Research Article

Development of the Audio Enhancement Method by Using the Reflected Signals in the Reverberant Environment

Shyang-Jye Chang and Hung-Wei Hsieh

*Department of Mechanical Engineering, National Yunlin University of Science and Technology, EM334B,
123 University Road, Section 3, Douliou, Yunlin 64002, Taiwan*

Correspondence should be addressed to Shyang-Jye Chang; changjye@yuntech.edu.tw

Received 29 September 2014; Accepted 26 December 2014

Academic Editor: Mo Li

Copyright © 2015 S.-J. Chang and H.-W. Hsieh. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper presents a new audio signal enhancement method based on reflection signal detection in the reverberant environment. This technique enhances the audio signal from the target sound source in time domain. The advantages of this technique are enhancing the target sound source by a simple algorithm and reducing the background noise effectively. The effects of distances between the speaker and microphone and the coefficient of correlation are discussed in this paper. The coefficient of correlation is rising from 0.17 to 0.63 as the distance between the speaker and microphone is varying from 30 to 90 cm. Also, some antinoise experiments are carried out in this study and the results show that the audio enhancement algorithm provides a good antinoise performance in different background noise levels.

1. Introduction

Speech is a very important communication tool for humanity and it is the most convenient way to deliver messages. Most communication equipment receives the speech signals by microphones and sends the signals to other equipment by some wireless technologies. When a speech signal is received by microphones, keeping the speech signal original is the basic requirement. The blur signals will make the system fail in some audio applications using the speech recognition technologies. Thus, the blur speech signals are not desired in any audio application. In the indoor environment, the speech signals will be reflected by the walls, tables, and ceilings and other objects cannot absorb voice completely. When a microphone is used to receive the speech signal in the indoor environment, the environmental noise and the reflected voice are all considered as noise. A popular method to reduce the noise is to set some sound-absorbing material in the indoor environment. For example, using sound-absorbing materials to build walls and ceilings is a common method to reduce the reflected voice. This method is widely used in the auditoriums and lecture theatres. The other way is reducing the noise

by means of signal processing technologies. The blur signals will be getting clear by the manipulation of noise-reduction algorithm in the signal processing system. This paper presents a new method to enhance the audio signal in the reverberant environment by using the reflected signal of the target sound source. Some antinoise experiments are also carried out in this study.

Blind source separation is a technology to separate a set of signals from a set of mixed signals. This technology can separate a set of signals without the information about the source signals or the mixing process. This problem is in general highly underdetermined. It is very useful in variety of conditions. Much of the literature in this field focuses on the separation of audio signals [1, 2]. The basic steps of blind sources separation are listed as follows: (i) transform mixed signals to the time-frequency domain, which is familiar with the name of spectrogram; (ii) apply instantaneous BSS algorithm for each frequency channel independently; (iii) determine the correspondence of separated components in each frequency based on temporal structure of signals; (iv) construct separated spectrogram of the source signals.

Beamforming, kind of spatial filter, is a signal processing technique used in sensor arrays for directional signal transmission or reception [3]. It can be used at both the transmitting and the receiving ends to achieve spatial selectivity. The beamforming technique can be used to receive the radio or sound waves. It has found numerous applications in radar, sonar, and acoustics [4, 5]. The beamforming technique can be used to extract sound sources in the indoor environment, such as multiple speakers in the cocktail party problem [6]. When using beamforming technique to receive speech signals in certain direction, the locations of the microphones should be known in advance.

The other popular method of reducing the effect of reverberation noise is spectral subtraction [7–9]. The popularity of this method is due to its relative simplicity and ease in implementation. In the spectral subtraction processing, an average signal spectrum and average noise spectrum are estimated in parts of the recording and subtracted from each other. Therefore, the average signal-to-noise ratio will be improved by this algorithm. It is assumed that the signal is distorted by a wide-band noise. The basic steps are listed as follows: (i) suppose the speech signal is corrupted by background noise; (ii) transform the sound signal into frequency domain from time domain; (iii) use the individual bands of complex sound signal to cut off the individual bands of noise signal; (iv) transform the sound signal into time domain from frequency domain and get the clear sound.

Among these algorithms, each still has some disadvantages. It is difficult to use blind source separation algorithm in the real-time system because of the large amount of data for calculation. The beamforming technique should use many microphones to build microphone array to raise the signal-to-noise ratio. It is also difficult to define the distribution of noises when using spectral subtraction algorithm. Therefore, it is urgently needed to develop a simple, low-cost audio enhancement algorithm for audio applications.

2. Materials and Methods

2.1. Concept. Sound, a kind of mechanical wave, travels through a medium. When the sound wave travels in the air, the reflection always occurs on the surfaces of objects. If one makes a voice in a mountain valley, a similar voice will be reflected from the cliff in a few seconds. If the distance of the mountain valley is more than approximately 17 meters, then the sound wave will be returned in more than 0.1 seconds. This time interval of 0.1 seconds is long enough for people to distinguish the original voice from the reflected one. The reflected voice is called “echo.” If the reflected voice returns in less than 0.1 seconds, the reflected voice is called “reverberation.” In the indoor environments, the combination of the sound wave and the reflected waves always causes reverberation effects. In this situation, two sound waves tend to combine as a new sound similar to the original one. Too much reverberation will affect the quality of audio signals received by the microphone. Moreover, the blur signals will affect the performance of audio applications, such as speech recognition and communication. Thus, it is

FIGURE 1: The scenario of receiving a target sound signal with a microphone.

FIGURE 2: The scenario of receiving a target sound signal and the reflected signals with a microphone.

very important to receive the clear audio signals by some technologies.

When a microphone is used to receive signal from the target sound source in the indoor environment, the noise signal will be received by the microphone as well. For traditional audio enhancement algorithm, each sound source in the indoor environment will be considered as a noise source except the target sound source. Thus, the reflected signal of the target source is also considered as a kind of noise. The scenario is shown as in Figure 1. Most audio enhancement technologies focus on the target source and cancel the other noise. The reflected signal from the target source will also be canceled since it is considered as the noise. However, the target signal and reflected signal from the target source are in fact from the same source. Rather than considering the reflected signal as the noise, it should be considered as a part of the target signal. In other words, if the reflected signal of the target source can be separated from the mixed signal, it can be used to enhance the target signal effectively by certain audio enhancement algorithms.

Basically, the reflected signal of the target source can be considered as the signal from the virtual target sources which is the same as the target source, as shown in Figure 2. In Figure 2, the two reflected signals of the target source are replaced by the signals from the two virtual target sources.

FIGURE 3: The flowchart of signal processing.

The locations of the virtual target sources are mirrored to the target source by the reflecting surfaces. Assuming that the sound signals are reflected perfectly on the reflecting surfaces, the signals from the virtual target sources will be quite similar to the signals from the target source. In the scenario of Figure 2, the microphone receives three target signals including one signal through the direct path and two signals through the reflecting paths. Therefore, the signals from the two virtual target sources can be used to enhance the target signal by certain audio enhancement algorithms.

2.2. Theory. When a sound wave travels in the air, the propagation path of the sound wave can be described by an impulse response function. In the reverberant environment, the propagation paths are different from the locations of the sound source and the microphone. In the case in which the location of the sound source, the location of the receiver, and the reflecting surface of the reverberation environment are maintained, then there exists only one set of functions of propagation paths between the sound source and the microphone. The set of functions of propagation paths describe the sound wave propagation path between the sound source and the microphone through the direct path and the reflected paths. By analyzing the original signal received by the microphone and the functions of the sound wave propagation paths, the signal through reflected path can be used to enhance the quality of the signal received by the microphone. In this paper, the technique of audio signal enhancement algorithm is based on time reversal method [10–12]. The purposes are to reduce background noise and to enhance the sound quality received by the microphone.

The time domain signal $m(t)$ and frequency domain signal $M(\omega)$ can be expressed as (1) and (2), respectively. Consider the following:

$$m(t) = p(t) \otimes s(t) \quad (1)$$

$$M(\omega) = P(\omega) \times S(\omega), \quad (2)$$

where $s(t)$ is the time domain signal of the sound source, $p(t)$ is the impulse response function of the reverberant environment, $m(t)$ is the time domain signal of the microphone, $\otimes$ is the convolution integration, and $\times$ is the multiplication.

The time reversal signal of microphone $r(t)$ can be expressed as

$$r(t) = m(-t). \quad (3)$$

According to Fourier transform, time reversal in time domain is equal to the phase conjugation in frequency domain, so that

FIGURE 4: The indoor experimental environment.

frequency domain of time reversal signal of microphone $R(\omega)$ can be expressed as

$$R(\omega) = M^*(\omega) = P^*(\omega) \times S^*(\omega), \quad (4)$$

where $*$ denotes the complex conjugate.

According to the signal processing of time reversal method, if the location of source is interchanged with sensor location and the source signal $s(t)$ is changed to time reversal signal of microphone $r(t)$, the frequency domain of time reversal signal of receiver $Z(\omega)$ can be expressed as the following equation:

$$\begin{aligned} Z(\omega) &= P(\omega) \times R(\omega) \\ &= P(\omega) \times P^*(\omega) \times S^*(\omega), \end{aligned} \quad (5)$$

where $P(\omega) \times P^*(\omega)$ is the function with the property of enhancement of signal. Figure 3 shows the flowchart of signal processing for time reversal method. According to the conclusion given above, the time reversal of original signal source can be obtained by time reversal method; that is, time reversal method can return the designate location of source signal and reduce the interference of noise signal.

2.3. Experiments. Figure 4 shows the indoor experimental environment with around eight meters in length, seven meters in width, and three meters in height. There are one conference table, some chairs, and cabinets in the room. The speaker, as the sound source, is set on the conference table. The microphone, as the receiver, is also set on the

FIGURE 5: The detailed arrangement of the experimental setting.

conference table and makes a distance from the speaker. The speaker and microphone are set at the same level of height and make a distance of 15 cm from the desktop of the conference table. Figure 5 shows the detailed arrangement of the experimental settings. In the figure, d is the distance between the speaker and microphone. In the experiment, the sampling frequency of the microphone is set as 44.1 kHz. Three sets of experiments with different sound source are carried out. The different distance between the speaker and microphone is considered as a parameter in the experiments, and it is set from 30 cm to 90 cm with the increment of 10 cm.

3. Results and Discussion

3.1. Performance Index. In this paper, the distance between the speaker and microphone is discussed on the influence of the performance of enhancement algorithm in the reverberant environment. The coefficient of correlation is used to quantify the performance of the audio enhancement algorithm. The definition of coefficient of correlation can be expressed as

$$\rho_{xy} = \frac{E[x(t) \times y(t)]}{(E[x(t)^2] \times E[y(t)^2])^{1/2}}, \quad (6)$$

where ρ_{xy} is coefficient of correlation, $E[*]$ is the statistical average, and $x(t)$ and $y(t)$ are the signals to be evaluated.

In this paper, $x(t)$ and $y(t)$ are the signal of sound source and the measured signal through the audio enhancement algorithm. The value of coefficient of correlation is between -1 and 1 . The value of coefficient of correlation ρ_{xy} is equal to 1 as the two signals are equal. It means that the two signals are completely positive correlative. The value of coefficient of correlation ρ_{xy} is equal to -1 as two signals are two opposite waveforms. It means that these two signals are completely negative correlative.

3.2. Experimental Results. Three different sound sources are used in the experiments. The first and the second sources are the speech signals of a woman and a man, respectively. The third source is a short music. A loud speaker is used to play the sound of the source, and the microphone is used to receive the sound signals in the experimental environment. The received signal is enhanced by the audio enhancement algorithm, and the coefficient of correlation between signal of sound source and the received signal through the audio enhancement algorithm can be obtained. Each sound source

TABLE 1: The experimental results table.

d (cm)	1st source	2nd source	3rd source	Average
30	0.22	0.15	0.13	0.17
40	0.39	0.29	0.33	0.34
50	0.50	0.40	0.45	0.45
60	0.59	0.48	0.52	0.53
70	0.63	0.52	0.55	0.57
80	0.66	0.58	0.60	0.61
90	0.67	0.60	0.61	0.63

FIGURE 6: The experimental results with different distances between the speaker and microphone.

is tested by seven different d , which are from 30 cm to 90 cm with the increment of 10 cm. The experimental results are listed in Table 1. Figure 6 shows the experimental results with different distances between the speaker and microphone. From the experimental results, the value of coefficient of correlation is increasing as the distance between the speaker and microphone increases. This result shows that the sound signal can be enhanced by increasing the distance between the speaker and microphone. Figure 6 also indicates that the coefficient of correlation of the first source (woman) is higher than those of the second and the third sources.

Some antinoise experiments are also carried out in this study. The experimental settings are all the same as the settings of pervious experiments except the air conditioner is turned on as the background noise. The distance between the speaker and microphone is set as 70 cm in the experiments. The noise level is set by adjusting the fan speed of the air conditioner. Figures 7, 8, and 9 show the results of performance of the antinoise test with the three different noise levels. The sound source signal corrupted by noise of low level and clear signal through the audio enhancement algorithm are shown in Figure 7. The sound source signal

FIGURE 7: The speech signal corrupted by low level noise and clear signal through the audio enhancement algorithm.

FIGURE 8: The speech signal corrupted by medium level noise and clear signal through the audio enhancement algorithm.

FIGURE 9: The speech signal corrupted by high level noise and clear signal through the audio enhancement algorithm.

corrupted by noise of medium level and clear signal through the audio enhancement algorithm are shown in Figure 8. The sound source signal corrupted by noise of high level and clear signal through the audio enhancement algorithm are shown in Figure 9. From the experimental results, it is found that the noise level in the signal is getting lower through the audio enhancement algorithm. The results draw a conclusion that the method presented in this paper provides a good performance in the reverberant environment.

4. Conclusions

This paper presents the audio enhancement algorithm which is based on reflection signal detection in the reverberant environment. Some experiments with different distance between the speaker and microphone are carried out and the results are discussed on the influence of the performance of enhancement algorithm in the reverberant environment. The coefficient of correlation is used as the performance index to evaluate the performance of the audio enhancement algorithm presented in this paper. The coefficient of correlation is rising from 0.17 to 0.63 as the distance between the speaker and microphone is varying from 30 to 90 cm. The experimental results show that the audio signal can be enhanced by increasing the distance between the speaker and microphone. The results of antinoise experiments also indicate that the audio enhancement algorithm provides a good antinoise performance as the audio signal is corrupted by background noise. Comparing with other audio enhancement algorithms, the most important advantages of this technique are small data calculating and not needing to use the microphone array. In the future, more parameters of the audio enhancement algorithm will be investigated.

Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

Acknowledgment

The authors would like to thank Ministry of Science and Technology (MOST) for their financial supports to the project (Grant no. NSC 102-2221-E-224-024).

References

- [1] A. J. Bell and T. J. Sejnowski, "An information-maximization approach to blind separation and blind deconvolution," *Neural Computation*, vol. 7, no. 6, pp. 1129–1159, 1995.
- [2] M. Zibulevsky and B. A. Pearlmutter, "Blind source separation by sparse decomposition in a signal dictionary," *Neural Computation*, vol. 13, no. 4, pp. 863–882, 2001.
- [3] B. D. van Veen and K. M. Buckley, "Beamforming: a versatile approach to spatial filtering," *IEEE ASSP Magazine*, vol. 5, no. 2, pp. 4–24, 1988.
- [4] S.-W. Gao and J. W. R. Griffiths, "Experimental performance of high-resolution array processing algorithms in a towed sonar

- array environment," *The Journal of the Acoustical Society of America*, vol. 95, no. 4, pp. 2068–2080, 1994.
- [5] P. S. Naidu and P. G. K. Mohan, "Signal subspace approach in localization of sound source in shallow water," *Signal Processing*, vol. 24, no. 1, pp. 31–42, 1991.
 - [6] O. Hoshuyama, A. Sugiyama, and A. Hirano, "A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters," *IEEE Transactions on Signal Processing*, vol. 47, no. 10, pp. 2677–2684, 1999.
 - [7] L.-P. Yang and Q.-J. Fu, "Spectral subtraction-based speech enhancement for cochlear implant patients in background noise," *Journal of the Acoustical Society of America*, vol. 117, no. 3, pp. 1001–1004, 2005.
 - [8] Z. Lin, R. A. Goubran, and R. M. Dansereau, "Noise estimation using speech/non-speech frame decision and subband spectral tracking," *Speech Communication*, vol. 49, no. 7-8, pp. 542–557, 2007.
 - [9] D. S. Brungart, P. S. Chang, B. D. Simpson, and D. Wang, "Isolating the energetic component of speech-on-speech masking with ideal time-frequency segregation," *Journal of the Acoustical Society of America*, vol. 120, no. 6, pp. 4007–4018, 2006.
 - [10] B. X. Zhang, C. H. Wang, and M. H. Lu, "Study of self-focusing in underwater by time reversal method," *Chinese Journal of Acoustics*, vol. 22, no. 1, pp. 22–32, 2003.
 - [11] G. Montaldo, M. Tanter, and M. Fink, "Real time inverse filter focusing through iterative time reversal," *The Journal of the Acoustical Society of America*, vol. 115, no. 2, pp. 768–775, 2004.
 - [12] K. G. Sabra, P. Roux, H.-C. Song et al., "Experimental demonstration of iterative time-reversed reverberation focusing in a rough waveguide. Application to target detection," *The Journal of the Acoustical Society of America*, vol. 120, no. 3, pp. 1305–1314, 2006.