Master of Science

Improvement of Speech Clarity in Noisy and Reverberant Conditions Based on Active Noise Control and Inverse Filter Algorithm

The Graduate School

of the University of Ulsan

Department of Mechanical Engineering

HANG SU

Improvement of Speech Clarity in Noisy and Reverberant Conditions Based on Active Noise Control and Inverse Filter Algorithm

Supervisor: Prof. Chang-Myung Lee

A Master's Thesis

Submitted to the Office of Graduate School of

University of Ulsan

in partial Fulfillment of the Requirements for the Degree of

Master of Science

by

Hang Su

School of Mechanical Engineering

Ulsan, Republic of Korea

June 2018

HANG SU 의 공학석사학위 논문을 인준함

심사위원장 이병룡 (인)

심사위원 김도중 (인)

심사위원 이장명 (인)

울 산 대 학 교 대 학 원

2018 년 6 월

ABSTRACT

In an indoor electronic meeting system, active noise control (ANC) system is generally used to reduce the ambient noise and improve the clarity of the received speech signal. Typical ANC methods like the filtered-x least mean square (FxLMS) algorithm attempt to control the background noise, which is a kind of additive noise. However, when the speaker of ANC system is used as the speech communication speaker, the nonadditive noise like reverberation, dispersion and attenuation caused by the propagation path is also necessary to be considered. As a common phenomenon that interrupts the communication, the speech reverberation is focused on in this thesis. To obtain the better approximation of the speech reverberation in the given room, the reverberation is simulated by image source model method (ISM) and added in the propagation model. Then, the ANC-IF (active noise control-inverse filter) method that combined the FxLMS algorithm and the FIF (fast inverse filter) algorithm is proposed to control the noise included the background noise and the speech reverberation in the electronic meeting system.

The performance of the FxLMS algorithm in the reverberant condition was first investigated. The noise source produced the factory noise as the background noise, and the ANC speaker generated the opposite phase noise sound wave at the same time. Then the algorithm is modified to add the speech signal to the system. Though the performance of the algorithm to reduce the noise in the reverberant condition is worse than in the non-reverberant condition, it is also contributive to improve the clarity of the received speech signal with the ANC system by reducing the background noise in the environment. The evaluation index of the simulation results demonstrates this

conclusion.

Secondly, we study on the inverse filter to improve the speech signal clarity by eliminating the speech reverberation caused by the propagation path. The inverse filter of the propagation path was calculated based on the fast inverse filter (FIF) algorithm. The original speech signal was preprocessed by the inverse filter to make the received speech signal with less reverberation through the propagation path. The propagation model was run to verify the correctness of the inverse filter algorithm. The simulation results indicate that the speech signal preprocessed by the inverse filter can be improved the speech signal clarity significantly. The same conclusion can be obtained by human subjective feeling.

At last, the ANC-IF method that combined the ANC system and the inverse filter was proposed and verified to reduce two types of noise simultaneously in the simulation model. The ANC speaker generated the signal that included both the opposite phase noise signal and the speech signal preprocessed by the inverse filter. Consequently, the speech signal would be received with less background noise and less speech reverberation. The simulation results indicate that the ANC-IF method can control the background noise effectively and reduce the speech reverberation significantly. Evaluation indexes also show the effectiveness of the proposed method.

ACKNOWLEDGMENTS

The deepest appreciation is presented for my advisor, Prof. Chang-Myung Lee. For his valuable academic guidance and generous financial support during my master's career.

The special thanks are offered to my thesis committee members, Prof. Byung-Ryoung Lee and Prof. Do-Joong Kim. Their valuable comments make this thesis more rational and precise.

The special thanks are also given to my cousin, Dr. Zhenhua Xu. His suggestions and encourages accompany on my entire master's career.

Thanks to my seniors, Huanyu Dong, Guangquan Hou, Qi Wu, Peng Wang and his girlfriend, Jue Wang. For their academic suggestions and care in life.

Thanks to my young school sister, Min Chen. I'm absolutely delighted that we can study and grow up together.

Heartful thanks to my parents. I can't accomplish the goal without your selfless support. You are my power source and inspire me to go further.

Sincere thanks to everyone I ever met, no matter encouragement or criticism you gave, both make me stronger.

CONTENTS

LIST OF FIGURES AND TABLES

ABBREVIATIONS

ANC	Active noise control
LMS	Least mean square
NLMS	Normalized least mean square
FxLMS	Filtered-x least mean square
ILC	Iterative learning control
RC	Repetitive control
PP	Primary path
SP	Secondary path
ISM	Image source method
RT	Reverberation time
RIR	Room impulse response
IF	Inverse filter
FIF	Fast inverse filter
ANC-IF	Active noise control-inverse filter
SPL	Sound pressure level
SNR	Signal-to-noise ratio
SegSNR	Segmental signal-to-noise ratio

NOMENCLATURES

u	Acoustic particle velocity
p	Acoustic pressure
ρ	Acoustic density
ρ_0	Ambient density
c_0	Sound speed
Nx	Number of simulation grid points in the x (column) direction
Ny	Number of simulation grid points in the y (row) direction
dx	Simulation grid point spacing in the x direction
dy	Simulation grid point spacing in the y direction
$d(n)$	Output signal of the primary path
$y'(n)$	Output signal of the secondary path
$s(n)$	Transfer function of the secondary path in time domain
$S(z)$	Transfer function of the secondary path in frequency domain
$y(n)$	Anti-phase noise signal
$W(z)$	ANC adaptive filter
$w(n)$	Coefficient vector of ANC adaptive filter
L	Order of the adaptive filter
$x(n)$	Reference signal to ANC
$x'(n)$	Modified reference signal
$\hat{s}(n)$	Estimated impulse response of the secondary path
μ	Step size
$s_p(n)$	Clean speech signal
$\hat{s}_p(n)$	Noisy speech signal
M	Number of frames
N	Length of a frame
$h(t)$	Room impulse response in time domain
$y(t)$	Original speech signal in time domain
$y'(t)$	Filtered speech signal in time domain
$H(\omega)$	Room impulse response in frequency domain

$Y(\omega)$	Original speech signal in frequency domain
$Y'(\omega)$	Filtered speech signal in frequency domain
$ih(t)$	Inverse filter response in time domain
$iH(\omega)$	Inverse filter response in frequency domain
$\mathrm{v}(n)$	Clean speech signal
$\hat{S}^{-1}(z)$	Inverse filter response
$e(n)$	Residual background noise

Chapter 1 Introduction

1.1 Background

Noise is the sound that undesired and made human uncomfortable and degraded the speech clarity [1]. Thus, many methods are proposed to reduce the noise. The methods can be broadly classified into two categories, the first is passive noise control, which is the traditional method such as enclosures, barriers, and silencers, the other is active noise control (ANC). The ANC method is suited for controlling noise in the low-frequency band that is just the frequency range where the most energy of speech is concentrated.

Lueg applied for the first patent for an active noise control (ANC) system in 1934, and since then, many ANC algorithms have been investigated. According to the work mode, ANC algorithms can be divided into narrowband feedback ANC [2][3], broadband feedback ANC [4][5][6], lattice ANC [7][8], frequency-domain ANC [9][10] and so on. Based on the calculation mode, ANC algorithms can be divided into least mean square (LMS) algorithm [11], filtered-x least mean square (FxLMS) algorithm, iterative learning control (ILC) algorithm [12], repetitive control (RC) [13] algorithm, etc.

ANC systems can reduce noise and decrease its influence on speech. Most research focused on the background noise [14], which is a kind of additive noise. However, the speech reverberation induced by the propagation path is a kind of the nonadditive noise and is also a common phenomenon that interrupts speech communication. The speech reverberation is focused on as an essential factor to be considered in this thesis.

Reverberation is the persistence of a sound after it is produced. The sound waves that propagate in a room are absorbed and reflected by walls and furniture. The sound persistence time is called the reverberation time which is one of the most critical factors in architectural acoustics. The reverberation time can be estimated using an empirical equation proposed by Sabine in the 1980s.

A speaker works as one part of an ANC system to generate a sound wave with the opposite phase of the noise to reduce the noise in the room. The speech signal can also be broadcast by this speaker, although the signal would be corrupted by the room impulse response (RIR). When the speaker plays the speech signal in a communication system, it is necessary to consider the speech reverberation caused by the propagation path. A speech signal corrupted by RIR may induce reverberation and stop people from effectively distinguishing words.

Neely established a stable and causal inverse filter that can be used to remove the influence of the RIR on a speech signal [15]. Based on Neely’s work, Kirkeby and Nelson proposed a fast inverse filter (FIF) algorithm [16][17][18]. The FIF algorithm is optimized with a faster calculation speed and better accuracy.

Researchers rarely study the control of the background noise and the speech reverberation simultaneously [19][20][21]. Richard improved the speech intelligibility in noisy and reverberant conditions by redistributing the energy of the speech [22]. Dong used the speech enhancement and an inverse filter to accomplish the same purpose [23]. However, in the case of an indoor electronic meeting system, as shown in Fig. 1-1, it is meaningful to improve the clarity of the received speech signal by controlling the two categories of noise simultaneously.

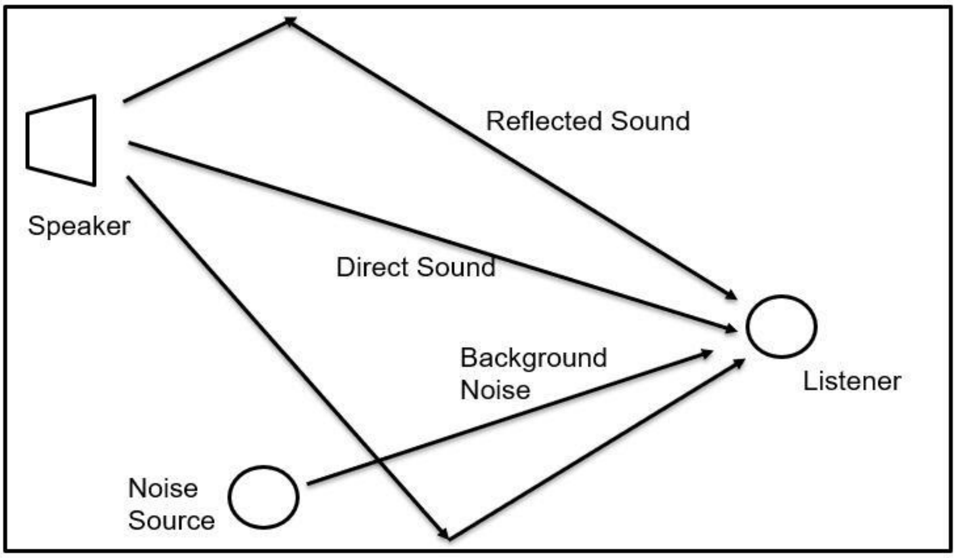


Fig. 1-1 Visualization of application scenario

1.2 ANC-IF method

In this thesis, a near-end denoise system with an ANC-IF method is established by combining the ANC method and the IF method. A classical ANC algorithm, the FxLMS algorithm, is used to generate the anti-phase sound wave to reduce the background noise in real time. The speech signal is preprocessed by an inverse filter based on the FIF algorithm to eliminate the speech reverberation. The RIR that is used in the FIF algorithm is acquired from the FxLMS algorithm. This means that the transfer function of the sound propagation path (RIR) is measured once and used in the two algorithms. A simulation model in MATLAB was established to validate the proposed method. The evaluation index of the simulation results indicates that the ANC-IF method can reduce the two types of noise considerably and improve the speech clarity significantly.

1.3 Outline

This thesis is organized as follows. Chapter 1 introduces the research background and reviews the relative academic achievements. Chapter 2 shows the simulation model method and partial parameters. Chapter 3 describes the ANC algorithm and emphasizes the performance of the FxLMS algorithm in the reverberant condition. Chapter 4 presents the inverse filter method in detail. The ANC-IF method, which combines the FxLMS algorithm and the FIF algorithm, is proposed in Chapter 5. Chapter 6 provides the conclusions of this thesis.

Chapter 2 Simulation Model

2.1 k-Wave toolbox

2.1.1 Introduction

The simulation model is established by the k-Wave toolbox of MATLAB. The k-Wave is an open source, third party, MATLAB toolbox designed for the time-domain simulation of propagating acoustic waves in 1D, 2D, or 3D. The toolbox has a wide range of functionality, but at its heart is an advanced numerical model that can account for both linear and nonlinear wave propagation and power law acoustic absorption [24][25]. The interface to the simulation functions has been designed to be both flexible and user-friendly, while the computational engine has been optimized for speed and accuracy.

2.1.2 Governing equations

When an acoustic wave passes through a compressible medium, there are dynamic fluctuations in the pressure, density, temperature, particle velocity, etc. These changes can be described by a series of coupled first-order partial differential equations based on the conservation of mass, momentum, and energy within the medium. Often in acoustics, these equations are combined into a single “wave equation" which is a second-order partial differential equation in a single acoustic variable (most often the acoustic pressure). For example, in the classical case of a small amplitude acoustic wave propagating through a homogeneous and lossless fluid medium, the first-order equations are given by [26].

$$\frac{\partial \mathrm{u}}{\partial t} = -\frac{1}{\rho_0}\nabla p, \qquad \text{(momentum conservation)} \tag{2-1}$$

$$\frac{\partial \rho}{\partial t} = -\rho_0 \nabla u, \qquad \text{(mass conservation)} \tag{2-2}$$

$$p = c_0^2 \rho, \qquad \text{(pressure-density relation)} \tag{2-3}$$

Here u is the acoustic particle velocity, p is the acoustic pressure, ρ is the acoustic density, ρ_0 is ambient (or equilibrium) density, and c_0 is the isentropic sound speed. These equations assume the background medium is quiescent (meaning there is no net flow and the other ambient parameters don't change with time) and isotropic (meaning the material parameters do not depend on the direction the wave is traveling). When they are combined, they give the familiar second-order wave equation.

$$\nabla^2 p - \frac{1}{c_0^2}\frac{\partial^2 p}{\partial t^2} = 0, \tag{2-4}$$

2.1.3 First-order simulation function

There are three simulation functions in the k-Wave Toolbox that implement the first-order k-space model for fluid media. These are named **kspaceFirstOrder1D, kspaceFirstOrder2D, and kspaceFirstOrder3D** and correspond to simulating wave propagation in one, two, and three dimensions as their names imply. In this case, "first-order" refers to the fact we are solving a system of coupled first-order partial differential equations. It is not related to the order of numerical accuracy of the solution, or to the order of the acoustic variables retained in the governing equations.

The simulation functions are called with four input structures; **kgrid, medium, source,** and **sensor**. The properties of the simulation are then set as fields for these structures in the form **structure.field**. The four structures respectively define the properties of the computational grid, the material properties of the medium, the properties and locations of any acoustic sources, and the properties and locations of the sensor points used to record the evolution of the pressure and particle velocity fields over time. When the simulation functions are called, the propagation of the wave-field in the medium is then computed step by step, with the acoustic field at the sensor elements stored after each iteration. These values are returned when the time loop has completed.

2.1.4 MATLAB parameters setting

In this simulation model, the domain size is 9m by 12m and divided into 720 by 1080 grid points. The sound speed is set to be isotropic. The noise source, the ANC speaker and the sensor are set as shown in Fig. 2-1. The four input structures are passed to **kspaceFirstOrder2D** which then calculates and returns the acoustic pressure recorded at each sensor point for each time step.

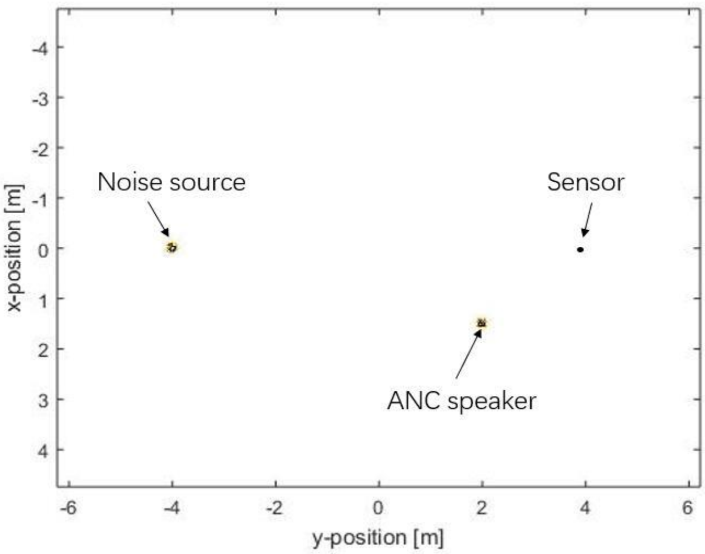


Fig. 2-1 Layout of the ANC system

Simulation partial parameters:

```
Nx = 720;                         % number of grid points in the x (column) direction
Ny = 1080;                        % number of grid points in the y (row) direction
dx = 9.0/Nx;                      % grid point spacing in the x direction [m]
dy = 12.0/Ny;                     % grid point spacing in the y direction [m]
source.p_mask(Nx/2,Ny/6) = 1;             % Noise Source position
source.p_mask(Nx*2/3,Ny*2/3) = 1;         % ANC Speaker position
sensor.mask(Nx/2, Ny*5/6) = 1;            % Sensor position
```

2.2 Image source model method

2.2.1 Introduction

To get the better approximation of the speech reverberation in the given room, the image source model method is used to model the reverberation. The image-source model (ISM) is a well-known technique that can be used in order to generate a synthetic room impulse response (RIR), i.e., a transfer function between a sound source and an acoustic sensor, in a given environment. Once such an RIR is available, a sample of audio data can be obtained by convolving the RIR with a given source signal. This provides a realistic sample of the sound signal that would be recorded effectively at the sensor in the considered environment.

Allen & Berkley were the first to propose a full implementation of the ISM technique in a landmark paper in 1979 [27]. Since then, this specific implementation has been used by many researchers.

However, Allen & Berkley's method suffers from several drawbacks, which are related to the way the image-source model is implemented. An improvement to Allen & Berkley's original ISM implementation was proposed by Peterson in [28], where each image-source impulse is implemented directly as a (truncated) fractional-delay filter (i.e., a sinc function). This effectively allows the exact representation of non-integer time delays for each image source. Note that the same result is also obtained by computing the transfer function in the frequency domain and then taking the inverse Fourier transform back into the time domain. This approach is of considerable importance as it provides a quick and easy way to generate RIRs with varying environmental characteristics, such as different reverberation times. Consequently, the ISM technique has been used intensively in many application domains of room acoustics and signal processing.

2.2.2 MATLAB code implementation

The MATLAB code on this section provides an implementation of the image-source method (ISM) described [29] for simulating reverberant audio data in small-room acoustics. The MATLAB parameters are set as similar as in the propagation model made by the k-Wave toolbox. The room size and the layout of the instruments are shown in Fig. 2-2.

The command ISM_RIR_bank(...) will create and save a bank of RIRs into the file ISM_RIRs.mat. This process can take quite a while depending on the number of RIRs to compute, size of the room, reverberation time, etc. The other command ISM_AudioData(...) takes the .mat file and the source signal vector as inputs and computes the sensor data AuData by convolution. This function also offers the possibility to write the resulting audio into a .wav or .mat file. The resulting audio data is a matrix where each column contains the signal generated for the corresponding microphone. Of course, the same process can also be used if the source remains stationary or in environments containing only one sensor.

MATLAB partial commands:

```
>> ISM_RIR_bank(my_ISM_setup, 'ISM_RIRs.mat');
>> AuData = ISM_AudioData('ISM_RIRs.mat', SrcSignalVec);
```

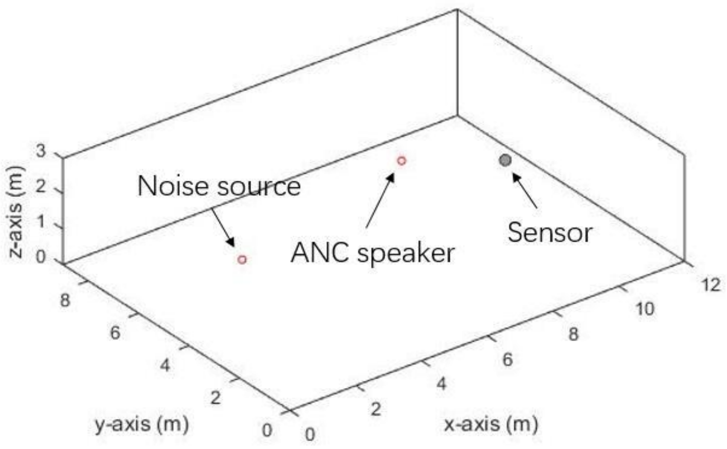


Fig. 2-2 Layout of the instruments in ISM

2.3 Conclusion

In this chapter, the numeric simulation model method is introduced to establish the simulation model for verifying the proposed method simply and conveniently. The simulation toolbox of MATLAB, k-Wave, is used to model the propagation path. The interface of the toolbox has been designed to be both flexible and user-friendly, while the computational engine has been optimized for speed and accuracy. Furthermore, in order to get the more realistic simulation result, the image source model method is used to simulate the reverberation in the given room. The ISM method has been used by many researchers and a lot of academic achievements are published. The background of above two methods and partial parameters in MATLAB are also introduced to make the reader understand the simulation model easily.

Chapter 3 Active Noise Control

Active noise control (ANC), is a method for reducing unwanted sound by generating the anti-phase sound wave of the noise and making the waves interfere at the control point.

Since the phase of the high-frequency wave changes rapidly, ANC technology is more suitable for controlling the noise in the low-frequency band. The energy of speech is also mostly concentrated in low-frequency range (about 30~3000Hz). Thus, ANC technology can be used in electronic meeting system to reduce the variant background noise in real time. In this thesis, an ANC system is used to improve the clarity of speech by controlling the noise in the frequency band of 30 to 3000 Hz. The sample rate is set as 8 kHz.

3.1 Filtered-x least mean square algorithm

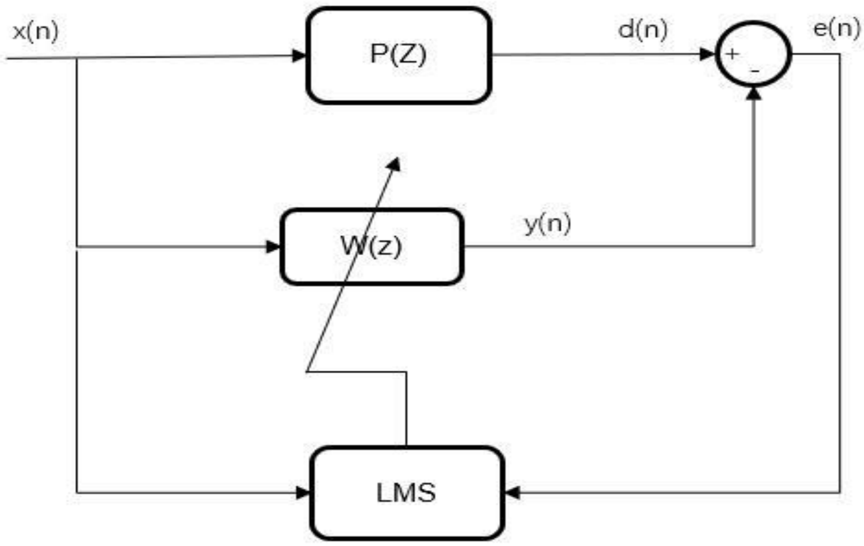


Fig. 3-1 ANC system block diagram based on LMS algorithm

The FxLMS algorithm is a widely used conventional algorithm based on the least mean square (LMS) algorithm [30]. The LMS algorithm is an adaptive filter algorithm that is used to mimic a desired filter by finding the filter coefficients that relate to producing the least mean square of the error signal (the difference between the desired signal and the actual signal) as shown in Fig. 3-1 [31][32]. It is a stochastic gradient descent method that the filter can be adapted based on the error at the current time. It was invented in 1960 by Stanford University professor Bernard Widrow and his Ph.D. student, Ted Hoff. Morgan proposed the FxLMS algorithm, where the LMS algorithm

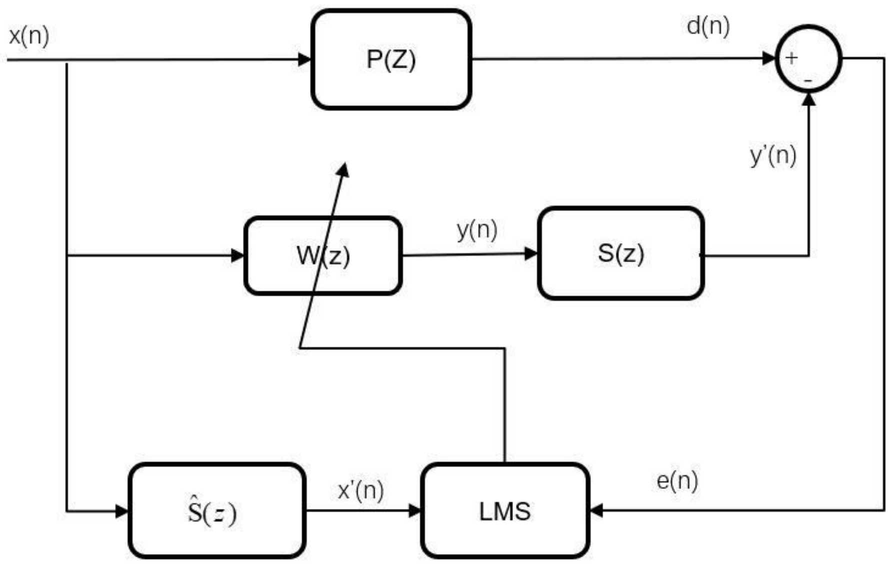


Fig. 3-2 FxLMS algorithm block diagram

is modified to consider the effect of the secondary path (SP) [33]. A block diagram of the FxLMS algorithm is shown in Fig. 3-2.

To consider the influence of the secondary path (between the speaker and the control point), the estimated transfer function of the secondary path is added in the reference signal path. The residual error signal is expressed as,

$$e(\mathrm{n}) = d(n) - y'(n) = d(n) - s(n) * y(n) , \quad (3\text{-}1)$$

where *d(n)* is the output signal of the primary path and *y'(n)* is the output signal of the secondary path at the error microphone location. *s(n)* is the transfer function of the secondary path *S(z)* (RIR in the inverse filter algorithm), and * denotes linear convolution. *y(n)* is the anti-phase signal calculated by the algorithm.

$$y(n) = w(n) * x(n)\,, \tag{3-2}$$

where *w(n)* is the coefficient vector of the ANC adaptive filter *W(z)*, and $w(n)=[w_0(n), w_1(n), w_2(n), \ldots, w_n(n)]$. *x(n)* is the reference input signal at time n, $x(n)=[x_0(n), x_0(n-1), x_0(n-2), \ldots, x_0(n-l+1)]$, and *l* denotes the order of the adaptive filter.

x'(n) is the modified reference signal which is formed by filtering the original reference signal with the estimated transfer function of the secondary path:

$$x'(n) = \hat{s}(n) * x(n)\,, \tag{3-3}$$

where $\hat{s}(n)$ is the estimated impulse response of the secondary path.

For step size μ, the coefficient vector of the ANC adaptive filter is updated as:

$$w(n+1) = w(n) + \mu x'(n) e(n)\,, \tag{3-4}$$

3.2 Simulation model

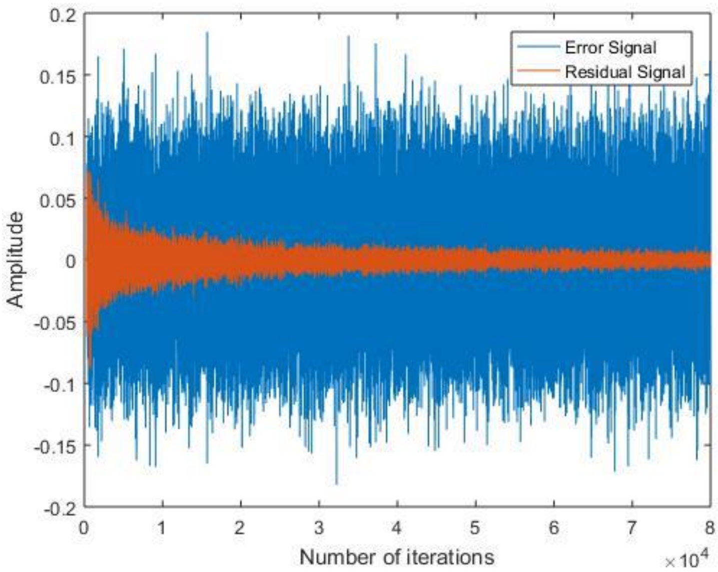


Fig. 3-3 Secondary path identification

A simulation model is established with the k-Wave toolbox in MATLAB. The density of grid points ensures the maximum frequency that the software can simulate. Gaussian white noise is used to estimate the transfer function of the secondary path. Factory noise acquired from the NoiseX-92 database [34] is used as the environmental noise. Normalized LMS adaptive filter method is used to identify the secondary path transfer function. The residual signal of the identification process is shown in Fig. 3-3.

3.3 Simulation result

A speaker was used to play white noise to measure the transfer function between the speaker and a microphone (the secondary path-*S(z)* in the FxLMS algorithm block diagram).

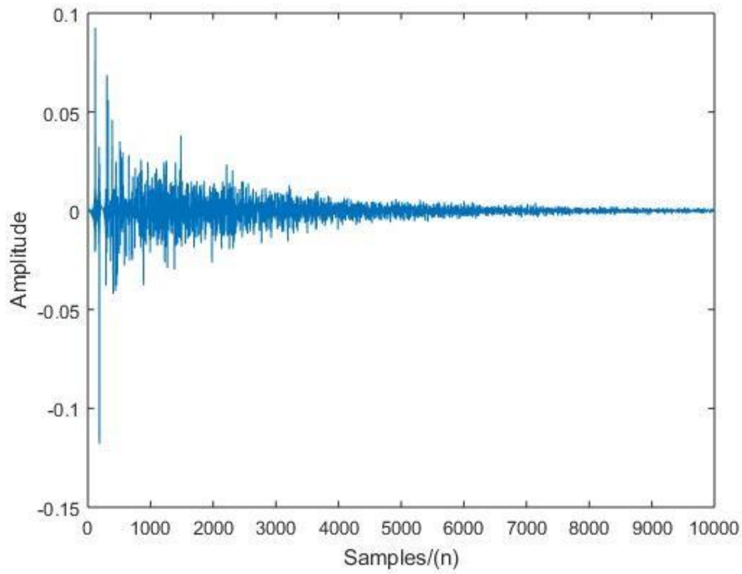


Fig. 3-4 Estimated transfer function of the secondary path

As shown in Fig. 3-4, the transfer function of the secondary path is identified between the speaker and the point to be controlled (the error microphone location). The secondary path transfer function is also the RIR when the propagation path is put in a room. To obtain a better approximation of the speech reverberation in the given room,

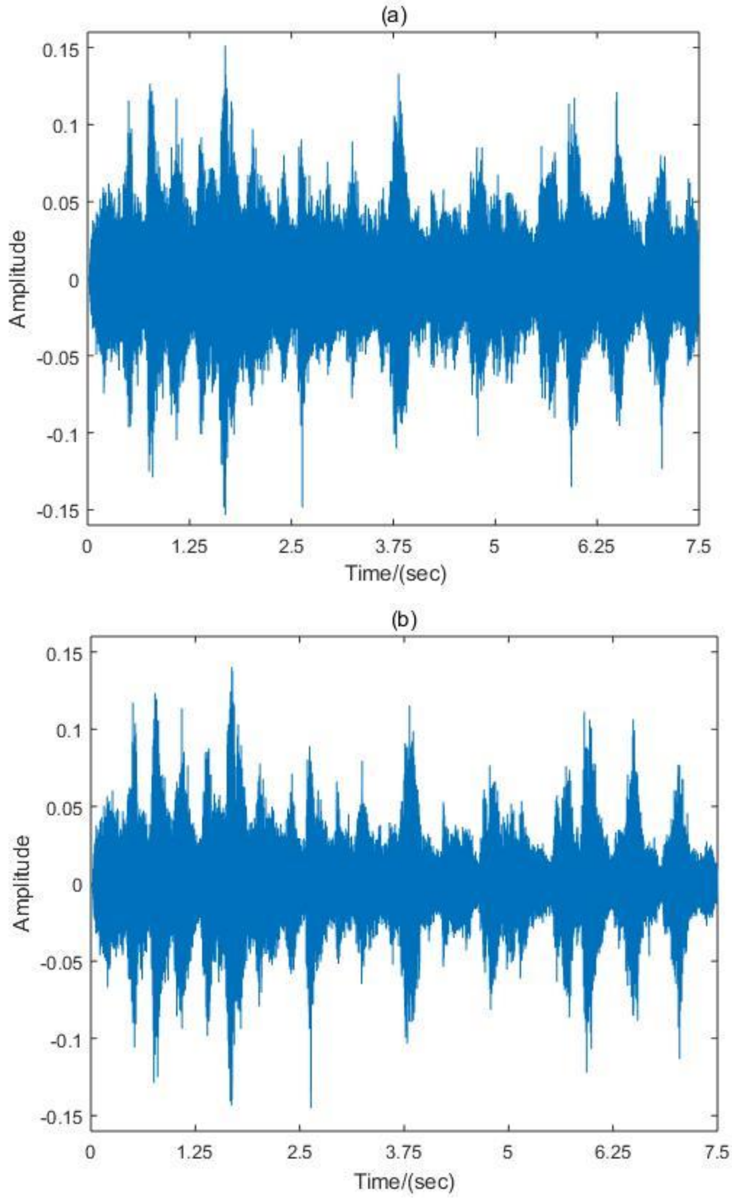


Fig. 3-5 Received speech signal with background noise.
(a) No ANC system, (b) ANC system.

the reverberation is simulated by the image source model method [35] and added to the propagation model. Tailing points are formed by the reflection of boundaries.

Signals obtained with and without the ANC system are shown in Fig. 3-5. When comparing the figures, the details of the speech signal are more obvious and recognizable when the ANC system is used. The same conclusion can be drawn subjectively when listening to the results.

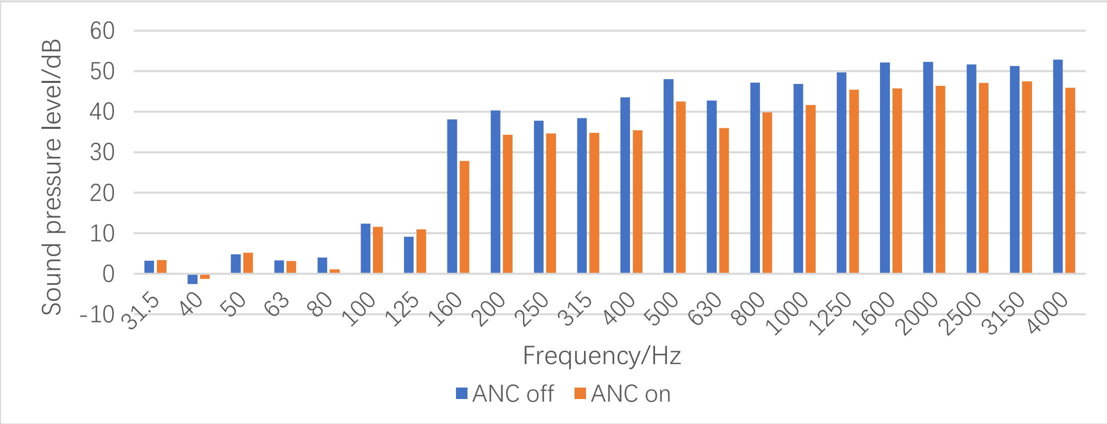


Fig. 3-6 SPL in the octave bands of the background noise

The sound pressure level (SPL) of the received background noise in the octaves and the decreased SPL when ANC worked are showed in Fig. 3-6 and Fig. 3-7 individually. The figures illustrate that the frequency range of 160-4000 Hz is where the ANC system works, and the SPL is decreased about 6 dB on average. This frequency range is also where the most energy of the speech signal is concentrated. Although the SPL does not decrease much, the improvement of the speech signal clarity is significant.

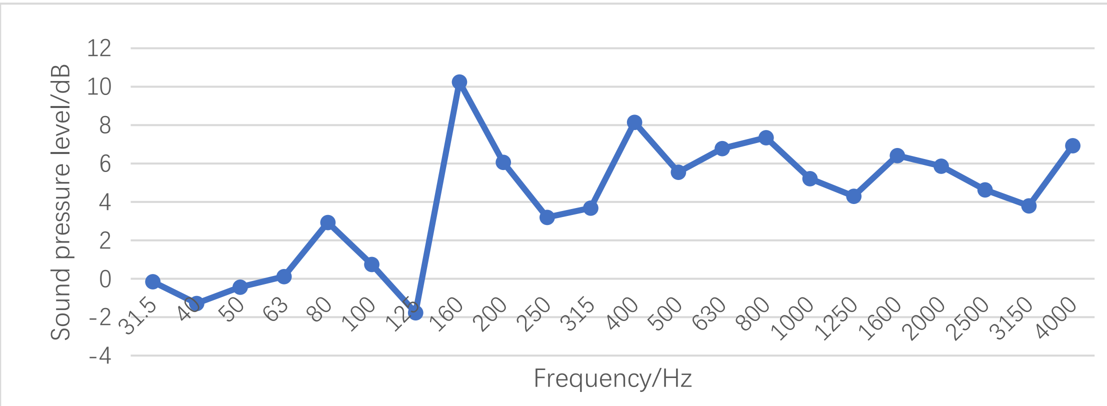


Fig. 3-7 Decreased SPL in the octave bands with ANC system

The signal-to-noise ratio (SNR) is a widely used index to evaluate the level of meaningful information in a signal with background noise. The SNR is defined as the power ratio between the signal with meaningful information and the pure noise signal:

$$SNR = 10\log_{10} \frac{\sum_{n} s_p{}^2(n)}{\sum_{n} \left[s_p(n) - \hat{s}_p(n) \right]^2}, \tag{3-5}$$

where $s_p(n)$ and $\hat{s}_p(n)$ are the clean speech signal and the noisy speech signal respectively.

Based on the SNR, the segmental SNR is proposed to consider human perception. The signal is divided into frames of about 20-40ms to obtain the approximate short time stationary signals. The segmental SNR is the mean SNR of the time frames. *M* denotes the number of frames, and *N* denotes the length of a frame.

$$SegSNR = \frac{10}{M} \sum_{i=1}^{M-1} \log_{10} \frac{\sum_{i=1}^{N-1} s_p{}^2(i,n)}{\sum_{i=1}^{N-1} \left[s_p(i,n) - \hat{s}_p(i,n) \right]^2}, \tag{3-6}$$

The evaluation index of the simulation results is presented in Table 3-1. The table shows that the clarity of the received speech signal is improved significantly when the ANC system is used.

Table 3-1 Evaluation index of the simulation results with ANC system

Noise	Factory 1		Factory 2	
Index	SNR/dB	SegSNR/dB	SNR/dB	SegSNR/dB
ANC off	3.58	-0.44	4.13	0.76
ANC on	8.77	4.38	8.02	4.08
Increased	5.19	4.82	3.89	3.32

3.4 Conclusion

The FxLMS algorithm is described in this chapter. The principle of the algorithm is introduced briefly. Furthermore, the performance of the FxLMS algorithm is investigated and shown by running the simulation model. Comparing the simulation results, the conclusion can be drawn that the FxLMS algorithm can reduce the background noise significantly although the denoising effect of the algorithm in the reverberant condition is worse than in the non-reverberant condition. The evaluation index also demonstrates the same conclusion.

Chapter 4 Inverse filter

4.1 Mathematical principle

The speech reverberation can be regarded as the convolution of the speech signal and RIR. The key point of de-reverberation is to deconvolute the received signal. In mathematics, deconvolution is an algorithm-based process used to invert the convolution in recorded data [36]. And it is easier to deconvolute in the frequency domain than the time domain because the deconvolution becomes division when the domain is changed.

For the room impulse response *h(t)*, the speech signal with reverberation received by the microphone can be express as:

$$y'(t) = y(t) * h(t), \tag{4-1}$$

where *y(t)* is the original speech signal, *y'(t)* is the speech signal filtered by the RIR, and * denotes the convolution.

Eq. (4-1) is changed to the frequency domain as follows:

$$Y'(\omega) = Y(\omega) \times H(\omega), \tag{4-2}$$

where $Y'(\omega)$, $Y'(\omega)$ and $H(\omega)$ are functions formed by the fast Fourier transfer (FFT) corresponding to *y(t)*, *y'(t)* and *h(t)*, respectively.

By taking the inverse of $H(\omega)$, the inverse filter response in the frequency domain, $iH(\omega)$, is obtained:

$$iH(\omega) = 1 / H(\omega), \tag{4-3}$$

The filtered speech signal processed by the inverse filter:

$$Y'(\omega) \times iH(\omega) = Y(\omega) \times H(\omega) \times iH(\omega) = Y(\omega), \tag{4-4}$$

Eq. (4-4) is then converted to the time domain with the inverse fast Fourier transform (IFFT):

$$y'(t) * ih(t) = y(t) * h(t) * ih(t) = y(t) \,, \tag{4-5}$$

the clean speech signal would be obtained in theory since *ih(t)* is used to counteract the effect of *h(t)* (RIR).

A modeling delay is also used in the calculation procedures to obtain a causal and stable inverse filter [17].

4.2 Simulation model

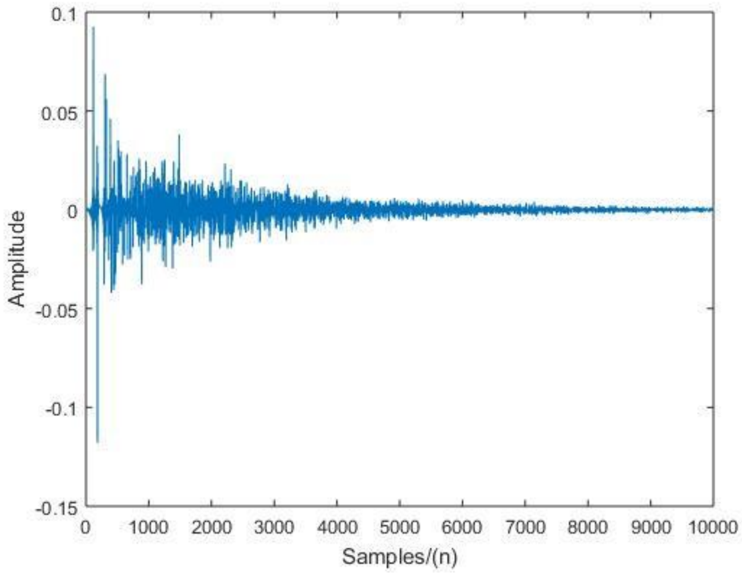


Fig. 4-1 Estimated room impulse response

The RIR that is identified between the speaker and the point we want to control (the error microphone location) is shown in Fig. 4-1. The propagation path of the RIR is also the secondary path in the ANC system, so the RIR is also the impulse response of the secondary path. The RIR is the same as the transfer function in Fig. 3-4. As mentioned, the reverberation is simulated by the image source model method, and T60 is set as 1.5s.

The inverse filter response based on the RIR is shown in Fig. 4-2. The inverse filter response is the reciprocal of RIR. The modeling delay is used to get a stable and causal filter [17]. Fig. 4-3 shows the convolution of the RIR and the inverse RIR. Although there is still some noise, the impulse signal is almost recovered. This means the reverberation is deconvoluted, and the original signal is recovered.

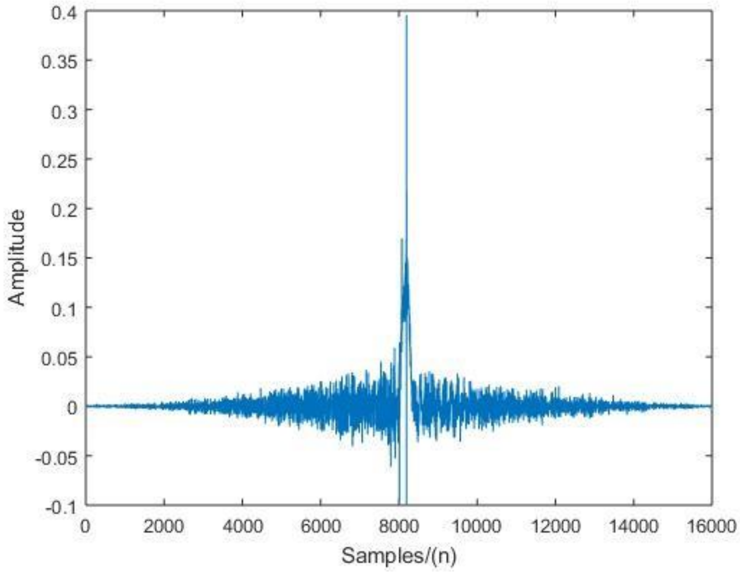


Fig. 4-2 Inverse room impulse response

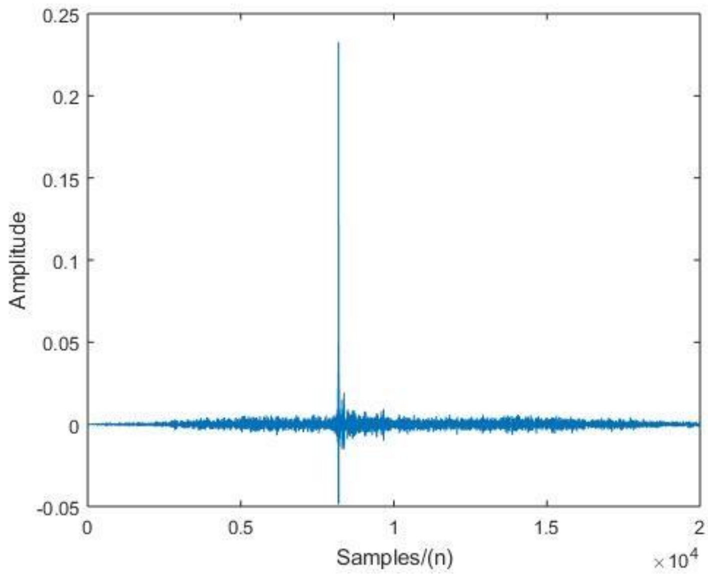


Fig. 4-3 Convolution of RIR and Inverse RIR

4.3 Simulation result

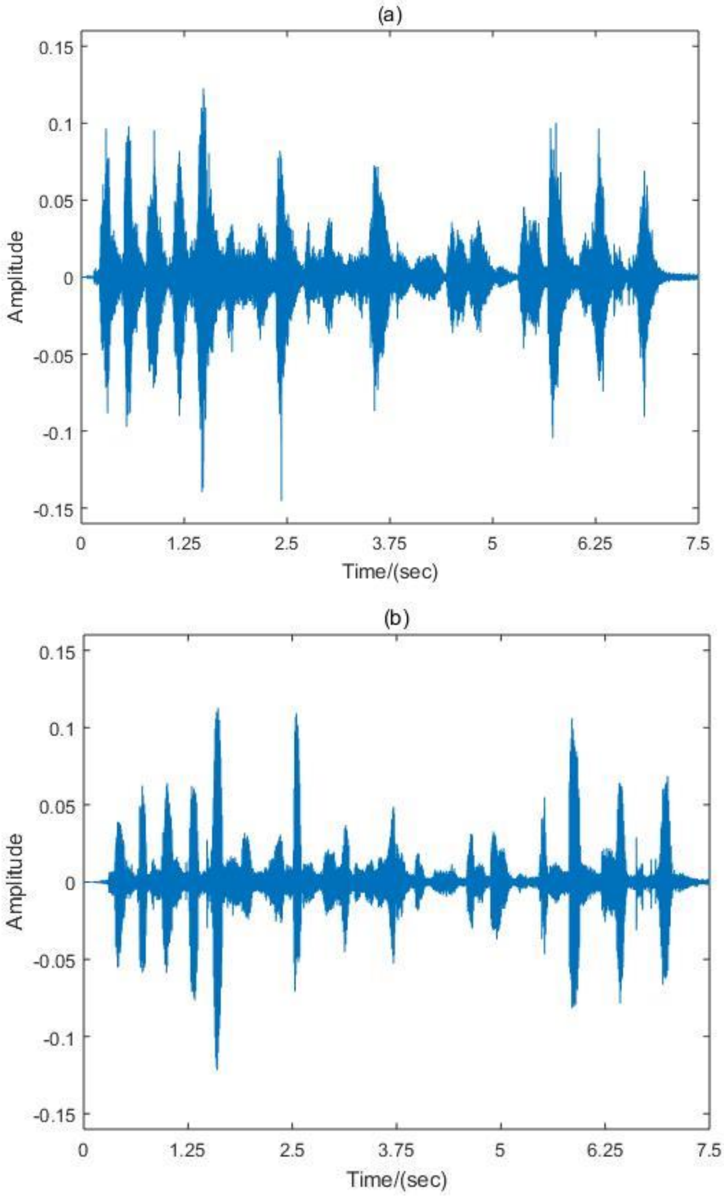


Fig. 4-4 Received speech signal with reverberation.
(a) No inverse filter, (b) Inverse filter.

The speech signals obtained with and without the inverse filter are shown in Fig. 4-4. The speech signal in Fig. 4-4 (b) is preprocessed by the inverse filter. Both the preprocessed and non-preprocessed speech signals by the inverse filter propagated in the simulation model, and the received speech signals with different processing methods are obtained. The signal in Fig. 4-4 (b) is more distinct and clearer than the one in Fig. 4-4 (a). This means there is less speech reverberation when the original speech signal is preprocessed by the inverse filter.

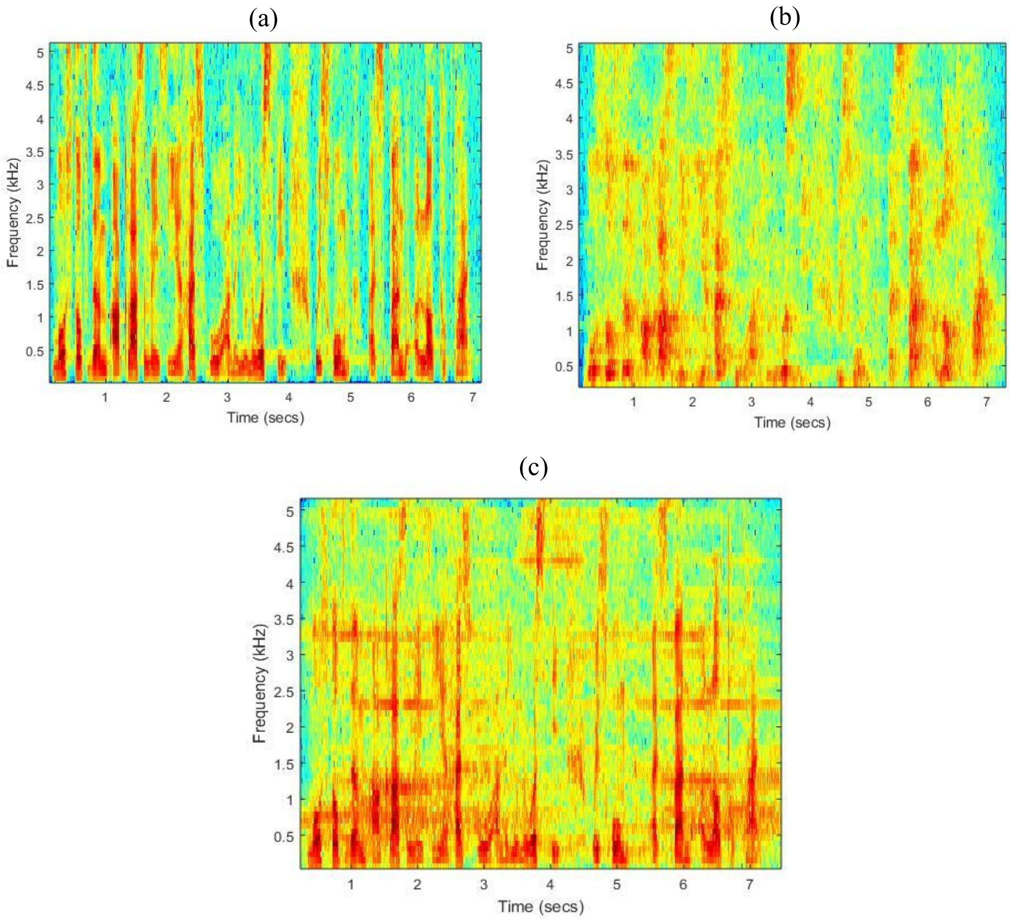


Fig. 4-5 Spectrograms of received speech signal in different situations.
(a) clean speech signal in the non-reverberant condition,
(b) speech signal in the reverberant condition,
(c) speech signal preprocessed by the inverse filter in the reverberant condition

Fig. 4-5 displays the spectrograms of the received speech in different situations. The clean speech signal received in the non-reverberant condition is presented in Fig. 4-5(a). Fig. 4-5(b) and Fig. 4-5(c) display the received speech signal preprocessed by the inverse filter or not in the reverberant condition respectively. Comparing the spectrograms, it is easy to observe that the energy of the speech signal is more intensive, and the trailing points are less when the speech signal is preprocessed by the inverse filter. It means that the speech is clearer and the speech reverberation is less.

However, some lines can also be observed which run through the entire time axes. It denotes there is the pure frequency noise is generated. Though the pure frequency noise influences the quality of the received speech signal little, it has an impact on the perceptual listening feeling. It will be discussed and solved in the subsequent studies.

4.4 Conclusion

This chapter introduces the inverse filter based on the fast inverse filter (FIF) algorithm. The theoretical mathematic equations are derived briefly. Inverse filter response is the reciprocal of RIR. The original speech signal is preprocessed by the inverse filter to reduce reverberation of the received speech signal. The room impulse response is acquired from the FxLMS algorithm because the propagation path of RIR is the same as the secondary path in the FxLMS algorithm. It means that the transfer function is measured once and used in two above algorithms. The speech spectrograms could demonstrate that the inverse filter can reduce the speech reverberation significantly.

Chapter 5 ANC-IF method

5.1 Introduction

Based on the simulation results, the ANC-IF method is proposed that combined the FxLMS algorithm and the FIF algorithm. Fig. 5-1 is the block diagram of the ANC-IF method.

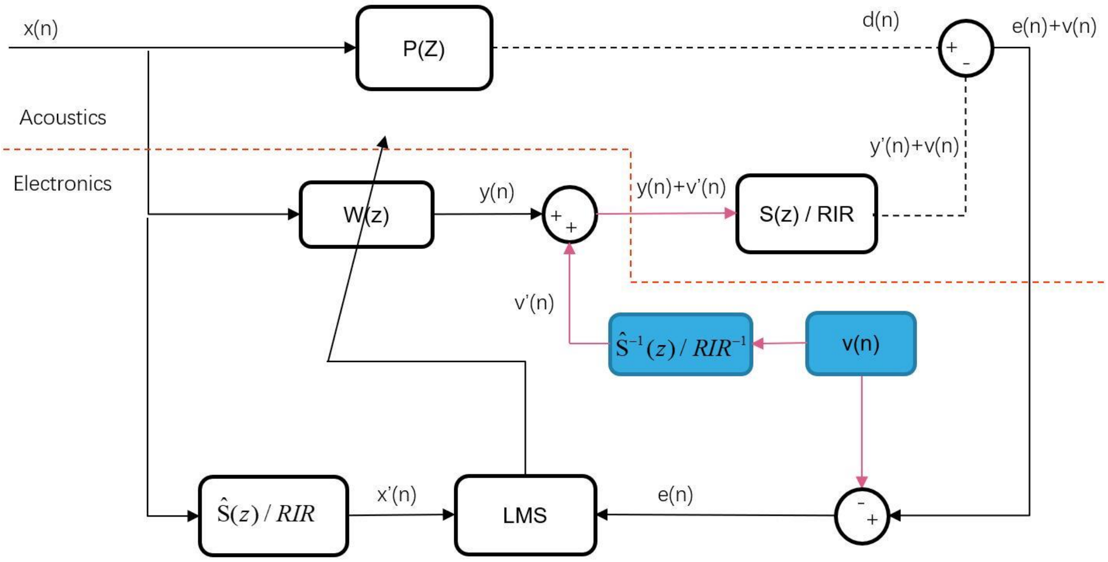


Fig. 5-1 Block diagram of the ANC-IF method

v(n) is the clean speech signal that the microphone should receive in an ideal situation. The FIF algorithm is added in the FxLMS algorithm to control the speech reverberation. $\mathrm{S}(z)$ is both the secondary path transfer function and RIR. $\hat{\mathrm{S}}^{-1}(z)$ is the reciprocal of $\hat{\mathrm{S}}(z)$, represents the inverse filter response:

$$\hat{\mathrm{S}}^{-1}(z) = 1/\hat{\mathrm{S}}(z)\,, \tag{5-1}$$

$$v'(n) = v(n) * \hat{\mathrm{s}}^{-1}(n)\,, \tag{5-2}$$

$$y'(n) = y(n) * \mathrm{s}(n)\,, \tag{5-3}$$

$$[y(n) + v'(n)] * \mathrm{s}(n) = y'(n) + v(n) * \hat{s}^{-1}(n) * s(n) = y'(n) + v(n)\,, \tag{5-4}$$

$$d(n) - [y'(n) + v(n)] = e(n) + v(n)\,, \tag{5-5}$$

Eq. (5-1) - (5-5) illustrate the mathematical principles of the proposed method. $e(n)+v(n)$ denote the received speech signal with residual noise and no reverberation as shown in Fig. 5-1.

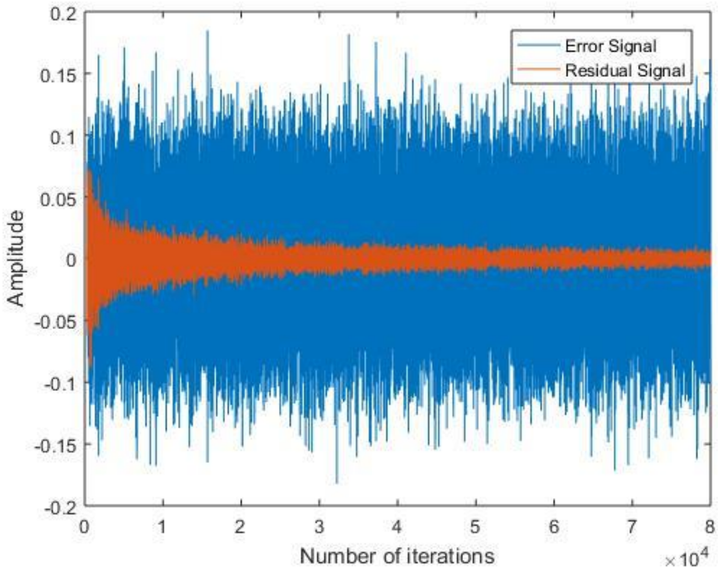


Fig. 5-2 Transfer function identification

Fig. 5-2 illustrates that the transfer function of the propagation path is identified. As shown in the figure, the value of the residual signal is decreased rapidly at the beginning of the identification process. It means the filter coefficients are quickly adjusted to adaptive the transfer function. Thus, at the end of the identification process, the residual signal is tiny than the error signal. It indicates that the filter could represent the transfer function of the propagation path accurately.

Fig. 5-3 displays the wave propagation in the simulation model. The noise source produces the factory noise as the background noise. Meanwhile, the speaker generates the sound wave that includes the denoising sound wave calculated based on the FxLMS algorithm and the speech signal preprocessed by the inverse filter. The sensor position is set at the point which is attempted to control the background noise and the speech reverberation.

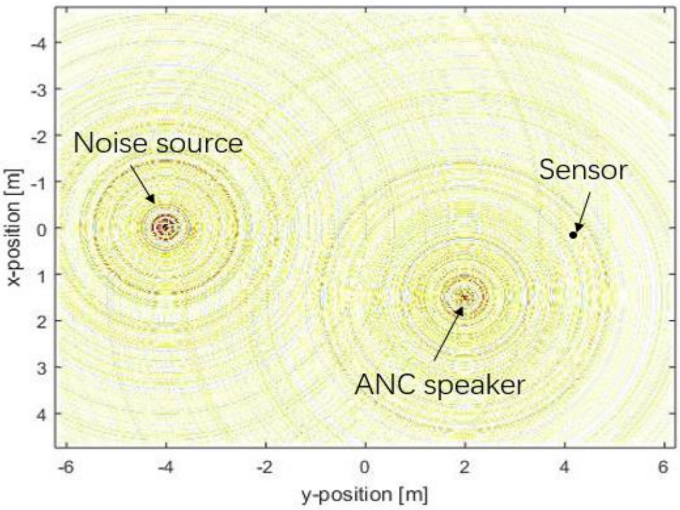


Fig. 5-3 Wave propagation diagram

5.2 Simulation result

Fig. 5-4 displays the received speech signal in the time domain in different situations. Comparing the figures, it can be deduced that the background noise is less when ANC system is used. Although the background noise has a serious impact on the distinction of the received speech signal, it can be observed that the speech signal preprocessed by the inverse filter is more approximate to the clean speech signal in the non-reverberant condition than the one which is not preprocessed.

The spectrograms corresponding to the signal in the time domain in Fig. 5-5 can also support the conclusion above mentioned. Observed the figures, there are fewer noisy points in Fig. 5-5 (c) than in Fig. 5-5 (b). This means the background noise is reduced significantly by the ANC system. The signal is more distinct and clearer in Fig. 5-5 (d) than Fig. 5-5 (c). This demonstrates the speech reverberation is decreased after preprocessing the speech signal by the inverse filter based on RIR. The waterfall spectrograms of the received speech in different situations are shown in Fig. 5-6 as the reference evidence.

The evaluation index of the simulation results with different methods is listed in Table 5-1. From the Fig. 5-7, the conclusion can be drawn that the evaluation index increased considerably after preprocessing the speech signal by the inverse filter even the ANC system is used.

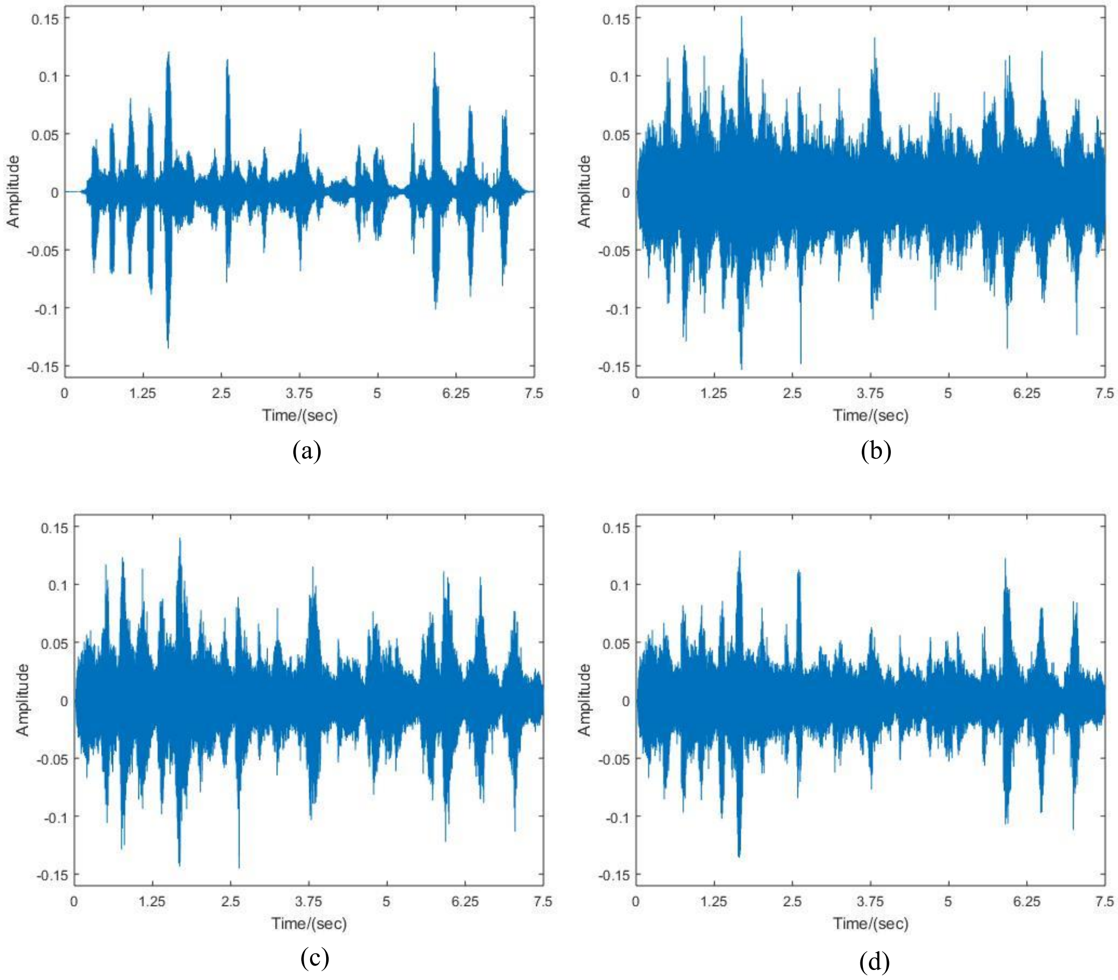
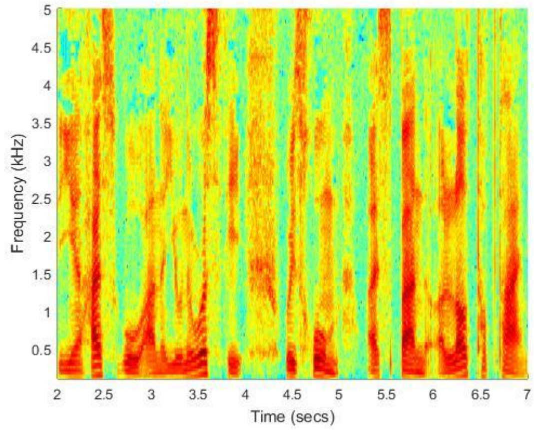
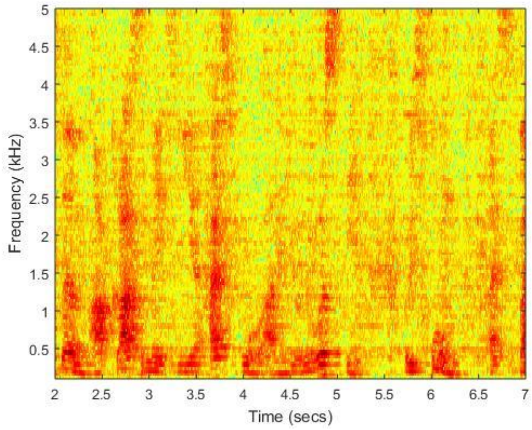


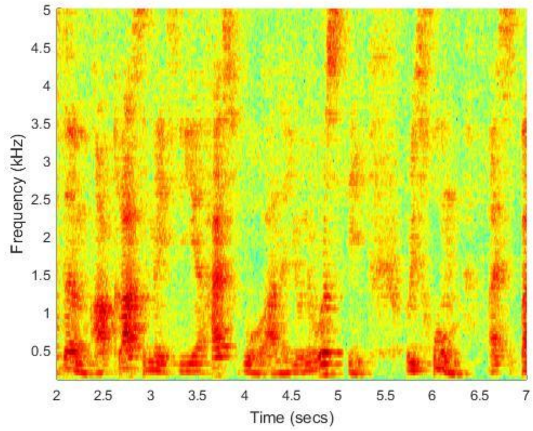
Fig. 5-4 Received speech signal in the time domain in different situations
(a) clean speech signal in non-reverberant conditions,
(b) speech signal in reverberant conditions with background noise,
(c) speech signal in reverberant conditions with background noise and ANC,
(d) speech signal in reverberant conditions with background noise and ANC and Inverse filter.



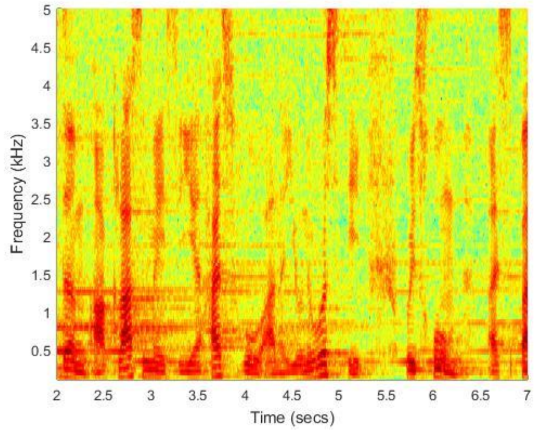
(a)



(b)



(c)



(d)

Fig. 5-5 Spectrograms of received speech signal in different situations.

(a) clean speech signal in non-reverberant conditions,

(b) speech signal in reverberant conditions with background noise,

(c) speech signal in reverberant conditions with background noise and ANC,

(d) speech signal in reverberant conditions with background noise and ANC and Inverse filter.

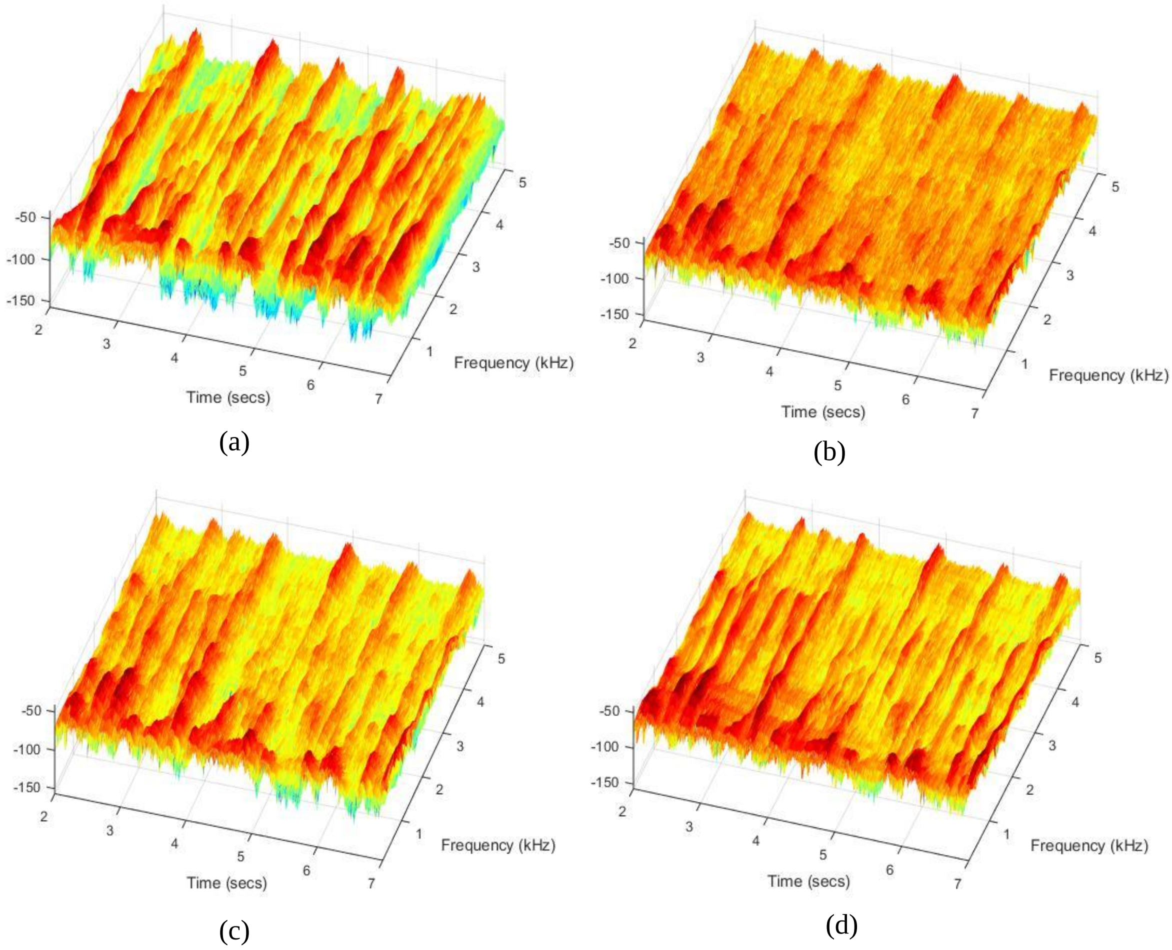


Fig. 5-6 Waterfall spectrogram of received speech signal in different situations.

(a) clean speech signal in non-reverberant conditions,

(b) speech signal in reverberant conditions with background noise,

(c) speech signal in reverberant conditions with background noise and ANC,

(d) speech signal in reverberant conditions with background noise and ANC and Inverse filter.

Table 5-1 Evaluation index of the simulation results with different methods

Noise	Factory 1		Factory 2	
Index	SNR/dB	SegSNR/dB	SNR/dB	SegSNR/dB
Original	3.58	-0.44	4.13	0.76
ANC	8.77	4.38	8.02	4.08
ANC+IF	13.62	7.38	12.82	7.02

Original means the noisy speech signal in the reverberant condition.

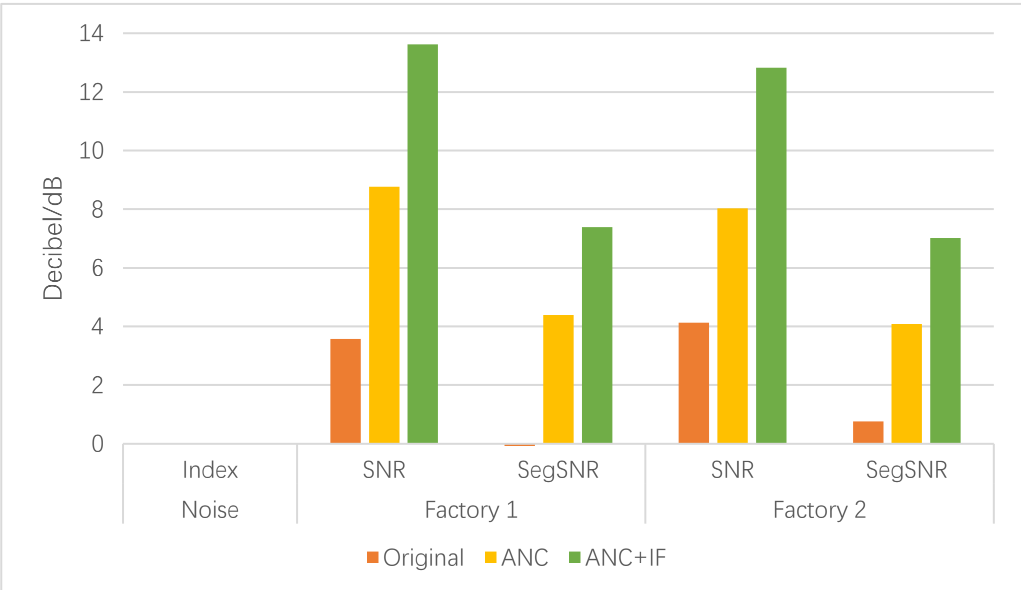


Fig. 5-7 Evaluation index of the simulation results with different methods

5.3 Conclusion

In this chapter, the ANC-IF method is proposed that combined the FxLMS algorithm and the FIF algorithm to control the background noise and the speech reverberation simultaneously. Since the secondary path in the FxLMS algorithm is also the propagation path of the RIR in the FIF algorithm, the transfer function of the secondary path is also used in the inverse filter as the RIR. Then the inverse filter is calculated based on the FIF algorithm and added in the FxLMS algorithm as the preprocessed filter of the speech signal. Moreover, the mathematical derivation illustrates the received signal only includes the speech signal with no reverberation and the residual noise denoised by ANC system. The simulation results demonstrate that the ANC-IF method is stable and efficient to control the background noise and reduce the speech reverberation in the room simultaneously. The conclusion, which the speech clarity is improved considerably by the ANC-IF method, is supported by the spectrograms of the received speech and the evaluation index.

Chapter 6 Conclusions

In this thesis, the ANC-IF method is proposed by combining the ANC method and the inverse filter method to reduce the background noise and speech reverberation simultaneously.

The propagation simulation model is established to investigate the performance of the method. To obtain the more realistic result, the reverberation is simulated by image source model method (ISM) and added to the propagation model. The noise source produces the factory noise that is acquired from NoiseX-92 database [34] as the background noise

The FxLMS algorithm used in the reverberant condition is simulated firstly. The fundamental principle of the algorithm is introduced, and mathematical equations are derived. While the anti-phase signal is calculated by the FxLMS algorithm to neutralize the background noise, the ANC speaker generates the opposite phase noise sound wave and the speech signal simultaneously. The simulation results demonstrate that the FxLMS algorithm is beneficial to improve the speech clarity by reducing the background noise in the environment, though the denoising performance of the algorithm in the reverberant condition is worse than in the non-reverberant condition. The evaluation index of the simulation results also proves this conclusion.

Secondly, the inverse filter based on the fast inverse filter (FIF) algorithm is used to preprocess the original speech signal for reducing the speech reverberation. The FIF algorithm is utilized to obtain the stable and causal inverse filter. Then the inverse filter is described and used to improve the speech clarity by eliminating the speech reverberation. Run the simulation model, and we can observe that the received speech signal has less reverberation in the propagation path when the original speech signal is preprocessed by the inverse filter. The simulation results indicate that the speech signal preprocessed by the inverse filter can improve the sound clarity significantly. The same conclusion could be obtained subjectively when listening to the results.

At last, the ANC-IF method is proposed to reduce two types of noise simultaneously by combining the ANC system and the inverse filter. Both the FxLMS algorithm and the FIF algorithm are classical and easy to implement. The bridge between the two methods is to estimate the transfer function of the secondary path between the speaker and the microphone in the FxLMS algorithm since it is just the RIR in the FIF algorithm. Thus, the ANC speaker generates the synthetical signal that includes both the opposite phase noise signal calculated by the FxLMS algorithm and the speech signal preprocessed by the inverse filter. The speech spectrograms can exhibit the difference of the received speech signal processed by different algorithms. Comparing the spectrograms, there are less noisy points in the received speech signal when the ANC system is used. While the original speech signal is preprocessed by the inverse filter, the energy of the received speech signal is more focused on the frames. It implies the speech signal is received with less background noise and less speech reverberation. The simulation results demonstrate that the ANC-IF method is stable and efficient to control the background noise and the speech reverberation in the room simultaneously. The evaluation indexes show that the ANC-IF method improved the speech clarity considerably.

Due to the reverberant condition, there are tailing points in the transfer function of the secondary path, which leads the convergent result of the adaptive filter coefficients is difficult to obtain. So, although the FxLMS algorithm works efficiently in the reverberant condition, the denoising effect of the algorithm is worse than in the non-reverberant condition. Further work will be done to improve the ANC algorithm to increase the denoising efficiency of the ANC-IF method.

Bibliography

[1]. Future noise policy, European commission green paper, Commission of the European Communities, Brussels, 1996

[2]. Liang Wang and Woon Seng Gan, Analysis of misequalization in a narrowband active noise equalizer system, Journal of Sound and Vibration, v.311, no.3-5, 2008, p.1438(9)

[3]. M. de Diego, A. Gonzalez, M. Ferrer, G. Pinero, Multichannel active noise control system for local spectral reshaping of multifrequency noise, Journal of Sound and Vibration 274 (2004) 249-271

[4]. Sen M. Kuo, Yajie Yang, Broadband adaptive noise equalizer, IEEE Signal processing letters 3(8), August (1996) 234-235

[5]. James W. Kamman, Koorosh Naghshineh, A comparison of open-loop feedforward and close-loop methods for active noise control using volume velocity minimization, Applied Acoustics, 57 (1999), 29-37

[6]. A. Roure, Self-adaptive broadband active sound control system, Journal of Sound and Vibration, 101(3), 1985, 429-441

[7]. Muhanmad Tahir Akhtar, Masahide Abe, Masayuki Kawamata, Wataru Mitsuhashi, A simplified method for online acoustic feedback path modeling and neutralization in multichannel active noise control system, Signal Processing, 89(2009), 1090-1099

[8]. S. Veena, S. V. Narasimhan, Improved active noise control performance based on Laguerre lattice, Signal Processing, 84(2004), 695-707

[9]. Sen M. Kuo, Mansour Tahernezhadi, Li Ji, Frequency-domain periodic active noise control and equalization, IEEE Transactions on Speech and Audio Processing, 5(4), 1997, 348-358

[10]. Sen M. Kuo, Ravi K. Yenduri and Abhijit Gupta, Frequency-domain delayless active sound quality control algorithm, Journal of Sound and Vibration, 318,2008, 715-724

[11]. Z. Qiu, C.-M. Lee, Z. H. Xu, A multi-resolution filtered-x LMS algorithm based on discrete wavelet transform for active noise control, Mechanical Systems and Signal Processing, 66-67(2016), 458-469

[12]. Zhao SD, Han HL, Shi WX. Generation on mechanism and characteristics of intermittent exhaust noise. Acad J Xian Jiaotong Univ (Chinese ed.), 1999, 33(3), 95–9

[13]. Longman RW, Iterative learning control and repetitive control for engineering practice, Int J Control 2000, 73(10), 930–54

[14]. Z. H. Xu, C. -M. Lee, Z. Qiu, A Study of the virtual microphone algorithm for ANC system working in audio interference environment, Journal of low frequency noise, vibration and active control, Vol. 33, No. 2, 189-206, 2014

[15]. ST Neely, JB Allen, Invertibility of a room impulse response, The Journal of the Acoustical Society of America, 66(1), 165-169(1979)

[16]. H Tokuno, O Kirkeby, PA Nelson, et al., Inverse filter of sound reproduction system using regularization, IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences, 80(5), 809-820(1997)

[17]. O Kirkeby, PA Nelson, H Hamada, et al., Fast deconvolution of multichannel systems using regularization, IEEE Transactions on Speech and Audio Processing, 6(2), 189-194(1998)

[18]. O Kirkeby, PA Nelson, Digital filter design for inversion problems in sound reproduction, Journal of the Audio Engineering Society, 47(7/8), 583-595(1999)

[19]. A Kusumoto, T Arai, K Kinoshita, et al., Modulation enhancement of speech by a pre-processing algorithm for improving intelligibility in reverberant environments, Speech communication, 45(2), 101-113(2005)

[20]. N Hodoshima, T Arai, K Kinoshita, et al., Improving syllable identification by a preprocessing method reducing overlap-masking in reverberant environments, The Journal of the Acoustical Society of America, 119(6), 4055-4064 (2006)

[21]. JB Crespo, RC Hendriks, Multizone speech reinforcement, IEEE/ACM Transactions on Audio, Speech and Language Processing, 22(1), 54-66 (2014)

[22]. RC Hendriks, JB Crespo, J Jensen, et al., Optimal near-end speech intelligibility improvement incorporating additive noise and late reverberation under an approximation of the short-time SII, IEEE Transactions on Audio, Speech and Language Processing, 23(5), 851-862 (2015)

[23]. Huan-Yu Dong, Chang-Myung Lee, Speech intelligibility improvement in noisy reverberant environments based on speech enhancement and inverse filtering, Journal on Audio, Speech, and Music Processing, vol. 3, 2018

[24]. B. E. Treeby and B. T. Cox, k-Wave: MATLAB toolbox for the simulation and reconstruction of photoacoustic wave fields, Journal of Biomedical Optics, vol.15, no.2, p.021314, 2010.

[25]. B. E. Treeby, J. Jaros, A. P. Rendell, and B. T. Cox, Modeling nonlinear ultrasound propagation in heterogeneous media with power law absorption using a k-space pseudospectral method, Journal of the Acoustical Society of America, vol.131, no.6, pp.4324-4336, 2012.

[26]. A. D. Pierce, Acoustics: An Introduction to its Physical Principles and Applications, Acoustical Society of America, 1989.

[27]. J. Allen, D. Berkley, Image method for efficiently simulating small-room acoustics, Journal of the Acoustical Society of America, vol.65(4), pp.943-950, 1979.

[28]. P. Peterson, Simulating the response of multiple microphones to a single acoustic source in a reverberant room, Journal of the Acoustical Society of America, vol.80(5), pp.1527–1529, 1986.

[29]. E. Lehmann and A. Johansson, Prediction of energy decay in room impulse responses simulated with an image-source model, Journal of the Acoustical Society of America, vol.124(1), pp.269-277, 2008.

[30]. S. M. Kuo, D. R. Mogran, Active noise control systems-algorithms and DSP implementations, Wiley, New York, 1996

[31]. P. M. Clarkson, Optimal and Adaptive Signal Processing, CRC Press, Boca Raton, FL, 1993

[32]. B. Widrow and S. D. Stearns, Adaptive Signal Processing, Prentice-Hall, Englewood Cliffs, NJ, 1985

[33]. D. R. Morgan, An analysis of multiple correlation cancellations loops with a filter in the auxiliary path, IEEE Transactions on Acoustics, Speech and Signal Processing, ASSP-28, 1980, 454-467

[34]. A Varga, HJ Steeneken, Assessment for automatic speech recognition: II. NOISE-92: A database and an experiment to study the effect of additive noise on speech recognition systems, Speech Communication, 12(3), 247-251(1993)

[35]. J. Allen and D. Berkley, Image method for efficiently simulating small-room acoustics, Journal of the Acoustical Society of America, vol. 65(4), pp. 943-950, April 1979.

[36]. O'Haver T., Intro to Signal Processing-Deconvolution, University of Maryland at College Park, 2007