MULTIPLE BEAMFORMING WITH CSP-BASED SOURCE LOCALIZATION

Takanobu NISHIURA †‡, Satoshi NAKAMURA †, Kiyohiro SHIKANO ‡

†) ATR Spoken Language Translation Research Laboratories
2-2-2 Hikaridai Seika-cho Soraku-gun Kyoto, 619-0288 Japan
Fax: +81-744-95-1308
‡) Graduate School of Information Science, Nara Institute Science and Technology
8916-5 Takayama-cho Ikoma-shi Nara, 630-0101 Japan
Fax: +81-743-72-5289
E-mail: takanobu.nishiura@slt.atr.co.jp, nakamura@slt.atr.co.jp, shikano@is.aist-nara.ac.jp

ABSTRACT

In real environments, the presence of room reverberations seriously degrades sound capture quality. To solve this problem, multiple beamforming, which forms directivity not only in the direction of the target sound source but also in the direction of reflection images, has been proposed by J. Flanagan, et al. However, it is difficult to apply the method practically, since it requires an accurate source position and room geometry. This paper proposes a new multiple beamforming algorithm which requires no prior information about the source position or room geometry. The proposed algorithm achieves multiple beamforming by estimating the position of the target sound source and the major reflection images by CSP (Cross-power Spectrum Phase) analysis. Evaluation experiments based on the image method are conducted. The proposed algorithm performs more effectively than the conventional single beamformer in a room with large reflection coefficients. However, the proposed algorithm performs almost the same as conventional single beamforming in a room with small reflection energy.

KEYWORDS: Microphone array, Multiple beamforming, Reflection sounds, Source localization
INTRODUCTION

Sound capture of distant sound sources is indispensable for teleconference and voice control systems. However, background noises and room reverberations seriously degrade the sound capture quality of distant sound sources in real acoustical environments. A microphone array has been adopted as a promising tool to deal with this problem. Super directivity toward the target sound source is formed by a microphone array based on beamforming algorithms such as single beamforming [1], multiple beamforming [2] and so on.

Multiple beamforming has been proposed to improve the degradation of single beamforming. The multiple beamforming synchronizes the phase of the sound signals not only in the target sound direction but also in major reflection image directions. However, these algorithms require either a measurement of the actual impulse responses from the sound source to the transducers or an estimation of the impulse responses by the image method using information about the source position and room geometry.

This paper proposes a new multiple beamforming algorithm which requires no prior information about the source position or room geometry. The proposed algorithm achieves multiple beamforming by estimating the direction of the target sound source and the major reflection images with CSP (Cross power Spectrum Phase) [3] analysis.

Multiple beamforming with acoustic source localization

Multiple beamforming was proposed by J. Flanagan et al. to improve the performance of the delay-and-sum beamformer. The multiple beamformer forms beams not only in the target direction but also in the major reflection image directions. The images are conventionally regarded as an interfering noise. However, multiple beamforming utilizes these major reflection images as a target sound. Fig. 1 illustrates the concept of multiple beamforming. Multiple beamforming requires room impulse response information. Therefore, it is necessary to measure the actual impulse responses or to estimate the impulse responses based on the image method from information about the sound source position and room geometry. This paper proposes a new multiple beamforming algorithm which requires no information about the sound source position or room geometry. The proposed method is based on an estimation of the directions of the source and major reflection images. Fig. 2 shows an overview of the proposed algorithm.

---

Fig. 1 Multiple Beamforming.

Fig. 2 Algorithm Overview.
Estimation of Delay of Arrival  

The direction of the target sound source can be determined by estimating the DOA (Delay Of Arrival) from multiple transducer outputs. Although the cross correlation function is normally used to estimate DOA, CSP analysis [3] is adopted in this paper. CSP provides stable estimation results by whitening the magnitude spectrum. Let $s_i(n), s_k(n)$ be signals detected by the transducer $i, k$. CSP is obtained by equation (1). In this paper only the largest four peaks of CSP are taken into consideration.

$$csp_{ik}(\tau) = \text{DFT}^{-1} \left[ \frac{\text{DFT} [s_i(n)] \text{DFT} [s_k(n)]^*}{||\text{DFT} [s_i(n)]|| ||\text{DFT} [s_k(n)]||} \right]. \tag{1}$$

Source Localization by DOA Clustering  

The position of the sound source is estimated by DOAs from each transducer pair. Fig. 3 illustrates this situation. In this case, many source-position hypotheses arise not only in desired positions but also in undesired positions. To solve this problem, we focus on only the source and the first order reflection images, since the first order images are assumed to be the most effective in multiple beamforming. Then, we apply clustering not to the cross point hypotheses but to the DOAs from each transducer pair. The positions are estimated only by the DOAs in the same cluster. This algorithm alleviates errors caused by undesired cross points. We also remove all DOAs that are shorter than the DOA from the sound source in order to reduce clustering errors. The positions of the source and the images are estimated by taking an average of the cross point hypotheses in the same DOA cluster.

Evaluation Experiments

Experiment Conditions  

The room acoustic impulse responses are simulated by the image method [4]. We improved the image method by interpolating the sampling period by the sinc function. Fig. 4 illustrates a two dimensional room simulated in the experiments. Six positions, indicated A-F are evaluated. The microphone array is placed adjacent to the wall. The sampling frequency is 48kHz. The reflection orders are 20. White Gaussian noise is used as a source sound.

Source Localization Experiments  

The microphone array used for source localization is composed of 12 transducers with 15cm equal spacing. Ten transducer pairs with 30cm

---

Fig. 3 Estimation of source and image positions.

Fig. 4 Experiment room.
spacing are used for source localization. Fig. 5 shows the correct and estimated positions of the source and the reflection images for source position B. The estimated positions outside the room are the reflection images. This result shows that the proposed source localization algorithm precisely estimates not only the source direction but also the reflection image direction.

**Multiple Beamforming Experiments**  A microphone array with 56 transducers is used for beamforming. Fig. 6 shows an example of responses to an input impulse signal. As an evaluation measure to compare the performance of single and multiple beamforming, the SNR (Signal to Noise Ratio) is used. Fig. 6 shows average and min-max SNRs in six source positions by single and multiple beamforming for the 1st reflection images.

Multiple beamforming based on the proposed estimation algorithm further improved SNR compared with single beamforming.

**Conclusion**

This paper proposes a new multiple beamforming algorithm which requires no prior information about the source position or room geometry by estimating the direction of a target sound source and major reflections based on CSP analysis. Simulation experiments were carried out to evaluate the proposed algorithm. The results show that the proposed algorithm accurately estimates the positions of the source and reflection images. An SNR evaluation was also investigated. It was confirmed that multiple beamforming based on the estimated positions of the source and reflection images effectively improved the SNR compared with the single beamforming algorithm while there are some estimation errors.

**REFERENCES**