2008 International Workshop on Nonlinear Circuits and Signal Processing
NCSP'08, Gold Coast, Australia, March 6-8, 2008

Estimation of Sound Source Direction Using Parabolic Reflection Board

Tetsuya Takiguchi†, Ryoichi Takashima‡ and Yasuo Ariki†

†Organization of Advanced Science and Technology
Kobe University
1-1 Rokkodai, Nada-ku, Kobe, 657-8501 Japan
Email: takigu@kobe-u.ac.jp Email: ariki@kobe-u.ac.jp

‡Graduate School of Engineering
Kobe University
1-1 Rokkodai, Nada-ku, Kobe, 657-8501 Japan
Email: takashima@me.cs.scitec.kobe-u.ac.jp

Abstract

This paper presents a new sound-source-direction estimation method using only a single microphone with a parabolic reflection board. 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 the parabolic reflection is able to estimate the sound source direction without any measurement in advance. Its effectiveness is confirmed by sound-source-direction estimation experiments in a room environment.

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, 5]), 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., [6, 7]). Also, sound source localization techniques focusing on the auditory system have been described in [8, 9]. 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 signal processing, and some techniques have been described (e.g., [10, 11, 12, 13]). In our previous work [1], we discussed a talker localization method using only a single microphone.

In [1], the acoustic transfer function is estimated from observed (reverberant) speech using a clean speech model without texts of user's utterance, where a GMM (Gaussian Mixture Model) is used to model the feature 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 the new method using the parabolic reflection that is able to estimate the sound source direction without any measurement in advance.

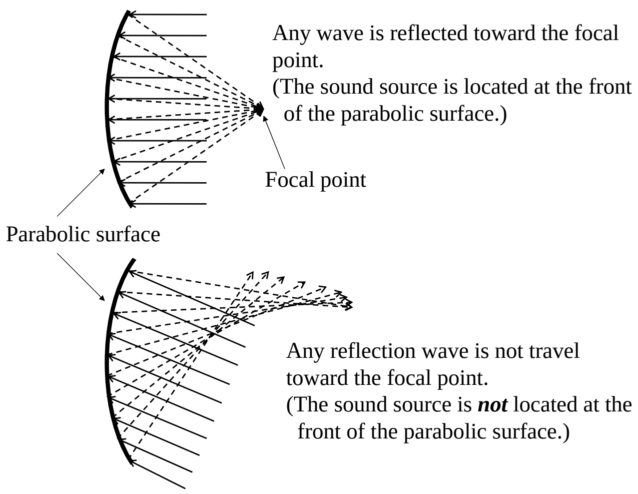

Figure 1: Concept of the parabolic reflection.

2. Parabolic Reflection Board

In this paper, a new active microphone with a parabolic reflection board is proposed for estimation of sound source direction, where the proposed 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), any reflection wave will not travel toward the focal point. Therefore, the use of the parabolic reflection board will be able to give us the differ-

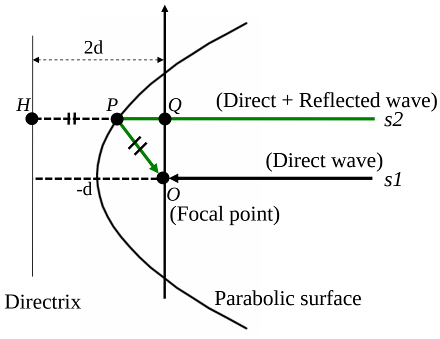

Figure 2: Observation signal at the focal point, where the input signal is coming from the front of the parabolic surface.

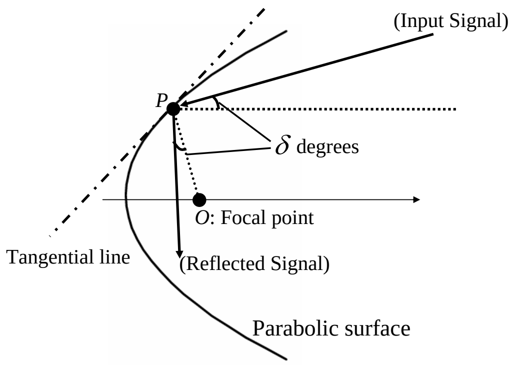

Figure 3: Observation signal at the focal point, where the input signal is coming from δ degrees.

ence of the power between the target direction and the non-target direction.

2.1. Signal Power Observed by Parabolic Reflection

Next, we consider about the signal power observed using the parabolic reflection. As shown in Fig. 2, the observation signal at the focal point can be expressed by the addition of the wave $s1$ and $s2$. The signal $s2$ reflects at a point P and arrives at the focal point. Therefore, the distance difference between the path $s1$ and $s2$ to the focal point can be expressed as follows:

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

where we use the definition of the parabola, $PH = PO$, 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

$$\tau = 2d/a. \quad (2)$$

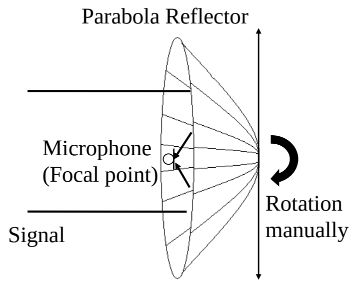

Figure 4: Active microphone with the parabolic reflection.

Here a represents the speed of sound. The above equation shows the time difference depends on only 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 reflection.) Therefore, under the assumption of no background noise, the observed signal at time t, $x(t)$, is given by

$$x(t) = s(t) + A \cdot s(t - \tau) \quad (3)$$

where $s(t)$, $s(t-\tau)$ and A represent the direct sound, the reflection sound, and the reflection coefficient, respectively. By applying the Fourier transform, the observed signal in the frequency domain is given by

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

The power spectrum is obtained as follows.

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

As shown in equation (5), the use of the parabolic reflection can increase the power gain of the signal arrived from the 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 the front of the parabolic surface), the direction of the reflected signal at the parabolic surface is out of δ degrees from PO.

2.2. Active Microphone

In this paper, as shown in Fig. 4, a new active microphone with the parabolic reflection board is built, where the microphone is 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 changes manually in 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.

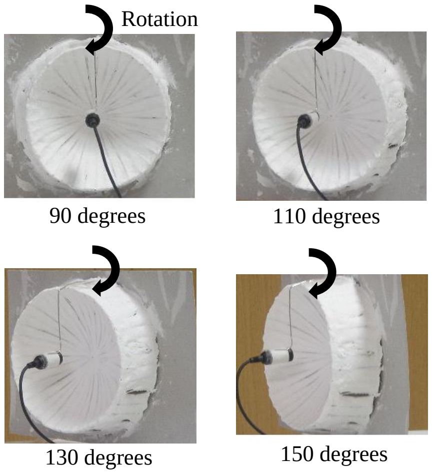

Figure 5: Rotated active microphone.

Then, the average power of the target signal at each angle is calculated. The direction having maximum power is selected as the sound source direction:

$$\hat{i} \;=\; \underset{i}{\operatorname{argmax}} \frac{1}{\Omega M} \sum_{m=1}^{M} \sum_{\omega=1}^{\Omega} \log |X_i(\omega; m)|^2. \qquad (6)$$

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

3. Experiment Result

Experiment of direction estimation was carried out in a real room environment. The microphone with the parabolic reflection shown in Fig. 4 is used for the experiments. The diameter is 12 cm, and the distance of the focal point is 3 cm. The microphone located at the focal point is the directional type. The directivity of the microphone is set up toward the parabolic surface because the reflection signal from

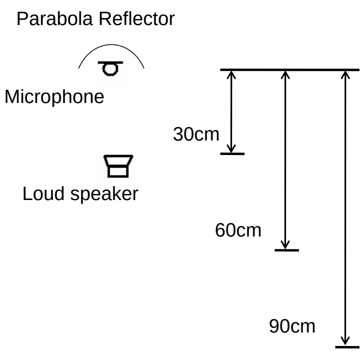

Figure 6: Experiment condition.

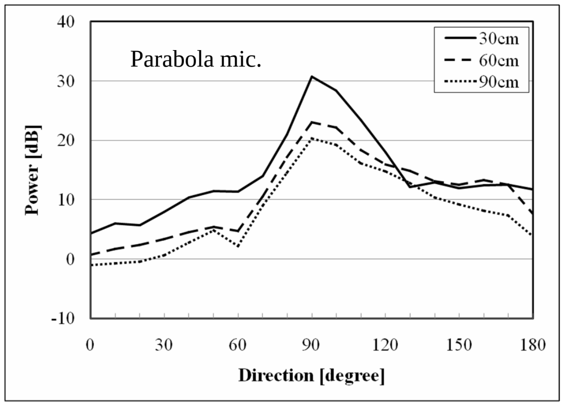

Figure 7: Average power versus angle of microphone.

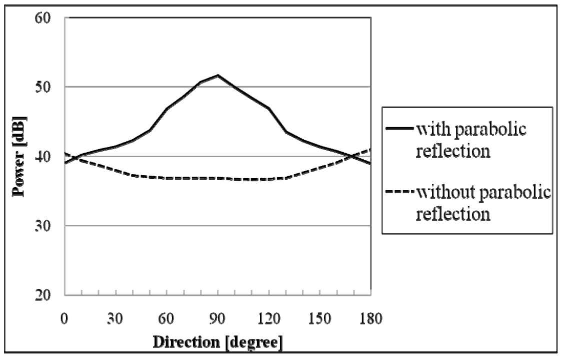

Figure 8: Comparison of the performance with/without the parabolic reflection.

the parabolic surface is more important than the direct signal in the proposed method.

The target sound source is located at 90 degrees and the source signal is the white noise (length: about 5 seconds). The distance from the sound source to the microphone is 30 cm, 60 cm, and 90 cm (Fig. 6). The angle of the microphone with the parabolic reflection changes manually from 0 degree to 180 degrees at an interval of 10 degrees. Then the average power of the target signal at each angle is calculated in five seconds. The sampling rate is 48 kHz, and a 32 msec Hamming window is used every 16 msec.

Fig. 7 shows the average log-power spectrum over the time-frequency domain versus the angle of the parabolic reflection. The experiment results show the average log-power spectrum at 90 degrees is the maximum value and the power is decreasing as the direction of the microphone with the parabolic reflection is farther from the direction of the target sound source.

Fig. 8 shows a comparison of the performance with / with-

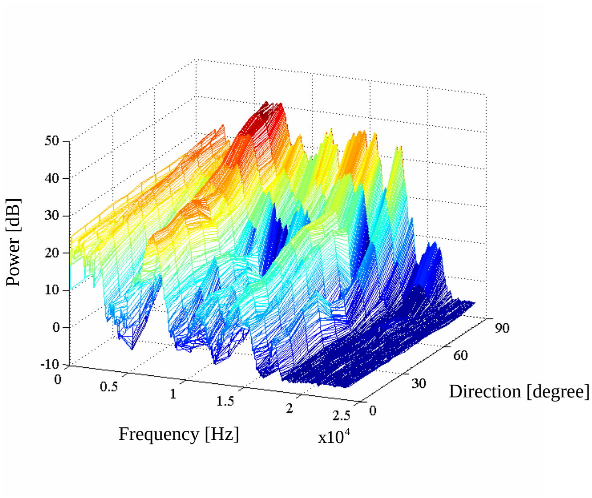

Figure 9: 3D power spectrum for each angle of the parabolic reflection.

out the parabolic reflection. As shown in Fig. 8, in comparison of the result of the proposed method (using the parabolic reflection) with that without the parabolic reflection (only directional mic.), the performance becomes better. In the case without the parabolic reflection, as the directivity of the microphone is set up opposite the sound source, the observed signal at 90 degrees causes a small decrease in the average power.

Fig. 9 shows the 3D power spectrum of the observed signal for each angle of the parabolic reflection. The result indicates that the effect provided by the parabolic reflection is not so much for the low-frequency components of the signal. This may be caused by the diffraction of the sound wave. The power spectrum except the low-frequency one becomes larger as the angle of the parabolic reflection is closer to 90 degrees.

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 the white noise signal. In future work, we will investigate about short signal (ex. speech) and the form of the parabolic reflection.

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] Y. Denda, T. Nishiura and Y. Yamashita, "Robust Talker Direction Estimation Based on Weighted CSP Analysis and Maximum Likelihood Estimation," IEICE Trans. on Information and Systems, Vol. E89-D, No. 3, pp. 1050-1057, 2006.

[6] 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.

[7] M. Takimoto, T. Nishino and K. Takeda, "Estimation of a talker and listener's positions in a car using binaural signals," The Fourth Joint Meeting ASA and ASJ, 3pSP33, p. 3216, 2006.

[8] 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.

[9] 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.

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

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

[12] G.-J. Jang. T.-W. Lee and Y.-H. Oh, "A Subspace Approach to Single Channel Signal Separation Using Maximum Likelihood Weighting Filters," Proc. ICASSP03, pp. 45-48, 2003.

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