ACTIVE MICROPHONE WITH PARABOLIC REFLECTION BOARD FOR ESTIMATION OF SOUND SOURCE DIRECTION

Tetsuya Takiguchi, Ryoichi Takashima and *Yasuo Ariki*

Organization of Advanced Science and Technology
Kobe University
1-1 Rokkodai, Nada-ku, Kobe, 657-8501 Japan

ABSTRACT

This paper introduces a concept of an active microphone that achieves a good combination of active-operation and signal processing, where a new sound-source-direction estimation method using only a single microphone with a parabolic reflection board is proposed. In our previous work [1], we proposed GMM (Gaussian Mixture Model) separation for estimation of the sound source direction, where the observed (reverberant) speech is separated into the acoustic transfer function and the clean speech GMM. However, the previous method required the measurement of speech for each room environment in advance. The new proposed method using parabolic reflection is able to estimate the sound source direction without any prior measurements. Its effectiveness is confirmed by sound-source-direction estimation experiments on white noise in a room environment.

Index Terms— Direction of arrival estimation, Acoustic reflection, Microphones

1. INTRODUCTION

Many systems using microphone arrays have been tried in order to localize sound sources. Conventional techniques, such as MUSIC, CSP, and so on (e.g., [2, 3, 4]), use simultaneous phase information from microphone arrays to estimate the direction of the signal arrival. There have also been studies on binaural source localization based on interaural differences, such as interaural level difference and interaural time difference (e.g., [5]). Also, sound source localization techniques focusing on the auditory system have been described in [6, 7]. This paper presents a new sound-source-direction estimation method using only a single microphone with a parabolic reflection board.

The problem of single-microphone source separation is one of the most challenging scenarios in the field of signal processing, and some techniques have been described (e.g., [8, 9, 10]). In our previous work [1], we discussed a sound source localization method using only a single microphone.

In that report, the acoustic transfer function was estimated from observed (reverberant) speech using a clean speech model without texts of the user's utterance, where a GMM (Gaussian Mixture Model) was used to model the features of the clean speech. This estimation is performed in the cepstral domain employing a maximum-likelihood-based approach. This is possible because the cepstral parameters are an effective representation to retain useful clean speech information. The experiment results of our talker-localization showed its effectiveness. However, the previous method required the measurement of speech for each room environment in advance. Therefore, this paper presents a new method that uses parabolic reflection that is able to estimate the sound source direction without any need for prior measurements.

In everyday life, if an interesting sound is almost inaudible, people usually adjust the angle of their ears so that their ears are facing the sound direction. That means that signal processing is occurring as one changes the angle of his ear in order to better hear an interesting sound. However, the conventional thought was that even if the microphone is unable to pick up a sound well, its position does not change. In this paper, we introduce the concept of an active microphone that achieves a good combination of active-operation and signal processing. The active microphone has a parabolic reflection board, which is extremely simple in construction. The reflector and its associated microphone rotate together, perform signal processing, and seek to locate the direction of the sound source, so that parabolic reflection can help to increase the power gain of the signal arriving at the front of the parabolic reflector.

2. ACTIVE MICROPHONE

2.1. Parabolic Reflection Board

In this paper, an active microphone with a parabolic reflection board is proposed for estimation of sound source direction, where the reflection board has a form of parabolic surface. As shown in Fig. 1, under the assumption of the plane wave, any parallel line (wave) for the axis of the parabolic surface is reflected toward the focal point. On the other hand, if the sound source is not located at 90 degrees (in front of the parabolic surface), no reflection wave will travel toward the focal point.

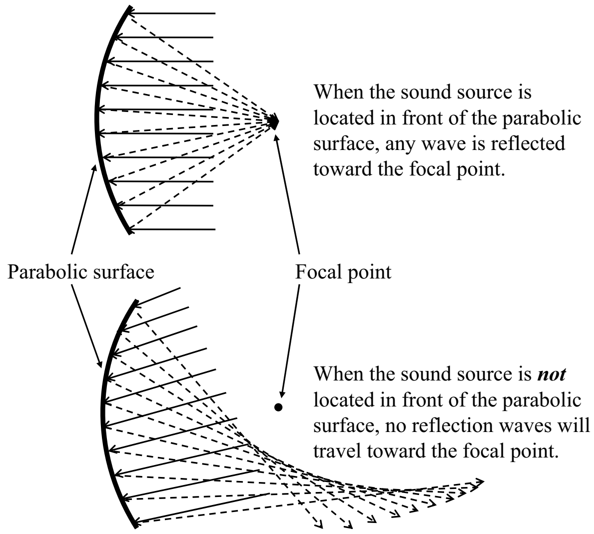

Fig. 1. Concept of parabolic reflection

Therefore, the use of the parabolic reflection board will be able to give us the difference of the power between the target direction and the non-target direction.

2.2. Signal Power Observed by Parabolic Reflection

Next, we consider the signal power observed using parabolic reflection. As shown in Fig. 2, the observed signal at the focal point can be expressed by the addition of the waves $s1$ and $s2$, where only two paths are considered for simplicity. The signal $s2$ reflects at a point P and arrives at the focal point. Therefore, the distance difference between path $s1$ and $s2$ to the focal point can be expressed as follows:

$$\begin{aligned} QP + PO &= QP + PH \\ &= 2d \end{aligned} \tag{1}$$

where the parabola is defined as $PH = PO$ and d represents the distance of the focal point. Therefore, the time difference to the focal point between the direct and reflection waves is given by

$$\tau = 2d/a. \tag{2}$$

Here a represents the speed of sound. The above equation shows that the time difference depends only on d. Therefore, we can understand that the time difference for all reflection paths (not only $s2$) is equal to $2d/a$. (There is no delay among reflections.) Therefore, assuming no background noise, the observed signal at time t, $x(t)$, is given by

$$x(t) = s(t) + A \cdot s(t - \tau) \tag{3}$$

where $s(t)$, $s(t - \tau)$ and A represent the direct sound, the reflection sound, and the reflection coefficient, respectively.

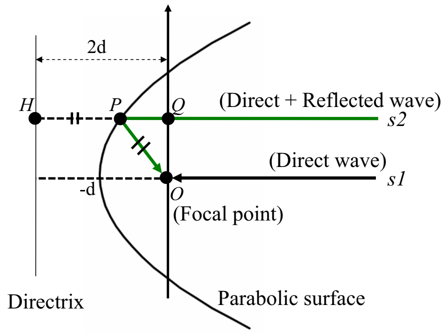

Fig. 2. Observed signal at the focal point, where the input signal is coming from directly in front of the parabolic surface

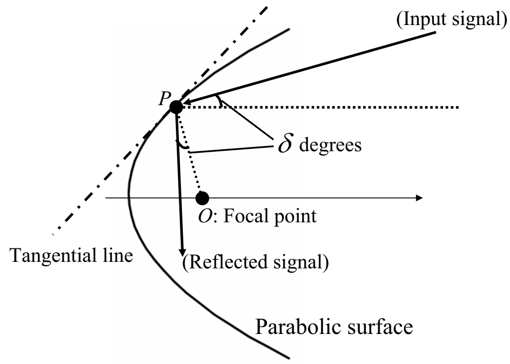

Fig. 3. Observed signal at the focal point, where the input signal is coming from δ degrees

By applying the Fourier transform, the observed signal in the frequency domain is given by

$$X(\omega) = S(\omega) + A \cdot e^{-j2\pi\omega\tau} S(\omega). \tag{4}$$

The power spectrum is then obtained as follows:

$$\begin{aligned} |X(\omega)|^2 &= |S(\omega)|^2 \cdot |1 + A \cdot e^{-j2\pi\omega\tau}|^2 \\ &= |S(\omega)|^2 \cdot H(\omega) \end{aligned} \tag{5}$$

As shown in equation (5), the use of parabolic reflection can increase the power gain of the signal arriving from directly in front of the parabolic board according to $H(\omega)$.

On the other hand, as shown in Fig 3, when the input signal is coming from δ degrees (not coming from directly in front of the parabolic surface), the direction of the reflected signal at the parabolic surface is off δ degrees from PO. Therefore, when the sound source is not located in front of the parabolic surface, the power gain will not increase.

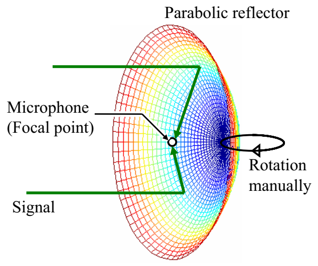

Fig. 4. Active microphone with parabolic reflection

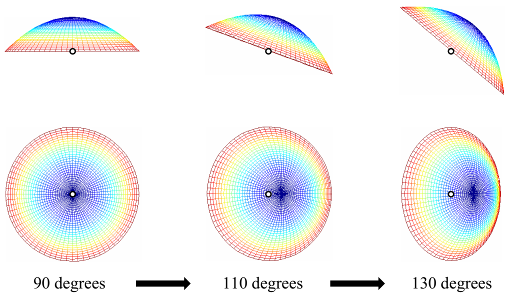

Fig. 5. Rotated active microphone

2.3. Estimation of Sound Source Direction

As shown in Fig. 4, a new active microphone with the parabolic reflection board was built, with the microphone located at the focal point. In order to obtain the power of the target signal observed at each angle, the angle of the microphone with the parabolic reflection was changed manually in research carried out for this paper (Fig. 5).

By applying the short-term Fourier transform to the target signal observed at a microphone angle, i, we obtain the power spectrum $|X_i(\omega; m)|^2$ at frame m. In this paper, the total number of frames at each direction is the same. The average power of the target signal at each angle was calculated, and the direction having maximum power was selected as the sound source direction:

$$\hat{i} = \underset{i}{\operatorname{argmax}} \sum_{m} \sum_{\omega} \log |X_i(\omega; m)|^2. \qquad (6)$$

Here, i is the angle of the parabolic reflection board (microphone).

3. EXPERIMENT RESULTS

The direction estimation experiment was carried out in a real room environment. The microphone with parabolic reflection shown in Fig. 4 was used for the experiments. The diameter was 12 cm, and the distance of the focal point was 3 cm. The

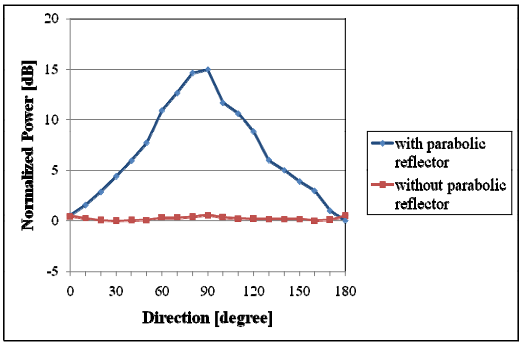

Fig. 6. Comparison of the performance with/without parabolic reflection

microphone located at the focal point is an omnidirectional type (SONY ECM-77B).

The target sound source is located at 90 degrees. The source signals are white noise with a length of about 5 seconds and the distance from the sound source to the microphone 40 cm, 70 cm, and 100 cm. The angle of the microphone with parabolic reflection was changed manually from 0 degrees to 180 degrees at an interval of 10 degrees. Then the average power of the target signal at each angle was calculated in five seconds. The sampling rate was 48 kHz, and a 32 msec Hamming window was used every 16 msec.

3.1. Performance of parabolic reflection

Fig. 6 shows a comparison of the performance with and without parabolic reflection. The source signal was white noise and the average power of the target signal at each angle was calculated in five seconds. The power was normalized so that the minimum was 0 dB. The distance from the sound source to the microphone was 70 cm. As shown in Fig. 6, the performance results of the proposed method (using parabolic reflection) are better than those obtained when parabolic reflection was not used (using only an omnidirectional microphone). The use of parabolic reflection improves the normalized power to 15 dB at 90 degrees, and the power decreases as the direction of the microphone becomes farther from the direction of the target sound source. On the other hand, when parabolic reflection is not used, the power remains almost unchanged.

3.2. Performance on different sound-source distances

In this section we describe a series of experiments performed on different sound-source distances using two active microphones having different size parabolic reflectors. One was the same one used in Section 3.1 (a small parabola), and the other was a parabolic reflector with a diameter of 24 cm and focal point distance of 9 cm (a large parabola).

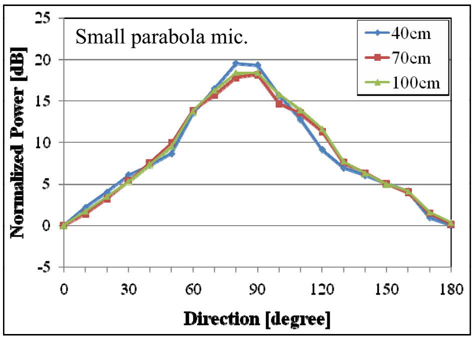

Fig. 7. Average power for different distances using a small parabolic reflector

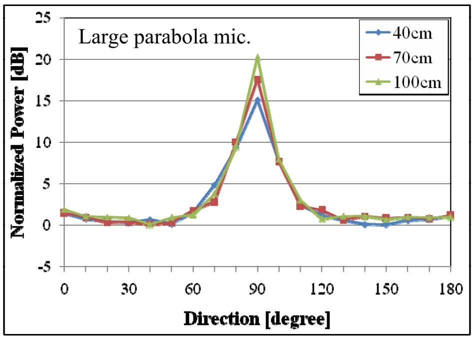

Fig. 8. Average power for different distances using a large parabolic reflector

Fig. 7 and Fig. 8 show that the average power of the target signal obtained for each angle for the small parabolic reflector and large parabolic reflector, respectively. The shape of the average power for the large parabola sharpens up at 90 degrees in comparison with that with small parabola because of the difference of the focal distance of parabolic reflectors.

Next, we will consider the power gain in relation to the distance of the sound source from the microphone and parabolic reflector. When the large parabola was used, the power gain increased as the distance increased, with the graph taking on a very sharp shape. This is thought to be due to the fact that the sound source does not form the ideal sound wave - a plane wave - that is created when the distance from the sound source is short. Even though the parabolic reflection board was turned directly toward the sound source, the sound waves were not reflected well.

4. CONCLUSION

This paper has described a sound-source-direction estimation method using a single microphone. The new proposed method using the parabolic reflection is able to estimate the sound source direction without any measurement in advance. In this paper, the experiment results in a room environment confirmed its effectiveness for a white noise signal. In future work, we will investigate the performance in noisy environments, such as with multiple sound sources.

5. REFERENCES

[1] T. Takiguchi, Y. Sumida, and Y. Ariki, "Estimation of Room Acoustic Transfer Function Using Speech Model," *IEEE Statistical Signal Processing Workshop*, pp. 336-340, 2007.

[2] D. Johnson and D. Dudgeon, "Array Signal Processing," Prentice Hall, 1996.

[3] M. Omologo and P. Svaizer, "Acoustic Event Localization in Noisy and Reverberant Environment Using CSP Analysis," Proc. ICASSP96, pp. 921-924, 1996.

[4] F. Asano, H. Asoh and T. Matsui, "Sound Source Localization and Separation in Near Field," IEICE Trans. Fundamentals, Vol. E83-A, No. 11, pp. 2286-2294, 2000.

[5] F. Keyrouz, Y. Naous and K. Diepold, "A New Method for Binaural 3-D Localization Based on HRTFs," Proc. ICASSP06, pp. V-341-V-344, 2006.

[6] O. Ichikawa, T. Takiguchi and M. Nishimura, "Sound Source Localization Using a Profile Fitting Method with Sound Reflectors," IEICE trans. on inf. and sys. E87-D(5), pp. 1138-1145, 2004.

[7] N. Ono, Y. Zaitsu, T. Nomiyama, A. Kimachi, and S. Ando, "Biomimicry Sound Source Localization with Fishbone," IEEJ Trans. Sensors and Micromachines, vol. 121-E, no. 6, pp. 313-319, 2001.

[8] T. Kristjansson, H. Attias and J. Hershey, "Single Microphone Source Separation Using High Resolution Signal Reconstruction," Proc. ICASSP04, pp. 817-820, 2004.

[9] B. Raj, M. V. S. Shashanka and P. Smaragdis, "Latent Direchlet Decomposition for Single Channel Speaker Separation," Proc. ICASSP06, pp. 821-824, 2006.

[10] T. Nakatani, B.-H. Juang, "Speech Dereverberation Based on Probabilistic Models of Source and Room Acoustics," Proc. ICASSP06, pp. I-821-I-824, 2006.