between the direct signals of two microphones, τ_1 and τ_2 denote the time difference of arrival from reflected signal arriving at two microphones, and the direct signal arriving at microphone #1, respectively. By calculating the "earliness" of delayed-sum signal corresponding to the peaks, we gain the correct DOA.[2)]

3. Experiment setup

The experiments were held in three environments. (a) Environment 1: At car parking space where Speaker is near to ground surface, (b) Environment 2: At car parking space where Speaker is far from ground surface, (c) Environment 3: In office where there are many of tables, chairs and bookshelves. For each condition, there exist some noises like wind, air conditioner, PC, and engines. The impulse responses are shown in **Fig.2**.

We estimate the DOA of the voice of a male speaking ("Ko-n-ni-chi-wa") sent from a speaker with using microphone array. Sampling rate was 96 kHz. In normal CSP analysis, the peak corresponds to the correct DOA not always the highest peak in

Table I Four peaks, (u, R_{CSP} (u; s_1, s_2)), and "earliness" of corresponding delayed-sum signals.

(a) Environment 1

DOA	Correct	Effect of reflected sounds		
u	12	25	-20	45
R_{CSP} $(u;s_1,s_2)$	0.0397	0.0787	0.0340	0.0170
earliness	2nd	3rd	1st	4th

(b) Environment 2

DOA	Correct	Effect of reflected sounds		
u	6	15	-32	52
R_{CSP} $(u;s_1,s_2)$	0.0767	0.0733	0.0367	0.0450
earliness	2nd	3rd	1st	4th

(c) Environment 3

DOA	Correct	Effect of reflected sounds		
u	3	13	-41	34
R_{CSP} $(u;s_1,s_2)$	0.0313	0.0285	0.0441	0.0495
earliness	2nd	3rd	1st	4th

CSP function since the reflected sounds have high correlation with the direct sounds. We distinguish the peak corresponds to correct DOA from other peaks as shown in **Fig.3**.

4. Results

Four peaks in CSP function are written in **Table I** for each of condition. The value of $R_{CSP}(u; s_1, s_2)$ corresponds to the correct DOA does not always marks the highest value. However, the "earliness" of the delayed-sum signals corresponds to each peaks were able to be defined correctly. Accordingly, we successfully estimate the DOA of the direct sound separated from the effect of reflected sound

5. Conclusions

The purpose of this research is to estimate the DOA correctly in reverberation environment. Although the CSP analysis not always estimates the correct DOA, with the method to evaluate the "earliness" of signal, we would distinguish the correct DOA without the effect of reflected sounds.

References

1. T. Nishiura, *et. al.* , Sys. and Comp. in Jpn,34 (5), 69-80, 2003.
2. K. Zempo, T. Ebihara, K. Mizutani: Proceedings of ICSV18, R42-2193, 2011.
3. K. Zempo, T. Ebihara, K. Mizutani: Proceedings of Atmn Mtg ASJ, 1-P-1, 2011 (in press).