Applying the Filtered Back-Projection Method to Extract Signal at Specific Position

Chia-Ming Chang and Chun-Hao Peng

Department of Computer Science and Engineering,
Tatung University,
Taipei, Taiwan
cmchang@ttu.edu.tw, g9106007@mail.ttu.edu.tw

Abstract— **In this paper, a back-projection method of extracting sound signal from multiple sources using microphone array is proposed. It is assumed that the signals received by microphone may be treated as a kind of non-straight projection. Therefore, the filtered back-projection method could be used to reconstruct the sound signal.**

The typical method used to enhance signal come from the focus location in multi-sources environment is the delay-and-sum method, whhich is based on the delay and shift operations. The filtered back-projection method usually is used to reconstruct the profile in the Computerized Tomography (CT).

The algorithm to apply filtered back-projection method to extract the signal located at specific position is dressed in the paper. Experiments are also presented to prove the fact that filtered back-projection method could be used to extract the location-specified signal from the multi-source environment.

1. Introduction

Microphone array is applied in many occasions nowadays. For example, concert, conference, speech, recording studio, etc [1]. Typically, delay and sum method is used for microphone array to enhance the signal come from specific location [2,3]. We propose a different method in which the same function based on the computerized tomography technique is used. Microphone array receives the mixed signal from all sources. The signal received by microphone can be treated as a special projection. The signal from specific location can be reconstructed using the filtered back-projection method. Then, the signal from the location we focus on is extracted by the system.

The organization of this paper is as follows. In Section 2, the delat-and-sum method and filtered back-projection method are addressed in brief. The discrete version of sound projection slice theorem is stated briefly in Section 3. The problems about the implementation of back-projection method to extract interested signal is discussed in Section 4. in Section 5, conclusions are addressed.

2. Relative researches

The typical method of enhancing signal received by microphone array is the delay-and-sum method. The signal comes from a specific position are received by microphone array. These received signals are different intensity and time delay. The differences are due to the difference in distance from the specified location to each microphone as shown in Fig. 1. Let L_k represent the distance from source to the k-th microphone, and L_{min} be the minimum distance of L_k. The delay time and wave propagation at-

Figure 1: Delay-and-sum method

tenuation are obtained from the follow equations,

$$D_k = (L_k - L_{min})/v \quad (\text{sec}) \tag{1}$$

$$A_k = L_k^2/L_{min}^2 \quad (\text{amplitude}), \tag{2}$$

where v is the velocity of sound. Applying D_k and A_k to the following equation, we can enhance the signal from the specific location, $S_e(t)$,

$$S_e(t) = \frac{1}{N}\sum_{k=1}^{N} A_k S_k(t + D_k), \tag{3}$$

where N is the numbers of microphone, and S_k is the signal received by the k-th microphone.

Computerized Tomography (CT) is a useful technique to reconstruct the profile image of an object without destroying it from the projections. A projection is a shadowgram obtained by illuminating an object by penetrating radiation.

The Radon transform and its inverse were formulated by J. Radon in 1917 [4]. The Radon transform provides the mathematical framework necessary for forth and going back between the spatial coordinates $\{x, y\}$ and the projection space coordinates $\{s, \theta\}$. The Radon transform of a function $f(x, y)$ at angle θ, denoted as $g(s, \theta)$, is defined as the line integral along a line inclined at an angle from the u-axis and at a distance s from the origin as shown in Fig. 2. The Radon transform is defined as

$$g(s, \theta) \stackrel{\triangle}{=} \mathcal{R}\{f\}$$
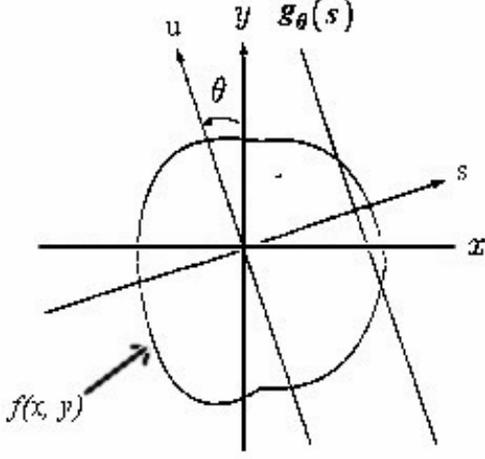
Figure 2: The Radon geometry

$$= \iint f(x,y)\delta(x\cos\theta + y\sin\theta - s)\,dxdy, \quad (4)$$

where $-\infty < s < \infty$ and $0 \le \theta < \pi$. The symbol $\mathcal{R}$, denoting the Radon transform operator, is also called the projection operator. In the rotated coordinate system (s, u), the relationship to (x, y) are

$$s = \quad x\cos\theta + y\sin\theta \quad (5)$$
$$u = -x\sin\theta + y\cos\theta. \quad (6)$$

Associated with the Radon transform is the back-projection operator denoted as β and is defined as

$$b(x,y) \overset{\Delta}{=} \beta\{g(s,\theta)\}$$
$$= \int_0^\pi g(x\cos\theta + y\sin\theta, \theta)\,d\theta. \quad (7)$$

The quantity $b(x,y)$ is called the back projection of $g(s,\theta)$. In polar coordinates it can be written as

$$b_p(s,\phi) = \int_0^\pi g(r\cos(\theta - \phi), \theta)\,d\theta \quad (8)$$

Eqs. (7) and (8) represent the accumulation of the ray-sums of all of the rays that pass through the point (x, y). For example, if

$$g(s,\theta) = g_1(s)\delta(\theta - \theta_1) + g_2(s)\delta(\theta - \theta_2) \quad (9)$$

that is, there are only two projections, then

$$b_p(r,\phi) = g_1(s_1) + g_2(s_2), \quad (10)$$

where $s_1 = r\cos(\theta_1 - \phi)$ and $s_2 = r\cos(\theta_2 - \phi)$.

In general, for a fixed point (x, y) or (r, ϕ), the value of back projection $\beta\{g\}$ is evaluated by integrating $g(s,\theta)$ over θ for all lines pass through that point.

The one-dimensional Fourier transform with respect to s of the projection $g(s,\theta)$ can be obtained from Eq. (4)

$$G_\theta = \mathcal{F}_1\{g_\theta(s)\}$$
$$= \iint f(s\cos\theta - u\sin\theta, s\sin\theta + u\cos\theta)$$
$$\cdot e^{-j2\pi rs}\,dsdu \quad (11)$$
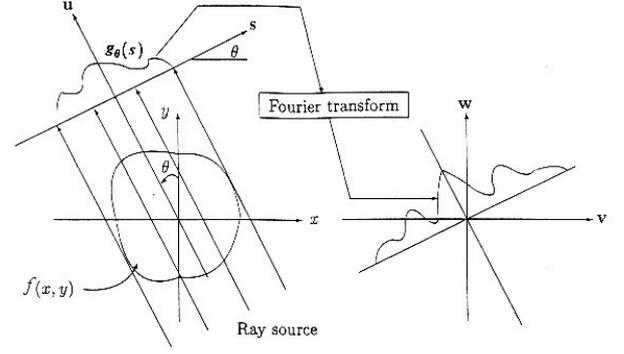
Figure 3: The Radon geometry

where $\mathcal{F}_1$ is the one-dimensional forward Fourier transformation and $G_\theta(r)$ is the Fourier transform of $g_\theta(s)$. Rotating the coordinate from $\{s, u\}$ to $\{x, y\}$ then Eq. (11) becomes

$$G_\theta(r) = \iint f(x,y)e^{-j2\pi(xr\cos\theta + yr\sin\theta)}\,dxdy$$
$$= F(r\cos\theta, r\sin\theta) \quad (12)$$

where $F(r\cos\theta, r\sin\theta)$ is a slice of two-dimensional Fourier transform of tomogram at angle θ as shown in Fig. 3.

If $0 \le \theta < \pi$ and $-\infty < r < \infty$, $F(r\cos\theta, r\sin\theta)$ represents the two-dimensional Fourier transform of the tomogram in polar form. The tomogram can be obtained by applying the two-dimensional inverse Fourier transform in polar form, that is,

$$f(x,y) = \int_0^\pi \int_{-\infty}^\infty F(r\cos\theta, r\sin\theta)$$
$$\cdot |r|e^{j2\pi r(x\cos\theta + y\sin\theta)}\,drd\theta \quad (13)$$

where $\int_{-\infty}^\infty \bullet\, dr$ represents the one-dimensional inverse Fourier transform, $\int_0^\pi g\,d\theta$ denotes the back projection as Eq. (8), and $F(r\cos\theta, r\sin\theta)|r|$ represent the filter operation in spatial domain. Combining the equations above, we have

$$f(x,y) = \int_0^\pi \mathcal{F}_1^{-1}\{\mathcal{F}_1\{g_\theta(s)\}|r|\}\,d\theta, \quad (14)$$

where $g_\theta(s)$ is the projection of $f(x,y)$ at angle θ, the $\mathcal{F}_1$ and $\mathcal{F}_1^{-1}$ are the one-dimensional Fourier and inverse Fourier transforms, respectively. And $|r|$ is the back-projection filter in frequency domain and $s = x\cos\theta + y\sin\theta$. The slice information we have are discrete in angle. Then, the Eq. (14) in discrete form is

$$f(x,y) = \sum_{0 \le \theta < \pi} \mathcal{F}_1^{-1}\{\mathcal{F}_1\{g_\theta(s)\}|r|\}. \quad (15)$$

Therefore, profile of the object, that is $f(x,y)$, could be reconstructed by summation of each filtered slice $g_\theta(s)$.

3. Sound Projection

We may consider that if the sound signal received by each microphone is a kind of projection, a non-straight projection, unlike the x-ray is doing. Then we can use the
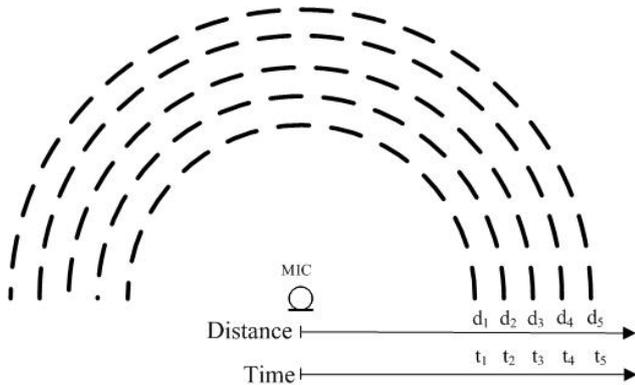
Figure 4: The sound projection.

Figure 5: System architecture.

back-projection technique to reconstruct the sound source signal. The signals come from point sources located on the concentric circules are received by microphones at the same time if the microphones is at the center. In other word, if the distances from sources to one microphone are the same, signal from different sources on the same concentric circule reach to the microphone at the same time.

Fig. 4 shows the relation between distance from microphone and received time, where d_k is the radius from center, the position of microphone, and t_k is the time that wave propagated to microphone. The received signals at time $T+t_k$ is the summation on radius d_k and is multiplied by a constant attenuation coefficient. We may consider that would exist a sound projection slice theorem, then the back-projection operation of sound could be rewrite from Eq. (15) and obtain

$$s(T) = \sum_{0 \leq k < M} \mathcal{F}_1^{-1}\{\mathcal{F}_1\{m_k(T)\}|\omega|\}, \qquad (16)$$

where T is a time period, $s(t)$ is the source signal, $m_k(t)$ is the signal received by the kth microphone, $|\omega|$ is the back-projection filter in frequency domain, and M is the size of microphone array. Next, we use the Eq. (16) to separate the source we focus on from mixed-signal received by microphone array.

4. Implementation

For the limitation of hardware, we use Matlab to simulate all of the system. The system structure is shown in Fig. 5. Assume there is a conference room with 4 meters width and 5 meters length without reverberant. A 16-channel microphone array is linearly arranged on the bottom side. Two point sources could be placed at any place in this room. The sine wave is used as the sound source in the begining of simulation.

The main part of the program is the filtered back-projection operation. The sequence is choosing the proper region of signal received from each microphone and filtering it, then average the filtered signal of each microphone in time domain like delay and sum method.

In Fig. 6, the simulation of filtered back-projection method with two sine wave signal sources is shown. The signals of two sources is the same as shown at the up-left plot in Fig. 6. The microphone receive the mixed-signal from source (the signal is interested) and noise (the signal isn't interested) is shown in bottom-left plot in Fig. 6. The
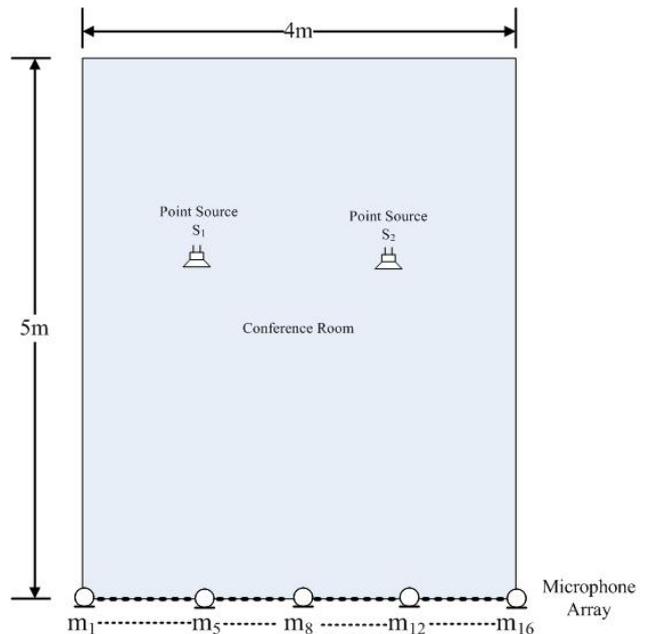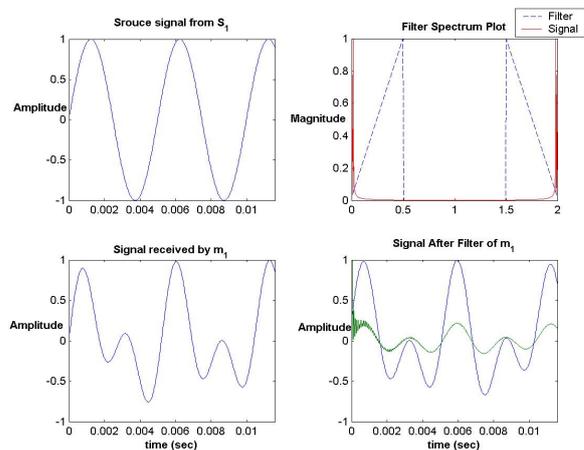
Figure 6: Simulation of filtered back-projection method with two sound sources.

result in spacial domain is shown at bottom-right plot in Fig. 6.

The equation used to measure the efficiency to extract the interesting signal is shown as

$$SNR_{signal} = 20\log\left(\frac{s}{s-r}\right) \quad \text{(dB)}, \qquad (17)$$

where s and r are the signals of the source and extraction result, respectively. This measurement shows the extracted signal compares with original signal.

Another measurement is used to measure the capability to reduce noisy signal is shown as

$$SNR_{noise} = 20\log\left(\frac{r-s}{\tilde{s}}\right) \quad \text{(dB)}, \qquad (18)$$

where s and r are the signals of the source and extraction result, respectively, and $\tilde{s}$ denotes the noisy signal.

The simulation results of sound source extraction are presented. First, sources in different frequencies are sim-
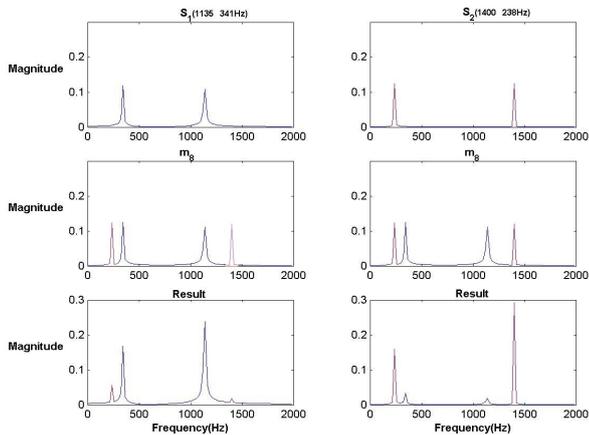
Figure 7: Simulation of two sound sources with two different frequency using filtered back-projection method with.

Figure 8: Simulation of two single frequency sources with signal fixed position and noise changed. The higher value means better performance.

Figure 9: Simulation of two dual frequencies sources with signal fixed position and noise changed. The higher value means better performance.

ulated. Second, experimation with changing the microphone array size and arrangement are shown. Finally, two pieces of real music are mixed and separated.

First, we verify the algorithm with two different frequencies in each source. The arrangement of components is the same as above. Let S_1 and S_2 are two dual frequency sources. The frequencies of S_1 are 1135 and 341 Hz. The frequencies of S_2 are 1400 and 238 Hz. In the Fig. 7 the spectrums of all the signal are shown. The top two figures are the spectrum of S_1 and S_2. The middle two figures are the signal received by microphone, m_8. The left side figure: S_1 is treated as the signal that we want then S_2 is noise. On the contrary, the right side figure S_2 is treated as signal. In the result, the magnitude of low frequency is smaller then high frequency. This is due to the filter we used is the high pass filter.

Next, we verify the efficiency and the sources position relation in two scenarios. Two point sources are placed in the conference room. One is treated as signal and the other is treated as interference. The signal source is placed at the fixed position and interference source is moved in the room area. When noise located at each point, SNR_{signal} and SNR_{noise} are calculated. The moving step of interference is 10 cm. In order to show the tendency of extraction efficiency, room size is changed to 6×6 meters and microphone array is still placed at the same place with 4 meters width. The single frequency and multi-frequency results are shown in Figs. 8 and 9.

The second scenario, we let the interference at a fixed position and move source in the room. Then also calculate the SNR with the same condition. Figs. 10 and 11 are the result of the single and multiple frequencies, respectively. From the four figures above, we could see the dark part means the simulation could extract signal we want clearly when signal source is in front of the noise source and near the microphone array. On the contrary, when signal is near the noise or far from to microphone array, the signal to noise ratio is less than 20 dB.

At final, we use two pieces of music in the simulation. The S_1 is female voice song and S_2 is the male voice song. The simulation system extract S_1 and S_2 from the mixed signal received by the 16-channel microphone array with the filtered back-projection method we proposed and delay-and-sum method. Choosing 3 periods of 0.05 second signal form two result music which is 30 seconds long to calculate the SNR. The comparion between filtered back-projection method and delay-and-sum method is shown in Table 1.
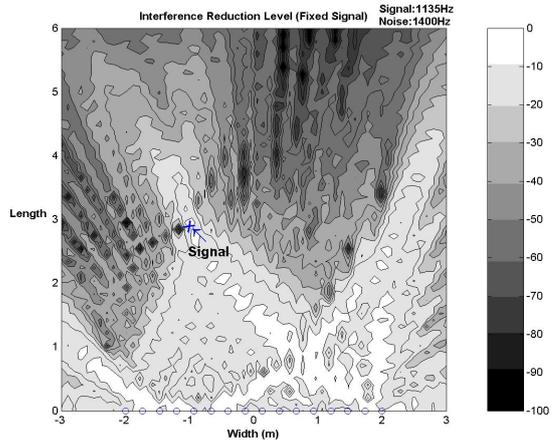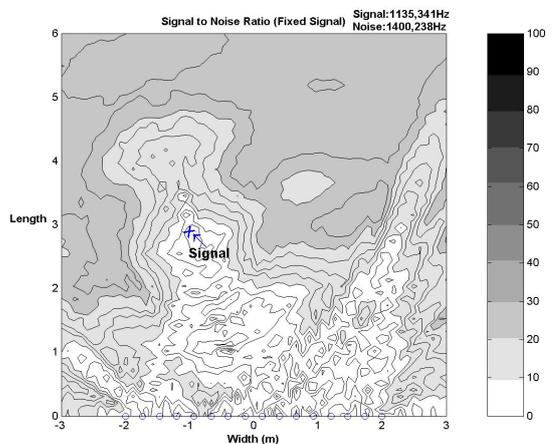
5. Conclusion

The simulation result shows the filtered back-projection method may reduce the noise and enhance the signal we wantt. As shown in figures, when the difference of distance from sources to microphone array in a proper range, the signal from the focus source could be extracted clearly. This system could be used in teleconference. Usually, the position of conferees is in front of the screen. According the Eq. 15, microphone array should be arranged around the half room. We consider the realization of the system would be too complex to setup hardware on multi-walls, so only arrange the microphone array under the projection screen. It would be a proper position to receive voice from speaker's height when sitting.
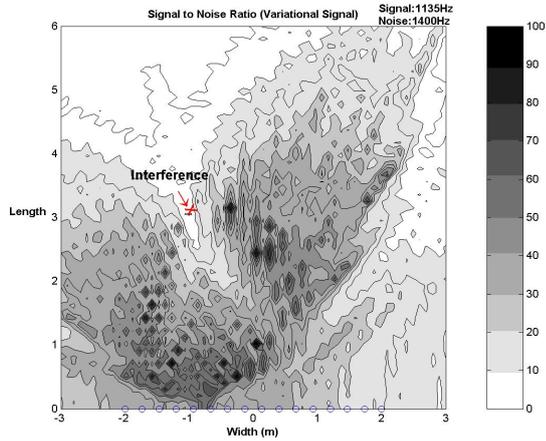
Figure 10: Simulation of two single frequency sources with noise fixed position and signal changed. The higher value means better performance.
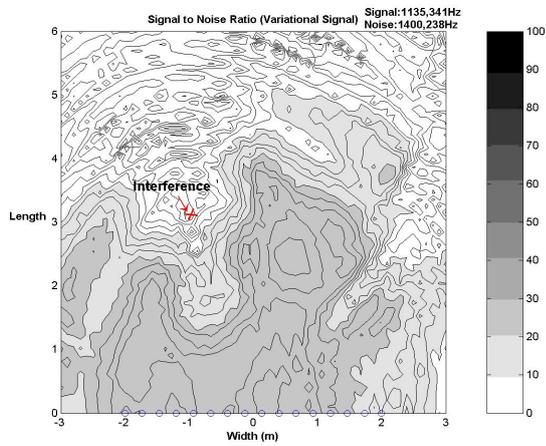
Figure 11: Simulation of two dual frequencies sources with noise fixed position and signal changed. The higher value means better performance.

Table 1: SNR_{signal} (dB) of extraction of music using two methods. The larger value means better performance.

Time	Back-projection		Delay-and-sum	
	S_1	S_2	S_1	S_2
t_1	8.34	21.17	10.15	13.26
t_2	18.78	29.06	11.26	13.74
t_3	13.96	9.37	26.19	16.19
t_4	37.18	5.47	10.56	2.13
t_5	10.49	17.48	10.23	12.68
Average	17.13		12.63	

References

[1] H. F. Silverman, W. R. Patterson, and J. L. Flanagan, "The huge microphone array," *IEEE Trans. on Concurrency*, vol. 6, no. 4, pp. 36–46, Oct.-Dec. 1998.

[2] Y. Tamai, S. Kagami, H. Mizoguchi, K. Sakaya, K. Nagashima, and T. Takano, "Circular microphone array for meeting system," vol. 2, pp. 1100–1105.

[3] D. Giuliani, M. Matassoni, and M. Omologo, "Hands free continuous speech recognition in noisy environment using a four microphone array," in *Proc. of IEEE Int'l. Conf. on Acoustics, Speech, and Signal Processing 1995 (ICASSP'95)*, vol. 1, Detroit, MI, May 1995, pp. 860–863.

[4] G. T. Herman, Ed., *Image Reconstruction from Projections: Implementation and Applications.* Berlin: Springer-Verlag, 1979.