



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

Ph.D. Thesis of the University of Ulsan

**Modified GSC Method to Reduce the Distortion of
Enhanced Speech Signal Using Cross-correlation and
Sidelobe Neutralization**

Department of Mechanical and Automotive Engineering

University of Ulsan

Ulsan, Korea

Hang Su

2021

**Modified GSC Method to Reduce the Distortion of
Enhanced Speech Signal Using Cross-correlation and
Sidelobe Neutralization**

Supervisor: Prof. Chang-Myung Lee

Author: Hang Su

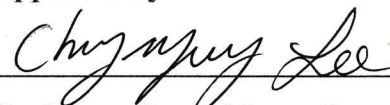
**Department of Mechanical and Automotive Engineering
University of Ulsan**

**A dissertation submitted to the faculty of the University of Ulsan in partial
fulfillment the requirement for the degree of Doctor of Philosophy in the
Department of Mechanical and Automotive Engineering.**

Ulsan, Korea






Dec. 15th, 2021

Approved by



Professor Chang-Myung Lee

SU HANG 의 공학박사학위 논문을 인준함

심사위원장	이병룡	
심사위원	김도중	
심사위원	박기서	
심사위원	이장명	
심사위원	전두환	

울산대학교 대학원

2021 년 12 월

ABSTRACT

Modified GSC Method to Reduce the Distortion of Enhanced Speech Signal Using Cross-correlation and Sidelobe Neutralization

Hang Su

Department of Mechanical and Automotive Engineering

The Graduate School

University of Ulsan

Online meetings are widely used today. A microphone array is used to extract the desired audio signal by analyzing the spatial information of signals. This dissertation proposes the modified GSC (generalized sidelobe canceller) method to reduce the distortion of the enhanced signal using cross-correlation and sidelobe neutralization.

The relative research background and academic achievement are displayed first. The fixed beamforming, LCMV (linearly constrained minimum variance), MVDR (minimum variance distortionless response), GSC, etc., are introduced. The GSC method is described in detail. The three parts of the GSC method could be optimized respectively to adapt to the various situations.

Next, the related fundamental knowledge about the proposed method is introduced. The formula of the beamforming pattern is derived first. Then the affection of the setup of a microphone array (radius of the microphone array and the number of the microphones) to the performance of the microphone array is displayed by figures. It illustrates that a larger diameter of the microphone array and a greater number of microphones in the microphone array could produce a better performance of spatial filtering. The correlation coefficient is also introduced and explained. The Person

correlation coefficient could be used to calculate the linear relationship between two variables and the rank correlation coefficient could give the variation tendency between two variables. The LMS (least mean square) algorithm is described as an adaptive algorithm of the adaptive filter. It also can be modified to adapt to various work situations.

Then the modified GSC method is proposed to reduce the distortion of the enhanced signal using cross-correlation and sidelobe neutralization. Distortion of the enhanced audio signal consists of two parts: the residual acoustic noise and the distortion of the desired audio signal, which means the damage to the desired audio signal. The modified GSC method is proposed to reduce both kinds of distortion if the desired signal is a nonstationary speech signal. First, the GSC-MCC (GSC method with the minimum cross-correlation coefficient) method is proposed that the cross-correlation coefficient between the canceling signal and the error signal of the LMS algorithm is added to the adaptive process of the GSC method. The cross-correlation coefficient is used to control the step size of the update process of the LMS algorithm to reduce the distortion of the enhanced signal while the energy of the desired signal frame is increased suddenly.

The sidelobe neutralization method is proposed to reduce the residual noise component in the output signal of the fixed beamforming method. This method could reduce the residual noise component effectively when the estimated noise direction is correct. The formula demonstrates that if the estimated noise direction is inaccurate, the amount of the noise component in the output signal of the beamforming method processed by the sidelobe neutralization method is still similar to the original noise component. Hence, the noise component of the beamforming output signal could be

decreased by subtracting the estimated noise signal to improve the denoising performance of the GSC method, which is referred to as the GSC-SN (GSC method with sidelobe neutralization) method.

Finally, the GSC-SN-MCC method is proposed by merging the above the GSC-MCC method and the GSC-SN method. The detailed performed steps are provided.

The experiment is performed in an anechoic chamber to validate the proposed method in various SNR (signal-noise ratio) conditions. The result demonstrates that the proposed method could reduce both kinds of noise effectively in various SNR conditions. The performance of the proposed method is like a parabolic curve with the SNR increasing, the best performance would be obtained when the energy of the speech and noise is almost equivalent, like the SNR is +5dB.

Furthermore, a simulated calculation with assumed inaccurate estimated noise directions is conducted based on the experiment data to inspect the robustness of the proposed method to the error of the estimated noise direction. The result shows that the denoising effect will be declined when the estimated noise direction becomes far away from the actual noise direction. However, even the estimated noise direction is opposite to the actual noise direction, the amount of the residual noise component is still similar to the original residual noise component, which implies the feasibility of the proposed method in practical cases.

The experiment data and calculation results indicate that the proposed method could reduce the distortion effectively under various SNR conditions and would not cause more distortion if the estimated noise direction is far from the actual noise direction.

ACKNOWLEDGMENTS

The heartfelt appreciation is presented to Prof. Chang-Myung Lee. For his academic guidance and financial support during my Ph.D. program.

The special thanks are offered to my dissertation committee members, Prof. Byung-Ryong Lee, Prof. Do-Joong Kim, Prof. Gi-Seo Park and Prof. Du-Hwan Chun. Their valuable comments make this dissertation more rational and precise.

The special thanks are also given to my cousin, Dr. ZhenHua Xu. His suggestions and encouragement inspire me to go further.

Thanks to my seniors, HuanYu Dong, GuangQuan Hou, Qi Wu and Peng Wang. Thanks for their academic suggestions and life care. I miss the time spent with them.

Thanks to my younger female schoolmate, Min Chen. I'm absolutely delighted that we can study and grow up together.

Thanks to my younger male schoolmate, Yu Shen. Thanks for his help in my study and life.

Heartful thanks to my parents. I can't accomplish the goal without their selfless support. They are my power source and the bay of my heart.

Sincere thanks to everyone I ever met, no matter what we have experienced, it will be a precious memory and treasure for me.

CONTENTS

ABSTRACT	i
ACKNOWLEDGMENTS	iv
CONTENTS.....	v
LIST OF FIGURES	vii
LIST OF TABLES	ix
ABBREVIATIONS	x
NOMENCLATURES.....	xii
Chapter 1 Introduction.....	1
1.1 Background	1
1.2 Methods of the microphone array.....	2
1.3 Outline.....	8
Chapter 2 Foundation Knowledge.....	9
2.1 Beamforming patterns	9
2.1.1 Introduction	9
2.1.2 Calculation formula.....	10
2.1.3 Influence factors	13
2.2 Correlation coefficient.....	22
2.2.1 Pearson correlation coefficient.....	22
2.2.2 Coefficient of rank correlation	25
2.3 LMS algorithm	28
2.4 Conclusion.....	30
Chapter 3 Modified GSC method.....	32
3.1 Conventional GSC method.....	32
3.2 TF-GSC method	33
3.3 GSC method with cross-correlation coefficient	35
3.4 GSC method with sidelobe neutralization	50
3.5 Proposed GSC-SN-MCC method.....	59

3.6 Conclusion.....	61
4. Experiment and Analysis.....	63
4.1 Experiment implementation	63
4.2 Experiment result analysis.....	64
4.2.1 Effect of the various SNR conditions	65
4.2.2 Effect of the inaccurate estimated noise direction.....	77
4.3 Conclusion.....	84
Chapter 5 Conclusions.....	86
Bibliography	90

LIST OF FIGURES

Figure 1-1. Diagram of the noise of the online meeting.	1
Figure 1-2. Diagram of delay-and-sum beamforming.	2
Figure 2-1. Diagram of UCA.	10
Figure 2-2. Bessel Function.	13
Figure 2-3. Beamforming Pattern of the microphone array with 4 microphones: (a) the radius is 0.1m; (b) the radius is 0.2m; (c) the radius is 0.4m.	15
Figure 2-4. Beamforming Pattern of the microphone array with 6 microphones: (a) the radius is 0.1m; (b) the radius is 0.2m; (c) the radius is 0.4m.	16
Figure 2-5. Beamforming Pattern of the microphone array with 8 microphones: (a) the radius is 0.1m; (b) the radius is 0.2m; (c) the radius is 0.4m.	18
Figure 2-6. Beamforming Pattern of the microphone array with 16 microphones: (a) the radius is 0.1m; (b) the radius is 0.2m; (c) the radius is 0.4m.	19
Figure 2-7. Correlation coefficients indicate relationships between variables.	23
Figure 2-8. Block diagram of the LMS algorithm.	29
Figure 2-9. Block diagram of the FxLMS algorithm.	30
Figure 3-1. Block diagram of the conventional GSC method.	33
Figure 3-2. Block diagram of the TF-GSC method.	34
Figure 3-3. Flowchart of the GSC with cross-correlation coefficient method.	36
Figure 3-4. Block diagram of the GSC method with the cross-correlation coefficient.	36
Figure 3-5. Clean speech signal and white noise signal on different SNR conditions.	39
Figure 3-6. Cross-correlation coefficient of different methods on different SNR conditions.	41

Figure 3-7 Comparison of the energy of the estimated noise frame with different methods on different SNR conditions: (a1~a6) all frames; (b1~b6) 50th to 100th frame.....	46
Figure 3-8. Distortion-noise ratio of frames with different methods from the 50th to 100th frame under different SNR conditions.	50
Figure 3-9. (a1~a6) Ratio of the noise processed using the sidelobe neutralization method with the different estimated noise directions and different actual noise directions. (b1~b6) Coefficient of sidelobe attenuation in the actual noise direction..	57
Figure 3-10. Block diagram of the GSC method with sidelobe neutralization.	59
Figure 3-11. Block diagram of the GSC-SN-MCC method.....	60
Figure 4-1. Experiment layout in the anechoic chamber.	64
Figure 4-2. Distortion-noise ratio of frames by different methods on different SNR conditions: (a1~a6) all frames; (b1~b6) from the 50th to 100th frame.....	69
Figure 4-3. Spectrum of the signal distortion of the 83rd frame by different methods on different SNR conditions: (a1~a6) entire frequency band; (b1~b6) from 500Hz to 2000Hz.	74
Figure 4-4. Distortion-noise ratio of frames (from 50th to 100th frame) by different methods when the estimated noise direction is different.	81

LIST OF TABLES

Table 2-1 The attenuation performance of microphone array with different hardware setups.....	21
Table 4-1. Average of the normalized distortion-noise ratio by different methods under various SNR conditions.....	76
Table 4-2. Normalized distortion ratio of the 83rd frame by different methods under various SNR conditions.....	76
Table 4-3. Average of the normalized distortion-noise ratio by different methods in various estimated noise directions.....	82
Table 4-4. Normalized distortion ratio of the 83rd frame by different methods in various estimated noise directions.....	83

ABBREVIATIONS

DSB	Delay-and-sum beamforming
MVDR	Minimum variance distortionless response
LCMV	Linearly constrained minimum variance
FB	Fixed beamforming
BM	Blocking matrix
ANC	Adaptive noise canceller
GSC	Generalized sidelobe canceller
DOA	Direction of arrival
TF-GSC	Transfer function GSC
DTF-GSC	Dual transfer function GSC
LMS	Least mean square
FxLMS	Filtered-x least mean square
ILC	Iterative learning control
RC	Repetitive control
NLMS	Normalized LMS
VAD	Voice activity detection
SN	Sidelobe neutralization
MSE	Mean square error
CC	Cross-correlation coefficient
ULA	Uniform linear array
UCA	Uniform circle array
URA	Uniform rectangular array
PCC	Pearson correlation coefficient
GSC-CC	GSC method with the cross-correlation coefficient
GSC-MCC	GSC method with the minimum cross-correlation coefficient
GSC-SN	GSC method with sidelobe neutralization
GSC-SN-MCC	GSC method with sidelobe neutralization and minimum cross-correlation coefficient

MEMS

Micro-Electro-Mechanical System

PCB

Printed circuit board

NOMENCLATURES

M	Microphone number of the microphone array
r	Radius of the microphone array
p_0	Received signal in the reference point of the microphone array
A_0	Amplitude of the received signal in the reference point
x_0	Distance from the sound source to the reference point
ω_k	Circular frequency
k	Wave number
c	Sound speed
f	Frequency
m	Sequence number of microphones
α_m	Local angle of microphones
Φ	Sound source incident angle
a_m	Relative distance of microphones to the origin
p_m	Received signal of microphones
v	Sound source direction vector
$\omega_{(s,m)}$	Weight of the microphone in the microphone array
Φ_s	Look direction of the microphone array
Φ_i	Incident direction of the sound source to the microphone array
ρ_{xy}	Pearson product-moment correlation coefficient
$Cov(x,y)$	Covariance of variables x and y
σ_x	Standard deviation of x
σ_y	Standard deviation of y
$Var(x)$	Variance of x
$Var(y)$	Variance of y
E	Mathematical expectation
τ_{xy}	Pearson product-moment correlation coefficient
\bar{x}	The mean of x
\bar{y}	The mean of y

$\rho_{R(x)R(y)}$	Spearman's rank correlation coefficient
$Cov(R(x), R(y))$	Covariance of the rank variables
$\sigma_{R(x)}$	Standard deviations of the rank variables $R(x)$
$\sigma_{R(y)}$	Standard deviations of the rank variables $R(y)$
d_i	Difference between the two ranks of each observation
τ	Kendall rank correlation coefficient
n_c	Number of the concordant pairs
n_d	Number of the discordant pairs
n	Total number of pair combinations
$x(n)$	Received signal of the microphone array
$y_m(n)$	Sum of the aligned signal of the beamformer
$y_b(n)$	Output canceling signal of the adaptive filter
$y_{out}(n)$	Final output signal of the microphone array
$z_b(n)$	Output signal of the blocking matrix
ω_f	Coefficients of the adaptive filter
$x(n)$	Received signals of the microphone array
$y_m(n)$	Sum of the aligned signals of the beamformer
$y_b(n)$	Output canceling signal of the adaptive filter
$y_{out}(n)$	Final output signal of the microphone array
$z_b(n)$	Output signals of the blocking matrix
ω_f	Coefficients of the adaptive filter
$Z_m(t, e^{j\omega})$	Fourier transform of the received time domain signals
W_0^H	Weights of the fixed beamformer
H^H	Weights of the blocking matrix
$H(e^{j\omega})$	Transfer function ratios of the microphones
$A_m(e^{j\omega})$	Acoustic transfer function to from the sound source to the m th microphone
$W_0(t, e^{j\omega})$	Weight of the fixed beamformer in frequency domain
$Y_{FBF}(t, e^{j\omega})$	Sum of the aligned signals of the beamformer in frequency domain

$U(t, e^{j\omega})$	Noise reference signals in frequency domain
$Y(t, e^{j\omega})$	Final output signal of the microphone array in frequency domain
$G_m(t, e^{j\omega})$	Coefficients of the adaptive filter in frequency domain
$\rho(k)$	Cross-correlation coefficient of the k th frame
$s_{speech}(k)$	Clean desired speech signal
$n_{residual}(k)$	Residual noise component of the output signal
$y_{final}(k)$	Output signal frame of the GSC-MCC method
$y_{out-corr}(k)$	Output signal frame of the GSC-CC method
$y_{out-conv}(k)$	Output signal frame of the conventional GSC method
y_{v1}	Output signal of the beamforming function in the direction v_1
y_{v2}	Output signal of the beamforming function in the direction v_2
y_{v1s2}	Output signal of the beamforming function in the direction v_1 which the sidelobe effect of the noise signal of the direction v_2 be subtracted
u_{ab}	Coefficient of the sidelobe attenuation when the sound source is placed at direction b and beamformed to direction a .
r_{v3}	Ratio that the residual noise signal processed by the sidelobe neutralization method with the inaccurate noise direction to by the fixed beamforming method
d	Minimum distance for the far-field condition
$distortion(k)$	Distortion of the k th frame of the enhanced speech signal
$p(k)$	Distortion-noise ratio as the distortion divided by the pure noise energy
r_p	Average of the distortion-noise ratio
$r_d(k)$	Distortion value of the k th frame

Chapter 1 Introduction

1.1 Background

Online meetings are widely used in the current time. The quality of meeting audio has become increasingly important. Though the audio hardware about recording and playing becomes more precise and higher dynamic range, some problems like interfering speech, background noise, and speech reverberation still are necessary to research further to improve the user experience.

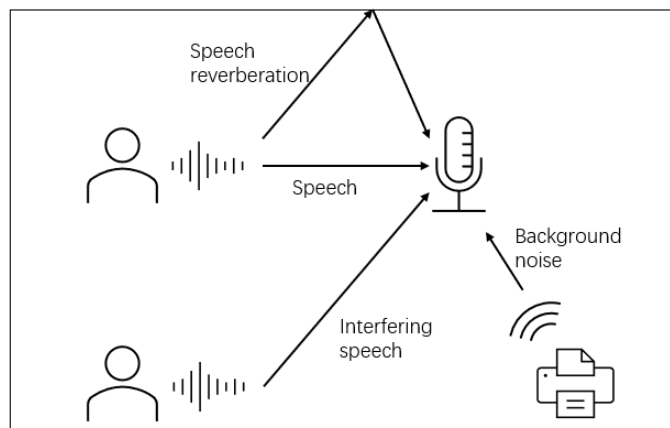


Figure 1-1. Diagram of the noise of the online meeting.

There are several thoughts to alleviate the problems. One is to use some signal-processing methods to improve the quality of meeting audio with a single-channel microphone, like the spectral subtraction method [1], Wiener filter [2], etc. However, these methods should work under several specific conditions, like noise is stationary and the correlation is low between the speech and the noise. Another thought is to deploy the microphone with every participant. It will raise the hardware cost and increase the meeting system complexity.

Then, the microphone array technology is proposed to solve the problems because all these problems are related to the spatial direction of the sound source. Therefore, a multichannel microphone array is used to extract the desired signal by analyzing the

spatial information of the received signals.

1.2 Methods of the microphone array

The primary method of a microphone array is the delay-and-sum beamforming (DSB) method presented by Flanagan et al. [3]. The multichannel signals are aligned at the looking direction by compensating with the proper delay, as shown in Figure 1-2. The aligned signals are summed to ensure that the gain of the looking direction is maximum. In other words, the signals from other directions are suppressed. By changing the delay of channels, the detected direction of the microphone array could be steered easily. Furthermore, the sound of multi-target in different directions could be obtained by steering the direction of the microphone array. At the same time, the DSB method is not required the particular demands on the sound source and is easy to combine with signal processing methods.

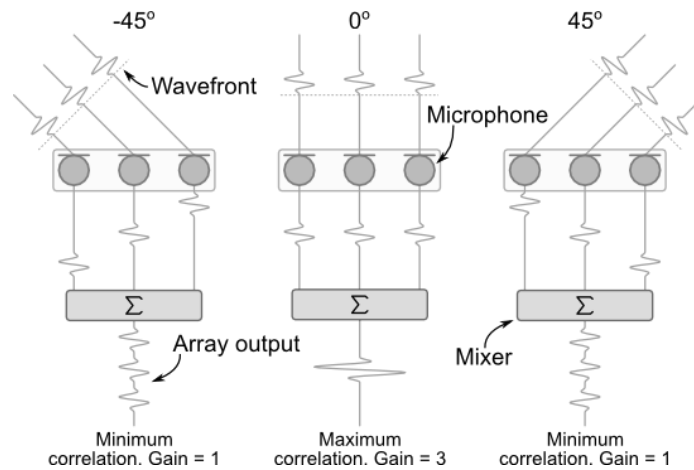


Figure 1-2. Diagram of delay-and-sum beamforming [4].

However, it is difficult to improve the efficiency of the microphone array unless to increase the number of microphones in the array. More microphones could suppress the non-looking direction signal more efficiently.

Several researchers studied to achieve better denoising performance without increasing the number of microphones [5]. Capon [6] proposed the minimum variance

distortionless response (MVDR) method in 1969. The weight coefficients of different channels were adjusted to preserve the gain of the looking direction and decrease the total power of the final output signal. Habets [7] modified the MVDR method to adapt the indoor acoustic conditions in 2010. This method could work effectively with the correlation noise signal in the diffuse sound field. Pan [8] analyzed the influence of the different incident angles of the sound for the MVDR method with a line microphone array in 2014. The experiment result demonstrated that the gain of the microphone array was mainly determined by the incident angle of the sound and the number of microphones in the microphone array. Vincent [9] proposed a speech enhance method by measuring the coherency based on ad hoc microphone arrays in 2016. Habets [10] proposed a speech distortion and interference rejection constraint beamformer method. This method was derived to minimize the ambient noise power subject to specific constraints that allow a tradeoff between speech distortion and interference-plus-noise reduction.

Frost [11] developed the MVDR method that extended the constraint in the MVDR method to linear equations. Therefore, the weight coefficients could be adjusted adaptively based on the constraints. This method was referred to as the linearly constrained minimum variance (LCMV) method. Shmulik [12] modified the LCMV method to apply two microphone arrays in the noisy reverberant environment in 2009. Both stationary and non-stationary noise could be suppressed. The authors reconstructed the LCMV method with the fixed beamforming (FB), blocking matrix (BM) and adaptive noise canceller (ANC). 2016 Norholm [13] used the harmonic linear chirp model to describe the voice model for speech enhancement with the LCMV method. The article showed that the higher SNR is obtained by the LCMV method than

the Wiener filter.

The constraints could be independent of the LCMV method to simplify the algorithm. Griffiths and Jim [14] proposed the generalized sidelobe canceller (GSC) method in 1982. GSC method consists of three parts. The first part is the same as a conventional beam-forming method to reduce the noise roughly. The second part is a blocking matrix to generate a reference noise signal by removing the desired signal from the original multichannel signals. The third part is an adaptive filter to estimate the noise component of the first part's output signal. Then, the estimated noise is subtracted from the output signal of the first part to obtain a cleaner desired signal. The GSC method converts the optimization goal of the LCMV method from the optimal weights with constraints to the optimal weights without constraints. Therefore, a simple adaptive algorithm could be adopted to reduce the residual noise in the final output signal.

The three parts of the GSC method could be optimized individually to improve the aggregate denoising performance of the adaptive algorithm. Hoshuyama et al. [15] and Lee et al. [16] modified the blocking matrix to decrease the sensitivity of the GSC method to a mismatch between the estimated and actual direction of arrival (DOA) of the desired signal. The robustness of the GSC method was increased by reducing the leakage of the desired signal into the reference noise signal [17].

Gannot et al. [18] considered the complicated acoustic environments and proposed a transfer function GSC (TF-GSC) method. The blocking matrix of the TF-GSC method was modified with the transfer function ratio (between different microphones in an array) to adapt the GSC method to reverberation conditions. Reuven et al. [19] combined the TF-GSC with an acoustic echo cancelation to improve the performance

of the algorithm in noisy and reverberant environments. Reuven et al. [20] also extended the TF-GSC method to a dual transfer function GSC (DTF-GSC) method for double talk scenarios to cancel the nonstationary interferences signal. Rombouts et al. [21] estimated a room impulse response and a desired speech signal model jointly to reduce the computational complexity dramatically for specific applications. Krueger, et al. [22] estimated the acoustical transfer function ratios in the presence of stationary noise and tracked the eigenvector adaptively to obtain better noise and interference reduction. This method relied on solving a generalized eigenvalue problem in each frequency bin.

The GSC method also could be optimized by using efficient complex value arithmetic to reduce the calculation complexity [23] and introducing the external microphones into the local microphone array to improve the performance of speech estimates [24]. Kim [25] proposed an improved generalized sidelobe canceller utilizing a phase-error filter for multi-channel signal enhancement. The experimental results showed that the proposed method provides better perceptual evaluation and intelligibility scores under multiple noise conditions.

The least mean square (LMS) algorithm is a common adaptive algorithm for the GSC method to estimate the noise signal. The LMS algorithm adjusts the filter coefficients to make the mean square of the error signal (difference between the canceling signal and the received noise signal) is least [26,27]. The stochastic gradient descent method is used to ensure that the error signal could be decreased iteratively in real-time, which was derived by Widrow and Hoff [28] in 1960. Based on the calculation mode, the LMS algorithm can derive to other branched algorithms, like filtered-x least mean square (FxLMS) algorithm [29], iterative learning control (ILC)

algorithm [30], repetitive control (RC) [31] algorithm, etc.

Several modified LMS algorithms were proposed for some practical consideration. Nagumo and Noda [32] introduced the normalized LMS (NLMS) algorithm to make the algorithm convergence speed independent of the reference signal power because the step size of the filter coefficient updating was inversely proportional to the reference signal power. Shan and Kailath [33] proposed the correlation LMS algorithm, which adjusted the step size of the filter coefficient updating to be proportional to the cross-correlation coefficient between the reference signal and the error signal. This step size control scheme made the adaptive algorithm more robust to disturbances such as the system-measured noise. Gitlin et al. [34] proposed the leaky LMS algorithm, in which a leakage factor was added to the adaptive weight update path. The leakage factor could avoid the overflow of the unconstrained weight to increase the stability of the algorithm.

Some studies introduced post-filtering methods to the microphone array for additional noise reduction to the beamforming output signal. Gannot [35] used multichannel post-filter in a diffused noise field to increase noise reduction, Cohen [36,37] discriminated the non-stationary noise and the voice by multichannel post-filter to improve the performance of the microphone array. Some studies [38-40] decomposed the multichannel signal into two subspace domains, desired signal subspace domain and noise signal subspace domain, based on statistic features (like the singular value decomposition of the covariance matrix of the multichannel signal). The signal in the noise subspace domain was removed or suppressed and the desired signal was restored from the desired signal subspace domain. Some researchers [41,42] attempted to extract the desired signal that had been contaminated by noise without a signal model or transmission model, which can be referred to as blind source separation or independent

component analysis, most were based on the statistic features or a neural network. This kind of research is mainly in the laboratory stage because of the complexity of the calculation and the feasibility in actual environments.

The distortion of the enhanced signal consists of two parts: the residual noise and the distortion of the desired signal, which means that the desired signal is damaged. Most researchers attempted to reduce the residual noise of the output of the microphone array. Whereas, the desired signal is sometimes a nonstationary signal, as in speech application. The varying speech signal may degrade the denoising performance of the adaptive algorithm. Some researchers [43] employed the voice activity detection (VAD) method to make the adaptive filter only be adjusted in speech-absent frames and avoid the affection of a varying speech signal. On the other hand, though some researchers modified the VAD method to improve its performance [44,45], the VAD method is not always feasible, especially when the background noise is strong or the sound field is complex.

In this study, the conventional GSC method was modified to reduce the enhanced speech distortion without the VAD method. The formula of the beamforming pattern was derived first. Then the sidelobe in the desired direction of the estimated noise source was calculated and subtracted from the output signal of the beamforming. This method was called the sidelobe neutralization (SN) method. Based on the excess mean square error (MSE) of the LMS algorithm [46] and the theoretical limits of the noise reduction performance [47], the SN method could reduce the noise component in the beamforming output signal of the GSC method to realize reducing the residual noise of the enhanced signal. To decrease the distortion that caused by the change of the amplitude of the speech signal, the step size of the adaptive algorithm in the GSC

method was controlled by the cross-correlation coefficient (CC) between the canceling signal and the error signal. This meant that when the canceling signal and the error signal are correlated, the step size would be big to change the weight coefficients of the adaptive filter rapidly. If the canceling signal and the error signal were uncorrelated, the step size would be small to slightly update the weight coefficients for reducing the distortion of the enhanced speech signal. This step size control method made the adaptive filter of the GSC method more robust to the energy variation of the desired signal.

1.3 Outline

This dissertation is organized as follows:

Chapter 1 introduces the research background and reviews the relative academic achievements.

Chapter 2 presents the foundation knowledge about beamforming patterns, correlation coefficient and the LMS algorithm.

Chapter 3 depicts the GSC method and modifies it with the cross-correlation coefficient and the SN method, respectively. Then the proposed modified GSC method is given by combining them.

Chapter 4 offers the implementation of the experiment and the experiment result analysis. The experiment is designed on different SNR conditions and different error angles of estimated incident noise direction.

Chapter 5 provides the conclusions of this study.

Chapter 2 Foundation Knowledge

2.1 Beamforming patterns

2.1.1 Introduction

Beamforming, also called spatial filtering, is a signal processing technique based on a sensor array to achieve spatial selectivity. It is achieved by combining elements in a sensor array so that signals at particular angles experience constructive interference while others experience destructive interference. The signal in the interesting direction would remain and interference in other directions would be rejected. The improvement compared with omnidirectional reception is known as the directivity of the array [48,49].

There are various kinds of microphone array configurations, like uniform linear array (ULA), uniform circle array (UCA), uniform rectangular array (URA), spiral-shaped array [50-52], etc. In most cases, the complex array configuration is designed to improve the performance of the array, higher dynamic range or more narrow spatial resolution. At the same time, it is inevitable to make the microphone array system more complex to implement in hardware and increase calculation burden in software.

For analysis and implementation easy, the UCA is chosen to validate the effectiveness of the algorithm in this study, as shown in Figure 2-1. At the same time, the elevation angle of the sound source is assumed as 0° , which means the sound source and the microphone array are in the same plane to simplify the calculation and display the beamforming pattern with more clarity. The beamforming pattern on the plane in which the elevation angle is 0° could be mapped to other horizontal planes (other elevation angles) with the appropriate trigonometric function about the elevation angle.

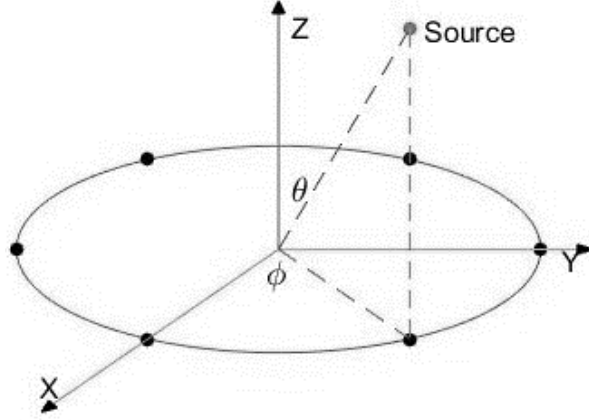


Figure 2-1. Diagram of UCA.

2.1.2 Calculation formula

In this study, the microphone array is assumed to be a uniform circle array (UCA) with M microphones deployed evenly on a circle with a radius of r . The reference received signal (p_0) of the microphone array in a reference point (most time is the array center point) is shown in Equation 2-1. A_0 is the complex amplitude of sound source generated, also related to the wave propagation mode. x_0 is the wave propagation distance from the sound source to the reference point of the microphone array. ω_k is the circular frequency of the signal, k is the wavenumber of the signal as shown in Equation 2-2, where c is the sound speed, f is the corresponding frequency of the signal.

$$p_0 = A_0 e^{j(\omega_k t + kx_0)}. \quad (2-1)$$

$$k = 2 \times \pi \times f / c. \quad (2-2)$$

Set the center point of the microphone array as the origin of the Cartesian coordinate system. The microphone is deployed uniformly around the origin. The m th microphone is placed at angle α_m as Equation 2-3. Assume the sound source is placed far enough from the microphone array. Thus, the sound field near the microphone array is a plane wave field. Φ is the sound source incident angle. x_m is the relative distance of the m th microphone to the origin. Thus, the signals of microphones received (p_m) are

represented by multiplying the proper phase shift as Equation 2-5. Furthermore, the phase shift of every microphone could be extracted as the sound source direction vector v as Equation 2-6.

$$\alpha_m = 2\pi(m-1) / M . \quad (2-3)$$

$$x_m = r \cdot \cos(\phi - \alpha_m) . \quad (2-4)$$

$$p_m = A_0 e^{j(\omega_k t + k(x_0 + x_m))} = A_0 e^{j(\omega_k t + kx_0)} e^{j(kx_m)} = p_0 e^{j(kx_m)} . \quad (2-5)$$

$$v = \left[e^{j(kx_1)}, e^{j(kx_2)}, \dots, e^{j(kx_m)}, \dots, e^{j(kx_M)} \right]^T . \quad (2-6)$$

The received signal of reference point could be derived by inverting the phase shift that relative distance (between the microphone and reference point) caused as follows:

$$p_0 = p_m e^{j(-kx_m)} . \quad (2-7)$$

Let $\omega_{(s,m)}$ is the weight of the microphone channel to adjust the looking direction Φ_s of the microphone array. Φ_i represents the actual incident direction of the sound source. Thus, the beamforming pattern could be derived using the following equations [53]:

$$\begin{aligned} B(\phi_s, \phi_i) &= \sum_{m=1}^M \omega_{(s,m)}^* V_{(i,m)} \\ &= \frac{1}{M} \sum_{m=1}^M e^{jkr(\cos(\phi_i - \alpha_m) - \cos(\phi_s - \alpha_m))} \\ &= \frac{1}{M} \sum_{m=1}^M e^{j2kr \sin\left(\frac{\phi_s - \phi_i}{2}\right) \sin\left(\frac{\phi_s + \phi_i}{2} - \alpha_m\right)} , \\ &= \frac{1}{M} \sum_{m=1}^M \exp(-jz \sin(\beta - \alpha_m)) \end{aligned} \quad (2-8)$$

where $z = 2kr \sin\left(\frac{\phi_s - \phi_i}{2}\right)$, $\beta = \frac{\phi_s + \phi_i}{2}$.

Based on the Jacobi-Anger identity:

$$\exp(jz \sin(\alpha)) = \sum_{n=-\infty}^{\infty} J_n(z) \exp(jn\alpha) , \quad (2-9)$$

Thus, the function of the beamforming pattern could be modified to Equation 2-10.

$$\begin{aligned}
B(\phi_s, \phi_i) &= \frac{1}{M} \sum_{m=1}^M \sum_{n=-\infty}^{\infty} J_n(z) \exp(-jn(\beta - \alpha_m)) \\
&= \frac{1}{M} \sum_{m=1}^M \sum_{n=-\infty}^{\infty} J_n(z) \exp\left(-jn\left(\beta - 2\pi \frac{m-1}{M}\right)\right) \quad , \\
&= \sum_{n=-\infty}^{\infty} J_n(z) \exp(-jn\beta) \frac{1}{M} \sum_{m=1}^M \exp\left(j2\pi(m-1)\frac{n}{M}\right)
\end{aligned} \tag{2-10}$$

where J_n is the n th-order Bessel function.

Let $b = n/M$, if b is an integer, the Equation 2-11 is workable.

$$\frac{1}{M} \sum_{m=1}^M \exp(j2\pi(m-1)b) = 1, \tag{2-11}$$

If b is not integer, based on the geometric series formula:

$$S_n = \frac{a_1(1-q^n)}{1-q}, \tag{2-12}$$

$$\text{Hence, } \frac{1}{M} \sum_{m=1}^M \exp(j2\pi(m-1)b) = \frac{1 - \exp(j2\pi n)}{1 - \exp(j2\pi b)} = 0, \tag{2-13}$$

Finally, the function of the beamforming pattern could be simplified as Equation 2-14.

If M is even, the equation could be simplified further to Equation 2-15.

$$B(\phi_s, \phi_i) = \sum_{b=-\infty}^{\infty} J_{bM}(z) \exp(-jbM\beta). \tag{2-14}$$

$$B(\phi_s, \phi_i) = J_0(z) + 2 \sum_{b=1}^{\infty} J_{bM}(z) \cos(bM\beta). \tag{2-15}$$

Consequently, the beamforming pattern could be formed by the Bessel functions.

Considering the property of the Bessel function as shown in Figure 2-2, the value of the second item of Equation 2-15 will be close to zero and can be neglected with the increasing number of microphones. Thus, the beamforming pattern could be approximated as a zero-order Bessel function.

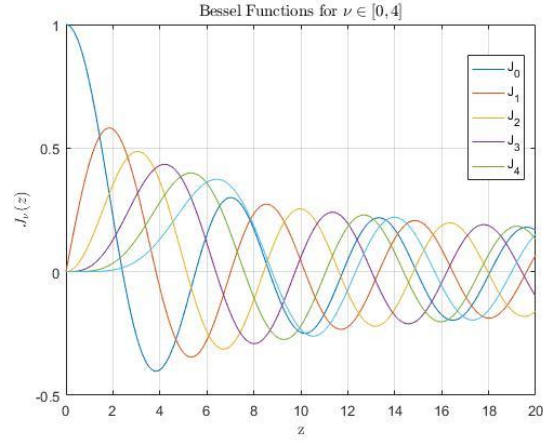


Figure 2-2. Bessel Function.

When $|z| < n$, the result of Bessel function $|J_n(z)|$ is small that makes the second item of Equation 2-15 be neglected. Thus, Equation 2-17 could be derived from Equation 2-16. It means that if the arc length between the microphones is less than the half wavelength of the interesting frequency, the beamforming pattern would be similar to a zero-order Bessel function and that no grating lobe generated.

$$M \geq \max(|z|) = 2kr . \quad (2-16)$$

$$2\pi r / M \leq \lambda / 2 . \quad (2-17)$$

2.1.3 Influence factors

The beamforming pattern could be compared by adjusting factors of the microphone array. The radius of the microphone array and the microphone number were changed to show the affection of the beamforming pattern. The beamforming pattern was calculated according to Equation 2-8. The radius of the microphone was set as 0.1m, 0.2m and 0.4m. The microphone number of the microphone array is 4, 6, 8, 16, respectively.

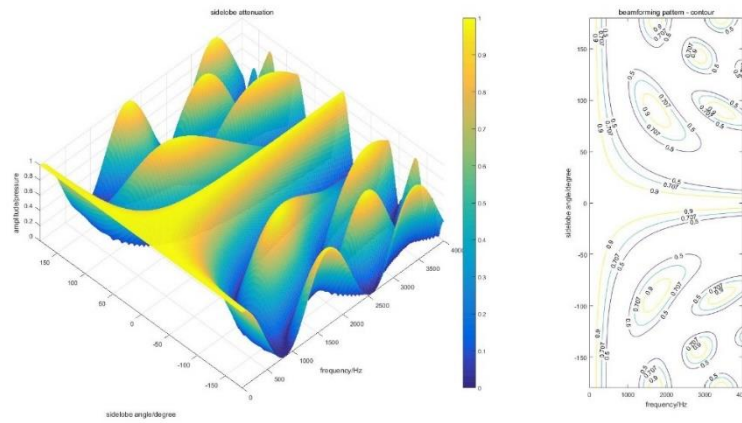
The beamforming patterns of the microphone array with different hardware setups are shown in Figure 2-3, Figure 2-4 and Figure 2-5, Figure 2-6. Figure 2-3-1 is selected to explain the meaning of the axes in the figure and the meaning of other figures is the

same as this figure.

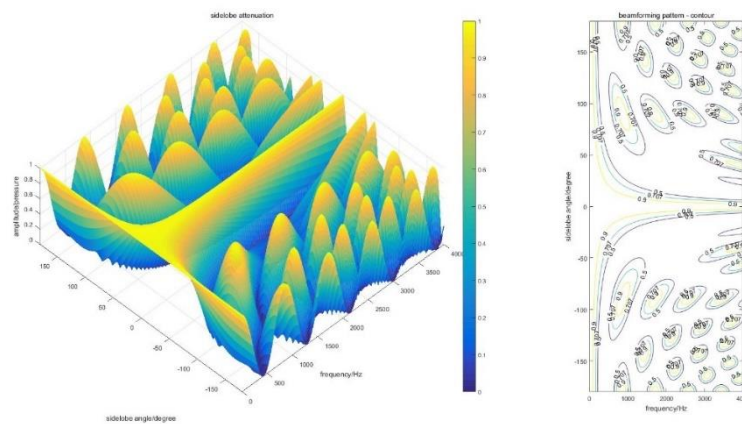
The beamforming pattern of the 4 microphones array with the 0.1m radius is shown in Figure 2-3(a). The left part of the figure is the 3-D beamforming pattern. The x-axis represents the angle of the sidelobe from -180° to 180° . The y-axis just shows the frequency of the signal from 0Hz to 4000Hz, because the energy of speech is concentrated in a low and medium frequency band. The z-axis means the coefficient of the pressure attenuation of the sidelobe. The right part of the figure displays the contour of the beamforming pattern. The ordinate axis represents the angle of the sidelobe and the abscissa axis represents the frequency of the signal. There are three contour curves shown in the figure. The yellow and blue curves in the figure mean the pressure attenuation of the sidelobe is 0.9 and 0.5 times to the desired direction of the microphone array. The green curve represents the pressure attenuation of the sidelobe is about 0.707 times to the desired direction of the microphone array, namely half the energy of the sidelobe signal is suppressed. This contour curve is selected to illustrate the efficiency of the microphone array with the different hardware setups. The sound speed is assumed as 343m/s.

(1) 4 microphones array

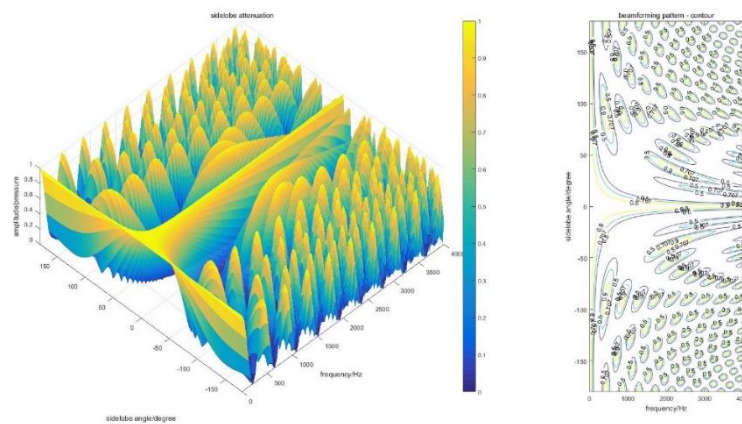
The number of microphones of the microphone array is 4. The radius of the microphone array is 0.1m, 0.2m, 0.4m, respectively. Based the Equation 2-17, if the radius of the microphone array is 0.1m, 0.2m, 0.4m, it means if the wavelength is more than 0.314m, 0.628m, 1.256m, namely the frequency is lower than 1092 Hz, 546 Hz, 273Hz, there will not grating lobes generated.



(a)



(b)



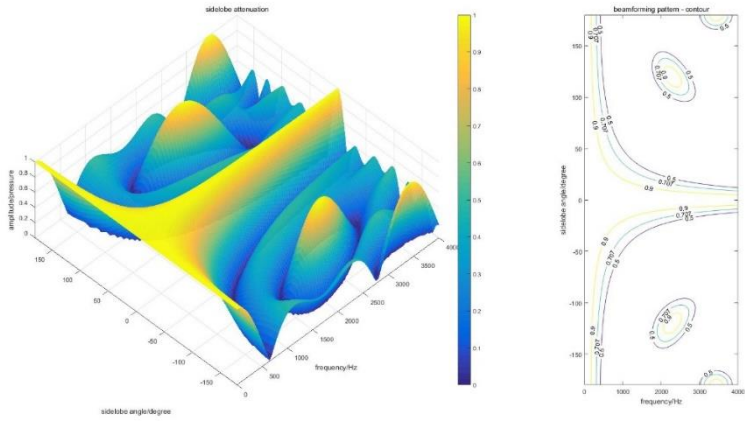
(c)

Figure 2-3. Beamforming Pattern of the microphone array with 4 microphones: (a) the radius is 0.1m; (b) the radius is 0.2m; (c) the radius is 0.4m.

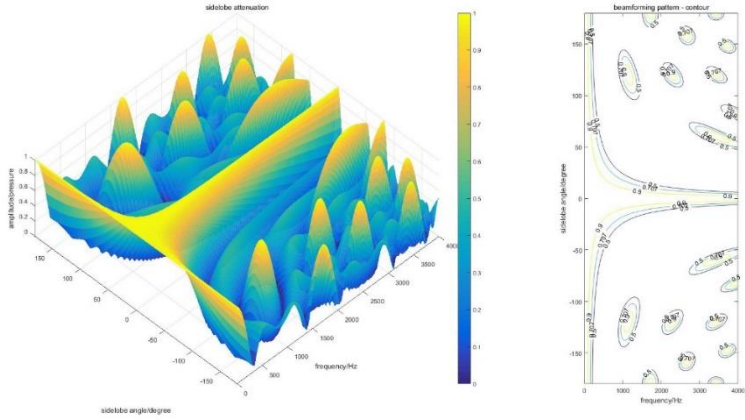
(2) 6 microphones array

The number of microphones of the microphone array is 6. The radius of the microphone array is 0.1m, 0.2m, 0.4m, respectively. Based the Equation 2-17, if the

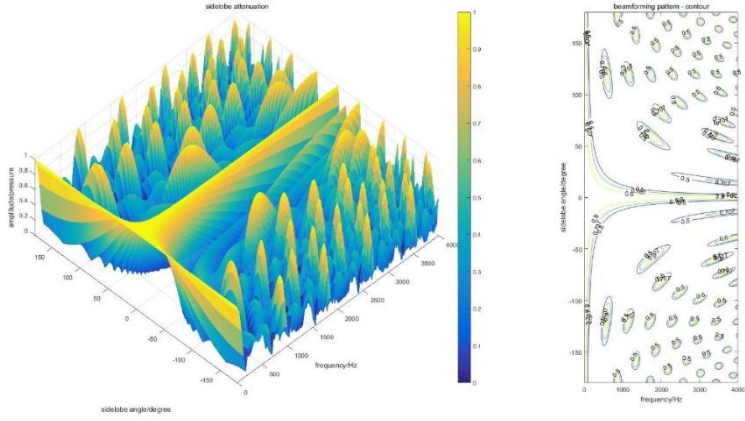
radius of the microphone array is 0.1m, 0.2m, 0.4m, it means if the wavelength is more than 0.209m, 0.419m, 0.837m, namely the frequency is lower than 1641Hz, 818Hz, 409Hz, there will not the grating lobes generated.



(a)



(b)

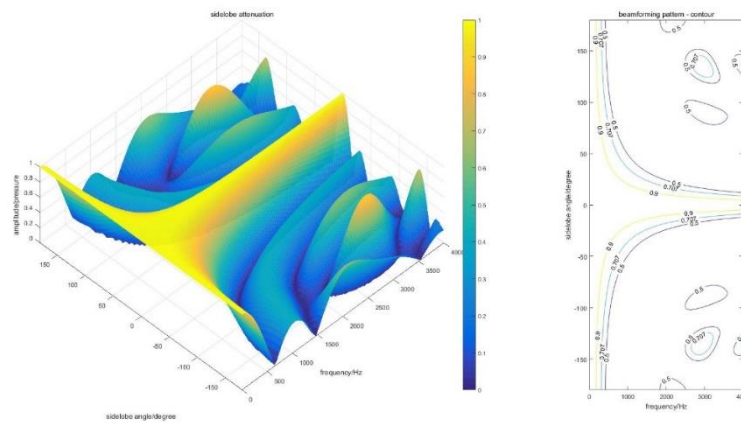


(c)

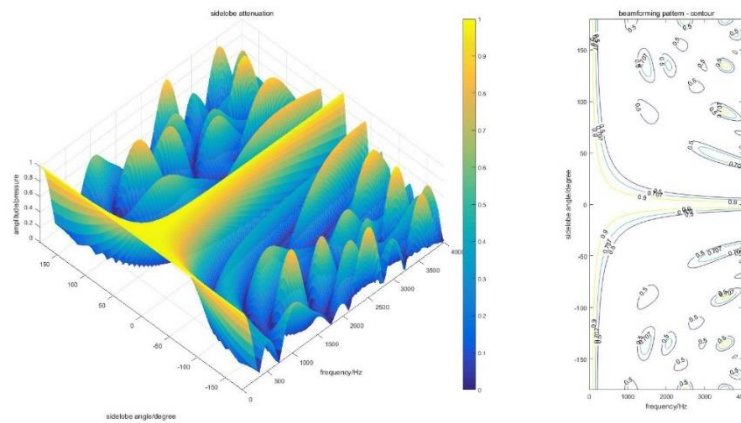
Figure 2-4. Beamforming Pattern of the microphone array with 6 microphones: (a) the radius is 0.1m; (b) the radius is 0.2m; (c) the radius is 0.4m.

(3) 8 microphones array

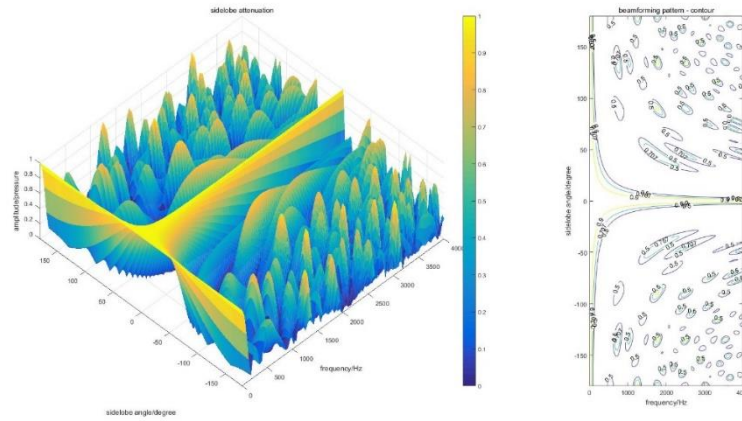
The number of microphones of the microphone array is 8. The radius of the microphone array is 0.1m, 0.2m, 0.4m, respectively. Based the Equation 2-17, if the radius of the microphone array is 0.1m, 0.2m, 0.4m, it means if the wavelength is more than 0.157m, 0.314m, 0.628m, namely the frequency is lower than 2184Hz, 1092 Hz, 546 Hz, there will not grating lobes generated.



(a)



(b)

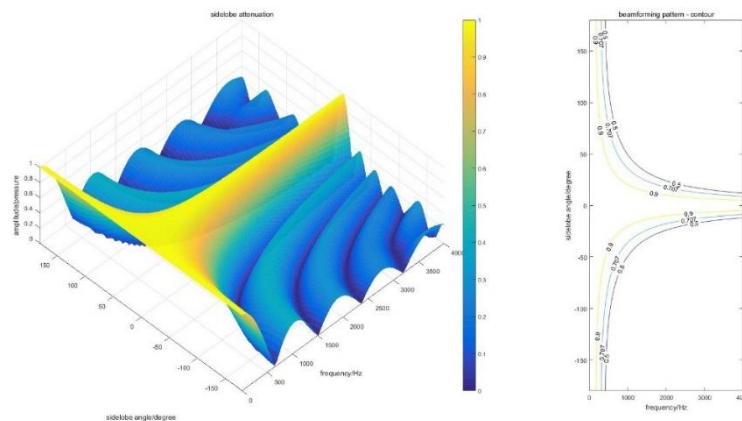


(c)

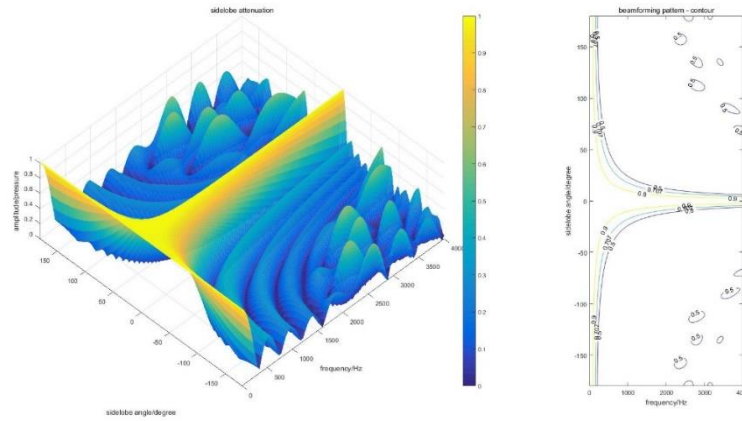
Figure 2-5. Beamforming Pattern of the microphone array with 8 microphones: (a) the radius is 0.1m; (b) the radius is 0.2m; (c) the radius is 0.4m.

(4) 16 microphones array

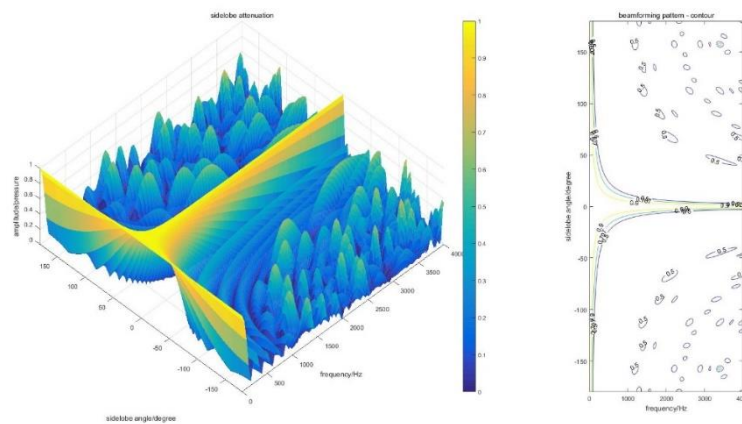
The number of microphones of the microphone array is 16. The radius of the microphone array is 0.1m, 0.2m, 0.4m, respectively. Based the Equation 2-17, if the radius of the microphone array is 0.1m, 0.2m, 0.4m, it means if the wavelength is more than 0.0785m, 0.157m, 0.314m, namely the frequency is lower than 4369Hz, 2184Hz, 1092Hz, there will not grating lobes generated.



(a)



(b)



(c)

Figure 2-6. Beamforming Pattern of the microphone array with 16 microphones: (a) the radius is 0.1m; (b) the radius is 0.2m; (c) the radius is 0.4m.

The beamforming pattern could be divided into 3 areas: main lobe area, low-frequency area and sidelobe area based on the 0.707 contour curve as shown in figures. In the main lobe area, most energy of the signal in the looking direction is preserved. In the low-frequency area, the wavelength of the signal is long that the phase shift of the received signal of the array is too small to distinguish the different incident directions. It means that the capability of the spatial filter of the microphone array is invalid. In the sidelobe area, most of the un-looking direction signal is suppressed according to the different attenuation factors. However, in several specific angles and specific frequencies, the grating lobes are generated and the attenuation factor is close to one. It means on a specific angle and specific frequency condition, the microphone

array loses the spatial filter capability as in the low-frequency area. When the microphone array loses the ability to attenuate the signal in an un-looking direction, it just works as a single channel microphone and should use the single-channel method to deal with the received signal.

Table 2-1 is established to exhibit the concrete value about the 3 areas with the different hardware setups. The first column of the table depicts the parameters of the microphone array. M04R01 means the microphone array consist of 4 microphones and the radius is 0.1m. The meaning of other items in the first column follows the same rules as the M04R01. The contour curve, which value is 0.707, means half the energy of the signal from the un-look incident direction is attenuated, is chosen as the criterion of spatial filter function of the microphone array is valid or not. The maximum frequency of the low-frequency area means when the frequency is lower than it, the signal attenuation will be less than 0.707, the spatial filter function of the microphone array is invalid. For comparison, the low-frequency area ratio is proposed that is the proportion of low-frequency area to the whole spectrum area.

The main lobe area is described by the angle of the sidelobe attenuation factor, which is 0.707 on several frequency bins (100Hz, 200Hz, 500Hz, 1000Hz, 2000Hz, 4000Hz). As similar to the low-frequency area ratio, the main lobe ratio is the proportion of the main lobe area (not including the overlap part with the low-frequency area) to the whole spectrum area. The grating lobes are shown in the figure if their attenuation factors are high than 0.707. The number of the grating lobes in which the attenuation factors are high than 0.707 is presented in the table. The grating lobe ratio is the proportion of the grating lobe area to the whole spectrum area. The total attenuation ratio considers the mentioned 3 areas ratio could be proposed as follows:

$$r_t = l - r_l - r_m - r_g. \quad (2-18)$$

Table 2-1 The attenuation performance of microphone array with different hardware setups

	Low frequency area (r_l)		Main lobe area (r_m)							Grating lobe area (r_g)		Total attenuation area ratio (r_t)
	max frequency /Hz	area ratio	100 Hz	200 Hz	500 Hz	1000 Hz	2000 Hz	4000 Hz	area ratio	number	area ratio	
M04R01	301.5	7.5%	/	/	$\pm 72^\circ$	$\pm 37^\circ$	$\pm 18^\circ$	$\pm 9^\circ$	14.4%	12	14.2%	63.9%
M06R01	301.5	7.5%	/	/	$\pm 73^\circ$	$\pm 36^\circ$	$\pm 18^\circ$	$\pm 9^\circ$	14.3%	4	4.0%	74.2%
M08R01	301.5	7.5%	/	/	$\pm 73^\circ$	$\pm 36^\circ$	$\pm 18^\circ$	$\pm 9^\circ$	14.3%	4	1.5%	76.7%
M16R01	301.5	7.5%	/	/	$\pm 73^\circ$	$\pm 36^\circ$	$\pm 18^\circ$	$\pm 9^\circ$	14.3%	0	0%	78.2%
M04R02	172.3	4.3%	/	$\pm 89^\circ$	$\pm 35^\circ$	$\pm 18^\circ$	$\pm 9^\circ$	$\pm 5^\circ$	9.0%	40	17.1%	69.6%
M06R02	172.3	4.3%	/	$\pm 91^\circ$	$\pm 35^\circ$	$\pm 18^\circ$	$\pm 9^\circ$	$\pm 4^\circ$	8.9%	18	7.2%	79.6%
M08R02	172.3	4.3%	/	$\pm 91^\circ$	$\pm 35^\circ$	$\pm 18^\circ$	$\pm 9^\circ$	$\pm 4^\circ$	8.9%	16	4.5%	82.3%
M16R02	172.3	4.3%	/	$\pm 91^\circ$	$\pm 35^\circ$	$\pm 18^\circ$	$\pm 9^\circ$	$\pm 4^\circ$	8.9%	0	0%	86.8%
M04R04	86.1	2.2%	$\pm 125^\circ$	$\pm 42^\circ$	$\pm 17^\circ$	$\pm 9^\circ$	$\pm 5^\circ$	$\pm 2^\circ$	5.9%	156	24.5%	67.4%
M06R04	86.1	2.2%	$\pm 126^\circ$	$\pm 42^\circ$	$\pm 17^\circ$	$\pm 9^\circ$	$\pm 4^\circ$	$\pm 2^\circ$	5.9%	68	9.9%	82.0%
M08R04	86.1	2.2%	$\pm 126^\circ$	$\pm 42^\circ$	$\pm 17^\circ$	$\pm 9^\circ$	$\pm 4^\circ$	$\pm 2^\circ$	5.9%	60	5.5%	86.4%
M16R04	86.1	2.2%	$\pm 126^\circ$	$\pm 42^\circ$	$\pm 17^\circ$	$\pm 9^\circ$	$\pm 4^\circ$	$\pm 2^\circ$	5.9%	4	0.09%	91.8%

The attenuation performance of the microphone array with different hardware setups is provided in Table 2-1. The table demonstrates that the low-frequency area and the main lobe area are only related to the radius of the microphone array and unrelated to the microphone number, just like a continuous microphone array. A bigger radius of the microphone array could generate the lower max frequency of the low-frequency area and narrower main lobe area. More microphone numbers could make the performance of the microphone array more similar to a continuous microphone array.

The less and smaller grating lobe area would be generated and the frequency of the lowest frequency of the grating lobe would be higher.

The phenomenon of the beamforming pattern coincides with Equation 2-15. The low-frequency area and the main lobe area correspond to the zero-order Bessel function in Equation 2-15, just be related to the frequency of the received signal and the radius of the microphone array. The grating lobe area corresponds to the high-order Bessel function in the equation. Less microphone number would lead to the considerable value of the high-order Bessel function in Equation 2-15, which means the ability of sidelobes suppression is impaired. When the microphone number increases, the frequency of the lowest frequency of the grating lobe would be higher and the grating lobe would be faded.

2.2 Correlation coefficient

2.2.1 Pearson correlation coefficient

The correlation coefficient is a statistical measure of the strength of the relationship between the relative movements of two variables. The values range between -1.0 and 1.0. A calculated number is greater than 1.0 or less than -1.0 means that there is an error in the correlation measurement. A correlation of -1.0 shows a perfect negative correlation, while a correlation of 1.0 shows a perfect positive correlation. A correlation of 0.0 shows no linear relationship between the movement of the two variables [55].

There are several types of correlation coefficients, but the one that is most common is the Pearson correlation coefficient. It measures the strength and direction of the linear relationship between two variables. It ignores many other types of relationships or correlations between two variables and cannot differentiate between dependent and

independent variables [56].

The Pearson correlation coefficient (PCC), the bivariate correlation [57], or colloquially simply as the correlation coefficient [58] is the ratio between the covariance of two variables and the product of their standard deviations. Thus, it is essentially a normalized measurement of the covariance, such that the result always has a value between -1 and 1.

As shown in Figure 2-7 [59], Values range between -1 (strong negative relationship) and +1 (strong positive relationship). Values at or close to zero imply a weak or no linear relationship. A value of exactly 1.0 means there is a perfect positive relationship between the two variables. For a positive increase in one variable, there is also a positive increase in the second variable. A value of -1.0 means there is a perfect negative relationship between the two variables. This shows that the variables move in opposite directions—for a positive increase in one variable, there is a decrease in the second variable. If the correlation between two variables is or close to 0, there is no or a weak linear relationship between them.

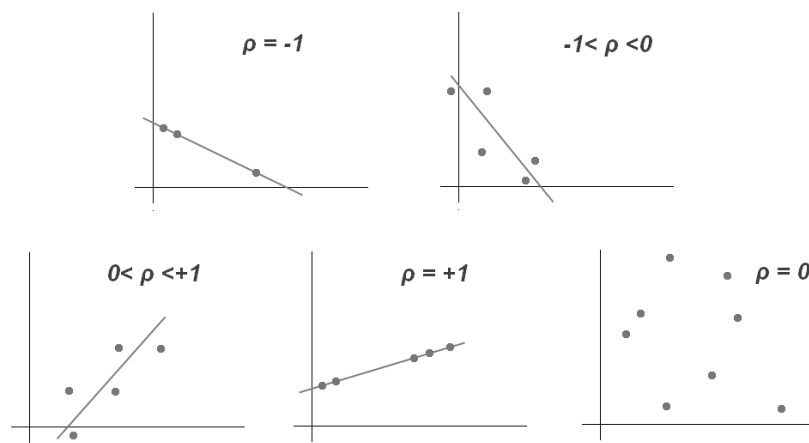


Figure 2-7. Correlation coefficients indicate relationships between variables.

The strength of the relationship varies in degree based on the value of the correlation coefficient. For example, a value of 0.2 shows there is a positive correlation

between two variables, but it is weak and likely unimportant. Analysts in some fields of study do not consider correlations important until the value surpasses at least 0.8. However, a correlation coefficient with an absolute value of 0.9 or greater would represent a very strong relationship [57].

To calculate the Pearson product-moment correlation, one must first determine the covariance of the two variables in question. Next, one must calculate each variable's standard deviation. The correlation coefficient is determined by dividing the covariance by the product of the two variables' standard deviations.

$$\rho_{xy} = \frac{Cov(x, y)}{\sigma_x \sigma_y} = \frac{Cov(x, y)}{\sqrt{Var(x)}\sqrt{Var(y)}} = \frac{E(XY) - E(X)E(Y)}{\sqrt{E(X^2) - E^2(X)}\sqrt{E(Y^2) - E^2(Y)}}, \quad (2-19)$$

Where,

ρ_{xy} = Pearson product-moment correlation coefficient,

$Cov(x,y)$ = covariance of variables x and y ,

σ_x = standard deviation of x ,

σ_y = standard deviation of y ,

$Var(x)$ = variance of x ,

$Var(y)$ = variance of y ,

E represents the mathematical expectation of the data set. X, Y represents the two data sets and x, y represents the sample in the corresponding data set.

Standard deviation is a measure of the dispersion of data from its average. Covariance is a measure of how two variables change together, but its magnitude is unbounded, so it is not easy to interpret. By dividing covariance by the product of the two standard deviations, one can calculate the normalized version of the statistic. This is the correlation coefficient.

When the Pearson correlation coefficient is applied to a sample, it may be referred to as the sample Pearson correlation coefficient [60]. The formula could be obtained as follows:

$$\tau_{xy} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (2-20)$$

Where,

τ_{xy} = Pearson product-moment correlation coefficient,

\bar{x} = the mean of x ,

\bar{y} = the mean of y .

2.2.2 Coefficient of rank correlation

The coefficient of rank correlation, also known as rank correlation coefficient, is a statistic obtained by arranging the sample values of two elements in order of data size and replacing the actual data with the sample values of each element. It is a statistical analysis index reflecting the degree of rank correlation. The commonly used rank correlation analysis methods include Spearman correlation coefficient and Kendall rank correlation coefficient [61].

If the rank correlation coefficient is positive, y increases as x increases; If the rank correlation coefficient is negative, y decreases as x increases. If the rank correlation coefficient is 0, it means that y has no tendency to increase or decrease as x increases. As x and y get closer and closer to strictly monotonic functions, the rank correlation coefficient becomes numerically larger. When the rank correlation coefficient is 1 or -1, it indicates that the strict monotonicity between and increases or decreases [62].

In practical application, sometimes the original data obtained do not have specific data performance and can only be used to describe a particular phenomenon by rank,

and to analyze the correlation between phenomena, only rank correlation coefficient can be used.

(1) Spearman correlation coefficient

Spearman's rank correlation coefficient, named after Charles Spearman is a nonparametric measure of rank correlation (statistical dependence between the rankings of two variables). It assesses how well the relationship between two variables can be described using a monotonic function [63].

Spearman's coefficient is appropriate for both continuous and discrete ordinal variables [64]. The Spearman correlation between two variables is equal to the Pearson correlation between the rank values of those two variables; while Pearson's correlation assesses linear relationships, Spearman's correlation assesses monotonic relationships (whether linear or not). If there are no repeated data values, a perfect Spearman correlation of +1 or -1 occurs when each of the variables is a perfect monotone function of the other.

The Spearman correlation coefficient is defined as the Pearson correlation coefficient between the rank variables [65] as follows:

$$r_s = \rho_{R(x)R(y)} = \frac{Cov(R(x), R(y))}{\sigma_{R(x)}\sigma_{R(y)}}, \quad (2-21)$$

Where,

$\rho_{R(x)R(y)}$ = the usual Pearson correlation coefficient, but applied to the rank variables,

$Cov(R(x), R(y))$ = the covariance of the rank variables,

$\sigma_{R(x)}$ = standard deviations of the rank variables $R(x)$,

$\sigma_{R(y)}$ = standard deviation of the rank variables $R(y)$.

Only if all n ranks are distinct integers, it can be computed using the popular

formula:

$$r_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)}, \quad (2-22)$$

$$d_i = R(X_i) - R(Y_i), \quad (2-23)$$

Where d_i is the difference between the two ranks of each observation, n is the number of observations.

(2) Kendall rank correlation coefficient

The Kendall rank correlation coefficient, commonly referred to as Kendall's τ coefficient, is a statistic used to measure the ordinal association between two measured quantities [66]. It is a measure of rank correlation: the similarity of the orderings of the data when ranked by each of the quantities [67,68]. Kendall's rank correlation provides a distribution-free test of independence and a measure of the strength of dependence between two variables. Spearman's rank correlation is satisfactory for testing a null hypothesis of independence between two variables, but it is not easy to interpret when the null hypothesis is rejected. Kendall's rank correlation improves upon this by reflecting the strength of the dependence between the variables being compared [69].

Kendall's rank correlation coefficient and Spearman's rank correlation coefficient assess statistical associations based on the ranks of the data. Kendall rank correlation (non-parametric) is an alternative to Pearson's correlation (parametric) when the data you're working with has failed one or more assumptions of the test. This is also the best alternative to Spearman correlation (non-parametric) when your sample size is small and has many tied ranks.

Kendall rank correlation is used to test the similarities in the ordering of data when it is ranked by quantities. Other types of correlation coefficients use the observations as

the basis of the correlation, Kendall's correlation coefficient uses pairs of observations and determines the strength of association based on the pattern of concordance and discordance between the pairs [70]. Intuitively, the Kendall correlation between two variables will be high when observations have a similar (or identical for a correlation of 1) rank between the two variables and low when observations have a dissimilar (or fully different for a correlation of -1) rank between the two variables.

Thus, the Kendall's rank correlation coefficient is defined as:

$$\tau = \frac{(n_c - n_d)}{n} = \frac{2*(n_c - n_d)}{(n_c + n_d)(n_c + n_d - 1)}, \quad (2-24)$$

Where,

n_c represents the number of the concordant pairs: ordered in the same way (consistency). A pair of observations is considered concordant if $(x_2 - x_1)$ and $(y_2 - y_1)$ have the same sign.

n_d represents the number of the discordant pairs: ordered differently (inconsistency). A pair of observations is considered discordant if $(x_2 - x_1)$ and $(y_2 - y_1)$ have opposite signs.

n represents the total number of pair combinations.

The explicit expression for Kendall's rank correlation coefficient is:

$$\tau = \frac{2}{n(n-1)} \sum_{i < j} \text{sgn}(x_i - x_j) \text{sgn}(y_i - y_j). \quad (2-25)$$

2.3 LMS algorithm

The LMS algorithm is an adaptive filter algorithm that is used to establish the desired filter by adapting the filter coefficients to reduce the least mean square of the error signal (the difference between the desired signal and the actual signal) as shown in Figure 2-8 [26][27]. It is a stochastic gradient descent method that the filter

coefficients can be adapted based on the error value at the current time. It was invented in 1960 by Stanford University professor Bernard Widrow and his Ph.D. student, Ted Hoff. Then, Morgan proposed the FxLMS algorithm, where the LMS algorithm was modified to consider the impact of the secondary path [71]. A block diagram of the FxLMS algorithm is shown in Figure 2-9.

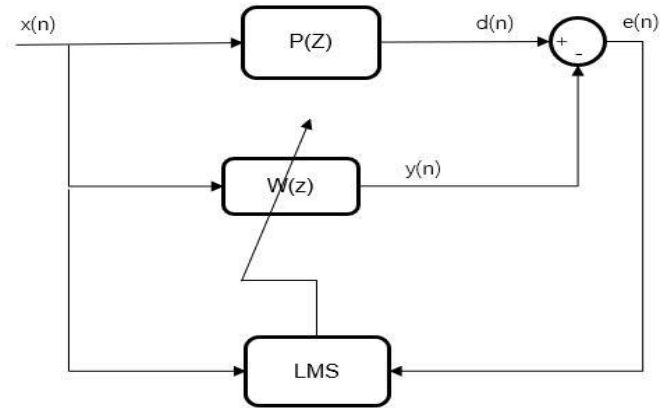


Figure 2-8. Block diagram of the LMS algorithm.

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

$$e(n) = d(n) - y'(n) = d(n) - s(n) * y(n), \quad (2-26)$$

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 $S(z)$, and $*$ denotes linear convolution. $y(n)$ is the anti-phase signal calculated by the algorithm.

$$y(n) = w(n) * x(n), \quad (2-27)$$

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

$$x'(n) = \hat{s}(n) * x(n), \quad (2-28)$$

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:

$$w(n+1) = w(n) + \mu x'(n)e(n). \quad (2-29)$$

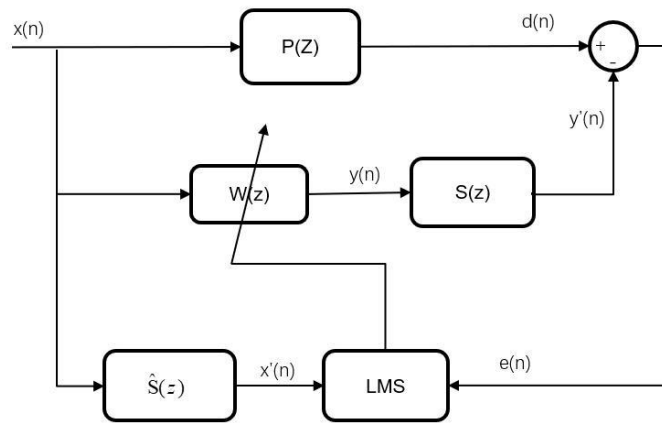


Figure 2-9. Block diagram of the FxLMS algorithm.

2.4 Conclusion

This chapter introduced the related foundation knowledge about the microphone array.

The formula of the beamforming pattern of the UCA was derived first. Then the influence factors of the beamforming pattern were displayed. The formula of the beamforming pattern could be formed by the Bessel functions. The simulation with different microphone setups demonstrated the validity of the derived formula of the beamforming pattern. The radius of the microphone array would affect the spatial resolution of the main lobe of the microphone array. It was because the zero-order Bessel function that determined the width of the main lobe of the beamforming pattern was related to the radius of the microphone array. The larger diameter would bring the

narrower main lobe. The high-order Bessel functions determined the attenuation of the sidelobe, the more microphones in the microphone array would make the beamforming pattern more like the continuous microphone array that avoiding the grating lobe being generated.

Then the correlation coefficient was introduced. The Pearson correlation coefficient shows the linear relationship of two variables. And the rank correlation coefficient, like the Spearman correlation coefficient and Kendall rank correlation coefficient, could show the variation trend of two variables concordant or discordant. The formulas of the correlation coefficient were given.

Last, the LMS algorithm was described. The LMS algorithm is a common adaptive filter algorithm to estimate the transfer function between the reference noise signal and the received noise signal. The LMS algorithm could be modified to the FxLMS algorithm by adding the estimated transfer function of the secondary path to the coefficient update process of the LMS algorithm to improve the converge performance.

Chapter 3 Modified GSC method

3.1 Conventional GSC method

The GSC algorithm can be divided into two transmission paths: primary path and auxiliary path, as shown in Figure 3-1. The primary path generates a denoising signal with residual noise by the fixed beamforming method. The auxiliary path estimates the noise component in the output signal of the primary path. ω_m is the weight vector to make the microphone array signal $x(n)$ aligned in the looking direction to suppress noise from other directions. $y_m(n)$ is the sum of the aligned signals as the beamforming output signal of the primary path. The auxiliary path estimates the noise component in the beamforming output signal of the primary path $y_m(n)$ as the canceling signal. The canceling signal $y_b(n)$ will be removed from the beamforming output signal. Hence, the enhanced signal $y_{out}(n)$ consists of the desired signal and residual noise signal is obtained.

The blocking matrix (BM) is adopted to generate the reference noise signal excluded the desired signal from the noisy signal in the auxiliary path for estimating the noise component of the beamforming output signal. Various methods can be used to design the blocking matrix for different purposes. The typical blocking matrix is as follows:

$$BM = \begin{bmatrix} 1 & -1 & 0 & 0 & \dots & 0 & 0 \\ 0 & 1 & -1 & 0 & \dots & 0 & 0 \\ \dots & & & & & & \\ 0 & 0 & 0 & 0 & \dots & 1 & -1 \end{bmatrix}_{(M-1)*M}$$

where M is the number of microphones in the microphone array. This matrix can block

the desired signal completely in theory and introduce a low calculation complexity. The output $z_b(n)$ of the blocking matrix is offered to the adaptive filter as the reference noise signal to estimate the noise component of the beamforming output signal. The LMS algorithm is commonly used as an adaptive filter to adjust the filter coefficients ω_f . Thus, the error signal of the LMS algorithm $y_{out}(n)$, which is also the output signal of the GSC method, is obtained by removing the estimated noise signal from the beamforming output signal. The filter coefficient ω_f is then updated based on the reference noise signal and error signal.

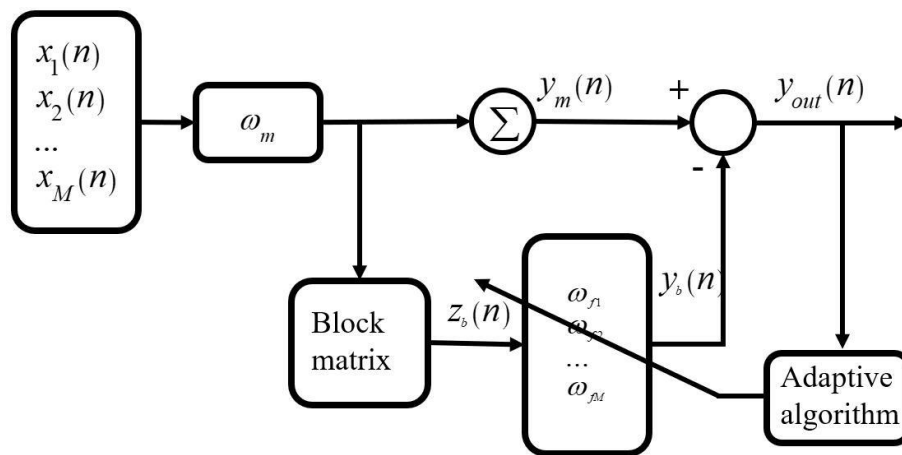


Figure 3-1. Block diagram of the conventional GSC method.

3.2 TF-GSC method

TF-GSC method is based on the conventional GSC method [54]. The block method of the conventional GSC method is modified with the transfer function ratios between different microphones in the array to adapt to the reverberation condition.

The block diagram of the TF-GSC method is shown in Figure 3-2. $Z_m(t, e^{j\omega})$ is the Fourier transform of the received time-domain signal. W_0^H is the weight of the fixed beamformer and H^H is the blocking matrix to the received signal.

The TF-GSC method would perform the following steps:

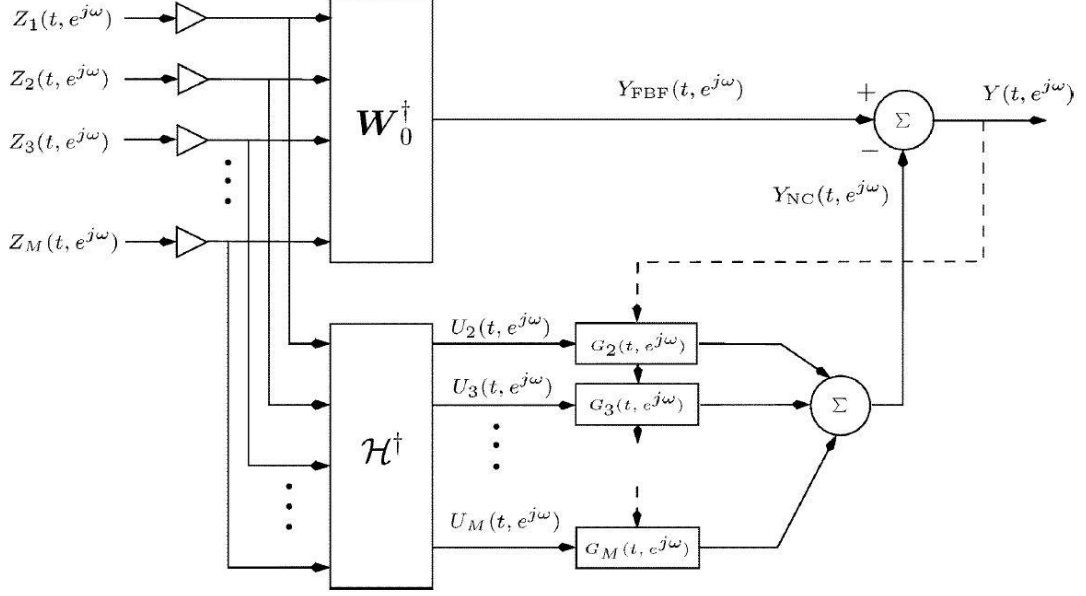


Figure 3-2. Block diagram of the TF-GSC method.

(1) The transfer function ratios $H(e^{j\omega})$ are calculated as Equation 3-1. $A_m(e^{j\omega})$ represents the acoustic transfer function from the sound source to the m th microphone in the microphone array.

$$H(e^{j\omega}) = \frac{A(e^{j\omega})}{A_1(e^{j\omega})} = \left[1, \frac{A_2(e^{j\omega})}{A_1(e^{j\omega})}, \frac{A_3(e^{j\omega})}{A_1(e^{j\omega})}, \dots, \frac{A_M(e^{j\omega})}{A_1(e^{j\omega})} \right]. \quad (3-1)$$

(2) The blocking matrix is constructed as the conjugation transpose matrix of the transfer function ratios $H(e^{j\omega})$. Thus, the blocking matrix satisfies Equation 3-2.

$$H^H(e^{j\omega})A(e^{j\omega}) = 0. \quad (3-2)$$

(3) The weight of the fixed beamformer $W_0(t, e^{j\omega})$ would be as following Equation 3-3. Therefore the output of the fixed beamformer $Y_{FBF}(t, e^{j\omega})$ is calculated as Equation 3-4.

$$W_0(t, e^{j\omega}) = \frac{H(e^{j\omega})}{\|H(e^{j\omega})\|^2} F(e^{j\omega}). \quad (3-3)$$

$$Y_{FBF}(t, e^{j\omega}) = W_0^H(e^{j\omega})Z(t, e^{j\omega}). \quad (3-4)$$

(4) The noise reference signals $U(t, e^{j\omega})$ could be calculated as follows:

$$U(t, e^{j\omega}) = H^H(e^{j\omega})Z(t, e^{j\omega}) = H^T(e^{j\omega})N(t, e^{j\omega}). \quad (3-5)$$

$$U(t, e^{j\omega}) = Z_m(t, e^{j\omega}) - \frac{A_m(e^{j\omega})}{A_1(e^{j\omega})} Z_1(t, e^{j\omega}); m = 2, \dots, M. \quad (3-6)$$

(5) Output signal $Y(t, e^{j\omega})$ of the TF-GSC method would be:

$$Y(t, e^{j\omega}) = Y_{\text{FBF}}(t, e^{j\omega}) - G^T(t, e^{j\omega}) U(t, e^{j\omega}). \quad (3-7)$$

(6) Final, the filter weights $G_m(t, e^{j\omega})$ would be updated as follows:

$$G_m(t+1, e^{j\omega}) = G_m(t, e^{j\omega}) + \mu \frac{U_m(t, e^{j\omega}) Y^*(t, e^{j\omega})}{P_{\text{est}}(t, e^{j\omega})}. \quad (3-8)$$

$$P_{\text{est}}(t, e^{j\omega}) = \rho P_{\text{est}}(t-1, e^{j\omega}) + (1-\rho) \sum_m |Z_m(t, e^{j\omega})|^2; m = 1, \dots, M-1. \quad (3-9)$$

3.3 GSC method with cross-correlation coefficient

In practice, the error signal of the GSC method consists of a residual noise signal and an enhanced desired signal. In speech application, the desired signal is often a nonstationary speech signal. When the amplitude of the speech signal increases suddenly, the amplitude of the error signal would also be increased synchronously even though the amplitude of the residual noise signal may remain stable. Hence, over-subtraction may be caused in the subsequent frames. Considering the speech signal usually is uncorrelated with noise signal in practical situations, the cross-correlation coefficient could be introduced in the adaptive filter weight update process to alleviate this problem. The flowchart of the GSC with cross-correlation coefficient method is displayed in Figure. 3-3. The solid line represents the flowchart of the conventional GSC method. The dotted line denotes the module of the cross-correlation coefficient method.

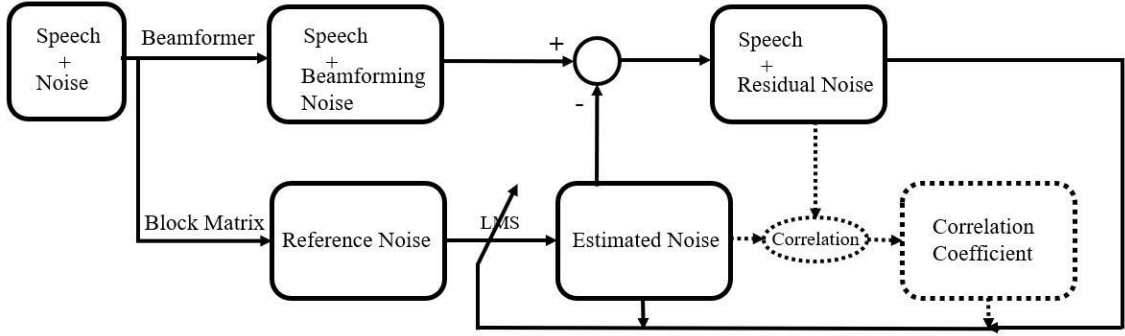


Figure 3-3. Flowchart of the GSC with cross-correlation coefficient method.

The GSC method with the cross-correlation coefficient (GSC-CC), in which the cross-correlation coefficient is introduced to the weight updating path, is shown in Figure 3-4.

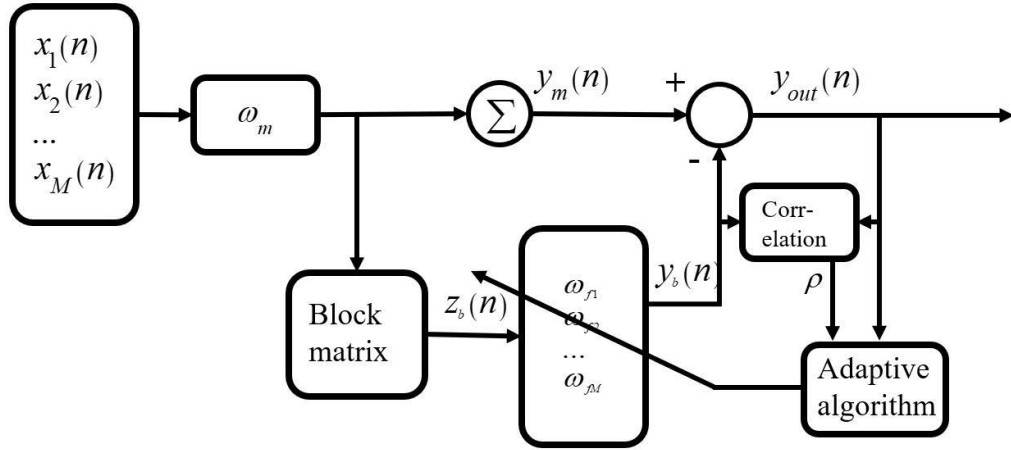


Figure 3-4. Block diagram of the GSC method with the cross-correlation coefficient.

The cross-correlation coefficient between the estimated noise $y_b(n)$ and the error signal $y_{out}(n)$ is used to control the step size of the adaptive filter weight update process, as expressed in the following equations:

$$w_f(n+1) = w_f(n) + |\rho(k-1)| \mu x(n) e(n). \quad (3-10)$$

$$\begin{aligned} \rho(k) &= E(y_b(k) y_{out}(k)) = E(y_b(k) (s_{speech}(k) + n_{residual}(k))) \\ &= E(y_b(k) s_{speech}(k)) + E(y_b(k) n_{residual}(k)) \end{aligned} \quad (3-11)$$

$$\text{If } y_b(k) \approx n_{beamform}(k), \quad (3-12)$$

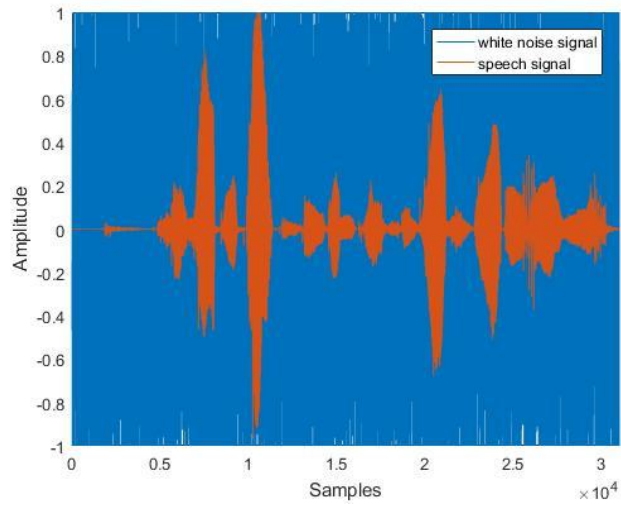
$$\text{Thus, } E(y_b(k) s_{speech}(k)) \approx 0, \quad (3-13)$$

$$\text{Finally, } \rho(k) \approx E(y_b(k) n_{residual}(k)), \quad (3-14)$$

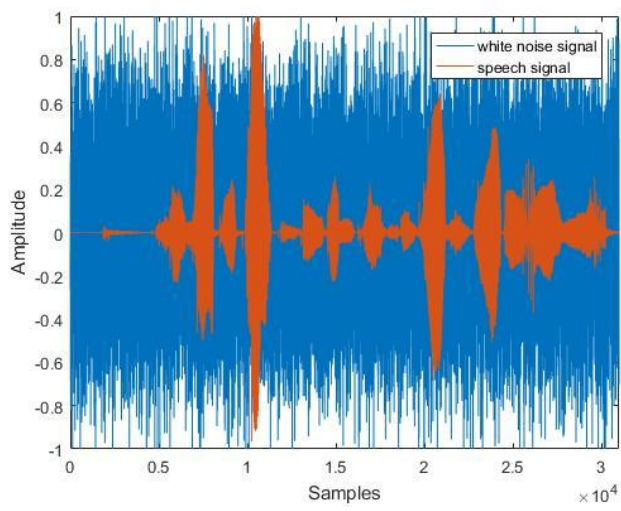
where $\rho(k)$ is the cross-correlation coefficient of the k th frame between the estimated noise $y_b(k)$ and the error signal $y_{out}(k)$. $s_{speech}(k)$ represents the clean desired speech signal and $n_{residual}(k)$ is the residual noise component of the output signal. Because of the speech signal is commonly assumed to be uncorrelated with the noise signal, the cross-correlation coefficient between the estimated noise and the speech signal is close to zero and can be neglected. Therefore, $\rho(k)$ is equal to the cross-correlation coefficient between the estimated noise signal $y_b(k)$ and the residual noise signal $n_{residual}(k)$ as Equation 3-14. This means that if the $\rho(k)$ is small, the residual noise is nearly uncorrelated with the estimated noise, so the weight coefficient will be changed only slightly. Otherwise, the weight coefficient will be changed considerably to reduce the residual noise rapidly.

Figures 3-5 and 3-6 illustrate the influence of the varying speech signal on the cross-correlation coefficient between the estimated noise signal and the output signal with the conventional GSC method and the GSC-CC method on different SNR conditions. The input original signal is a speech signal extracted from the TIMIT database [72] and the noise signal is the white noise obtained from the NOISEX-92 database [73].

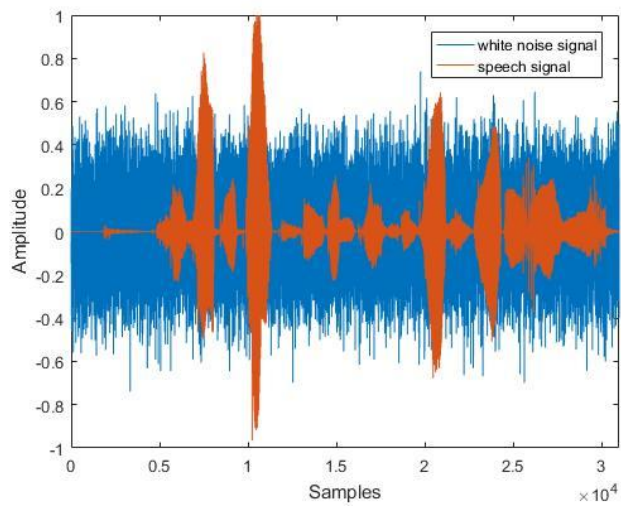
Figure 3-5 shows the original speech and noise signal, in which the signal-to-noise ratio (SNR) is equal to -15dB, -10dB, -5dB, 0dB, 5dB, 10dB (the ratio of the energy of the whole speech signal and the energy of the noise signal during the same time), respectively. Detailed information about the experimental setup is depicted in Chapter 4. Figure 3-6 presents the cross-correlation coefficient between the estimated noise signal and the output signal with the conventional GSC method and the GSC-CC method on different SNR conditions.



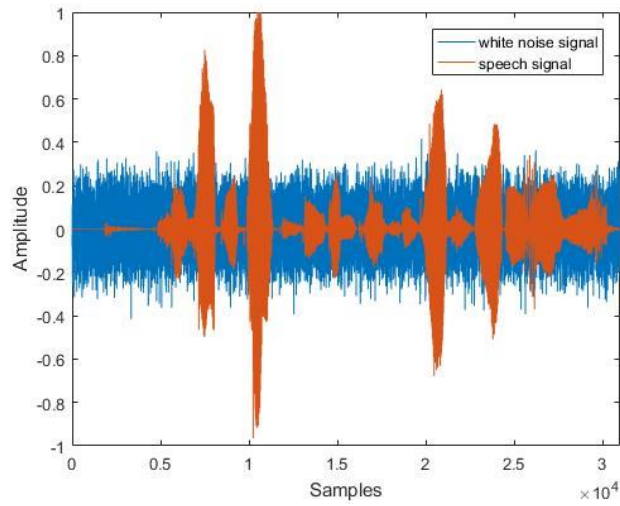
(a) SNR = -15dB



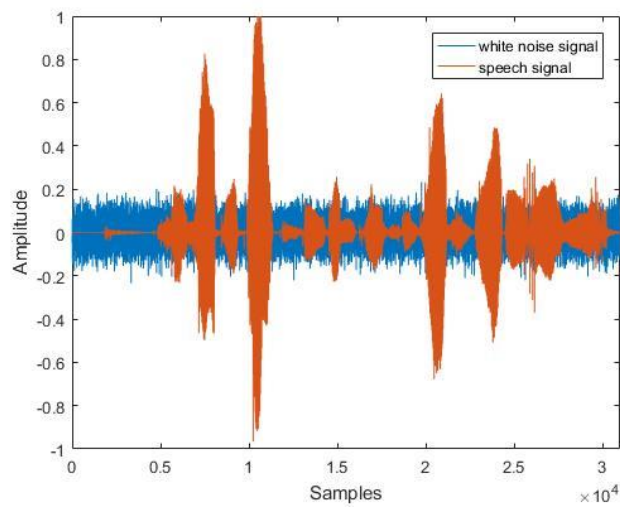
(b) SNR = -10dB



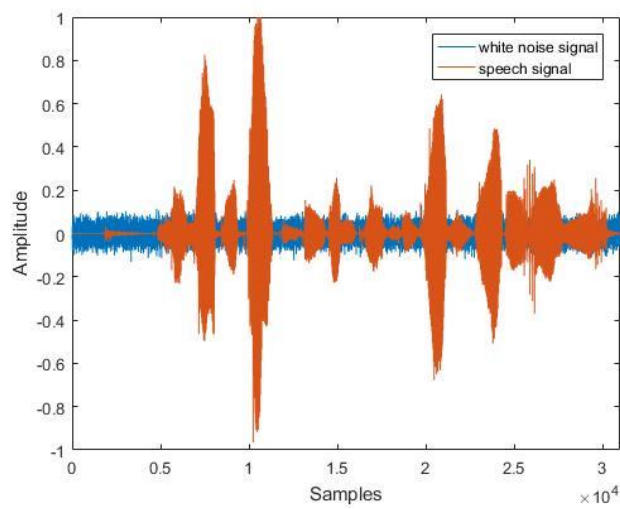
(c) SNR = -5dB



(d) SNR = 0dB

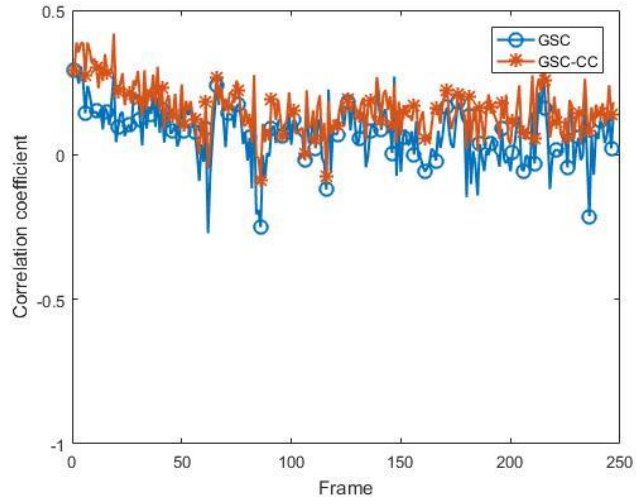


(e) SNR = +5dB

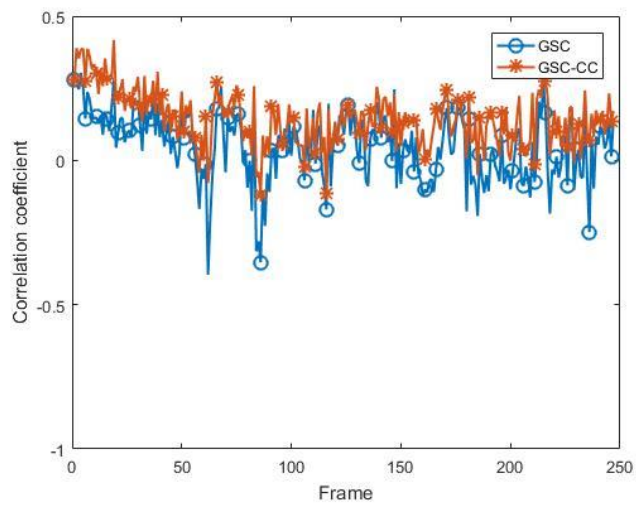


(f) SNR = +10dB

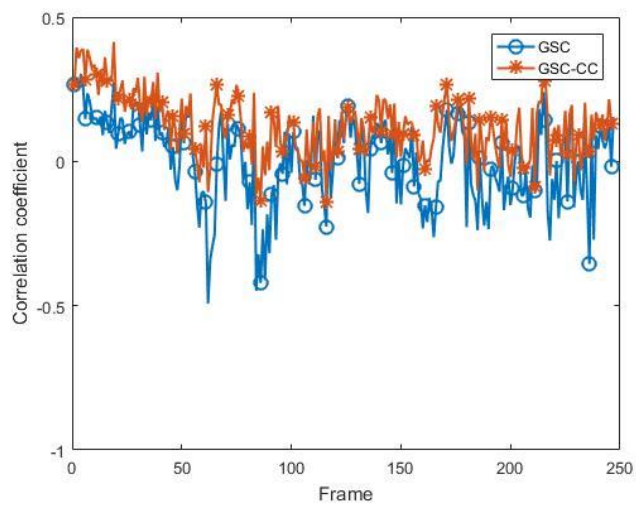
Figure 3-5. Clean speech signal and white noise signal on different SNR conditions.



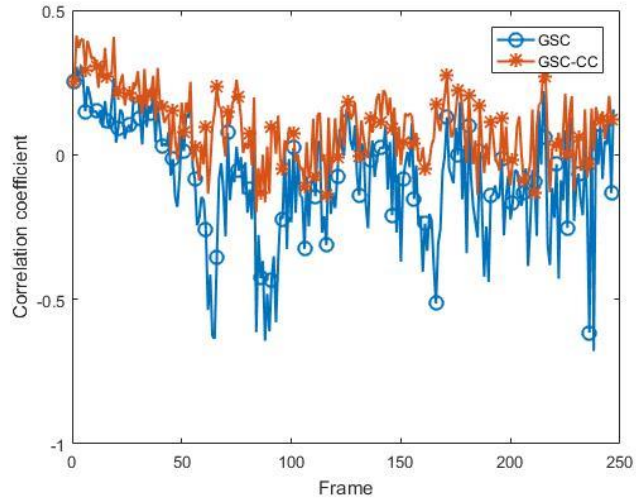
(a) SNR = -15dB



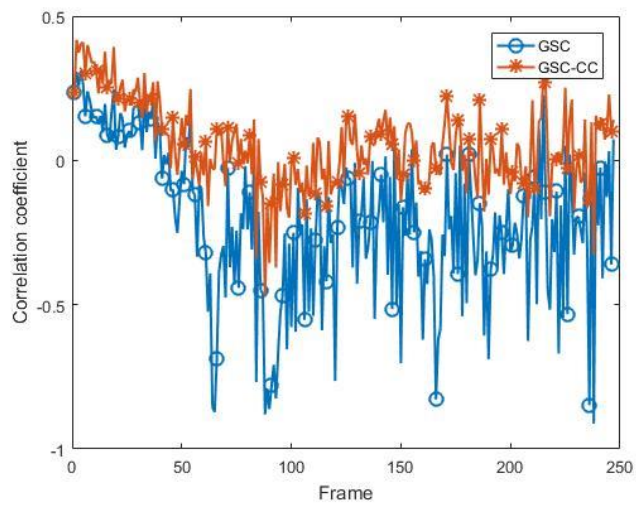
(b) SNR = -10dB



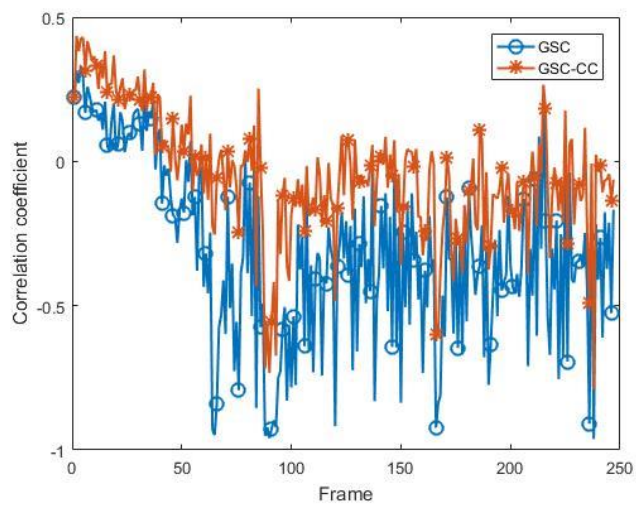
(c) SNR = -5dB



(d) SNR = 0dB



(e) SNR = +5dB



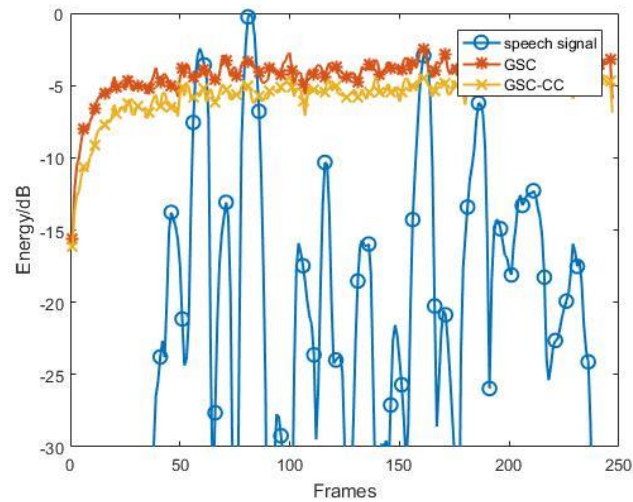
(f) SNR = +10dB

Figure 3-6. Cross-correlation coefficient of different methods on different SNR conditions.

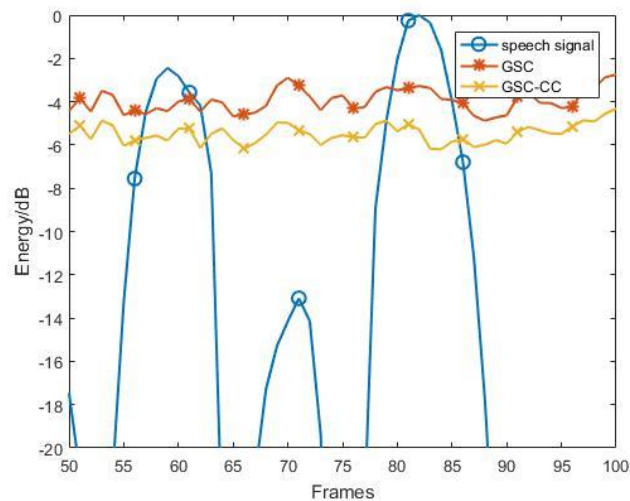
The figures illustrate that the negative peak of the cross-correlation coefficient by

the conventional GSC method is generated when the amplitude of the speech signal is increased suddenly, which implies the phenomenon of over-subtraction. With the increase of the SNR, the speech signal emerges from the white noise and more peaks of the cross-correlation coefficient are generated, some time is closed to -1, which implies the phenomenon of over subtraction becomes more serious.

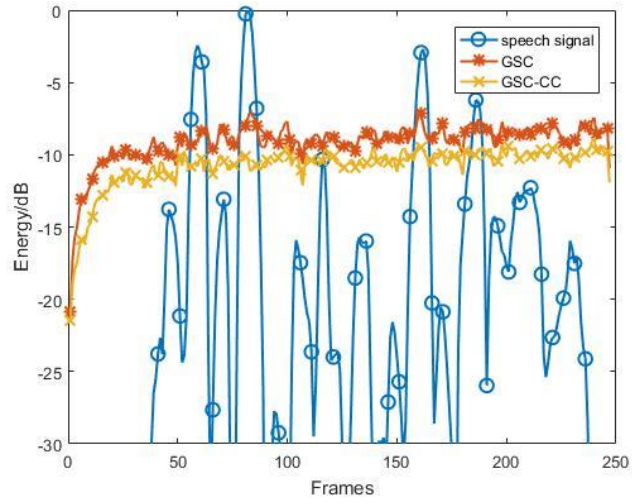
Figure 3-7 displays the influence of the varying speech signal on the estimated noise signal. The figures present the energy of the estimated noise frame with different methods on different SNR conditions.



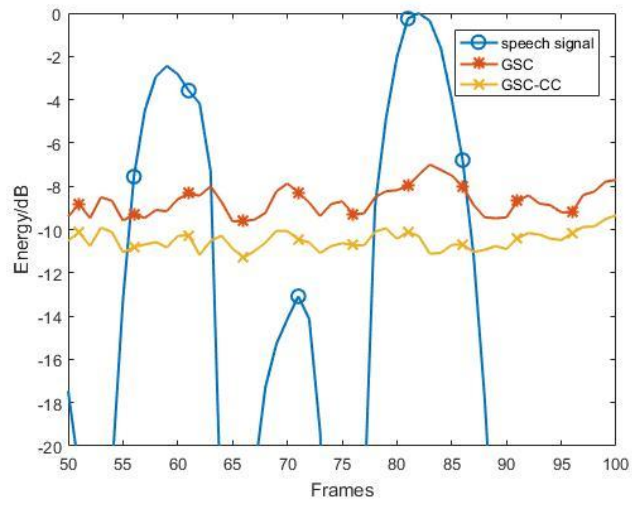
(a1) SNR = -15dB



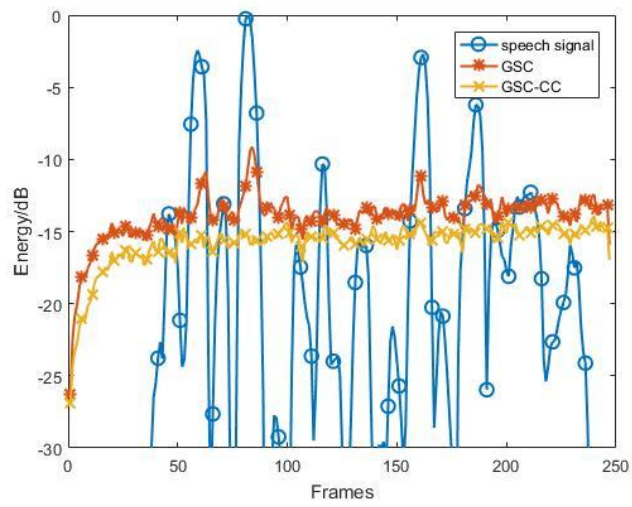
(b1) SNR = -15dB



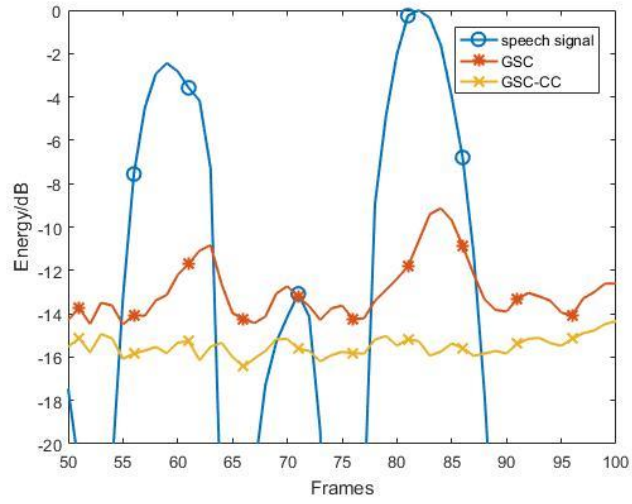
(a2) SNR = -10dB



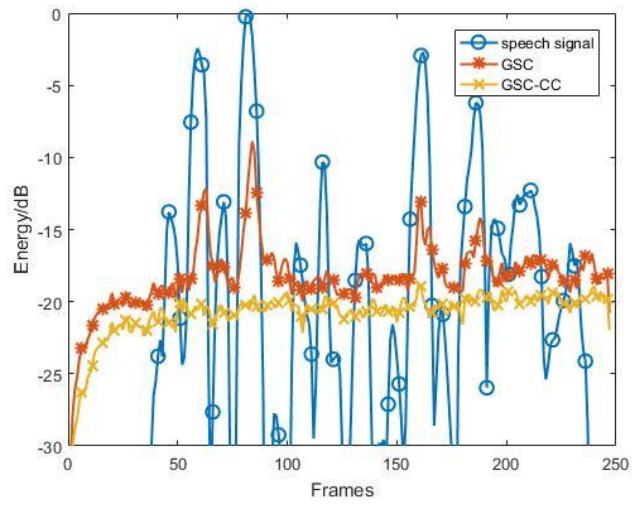
(b2) SNR = -10dB



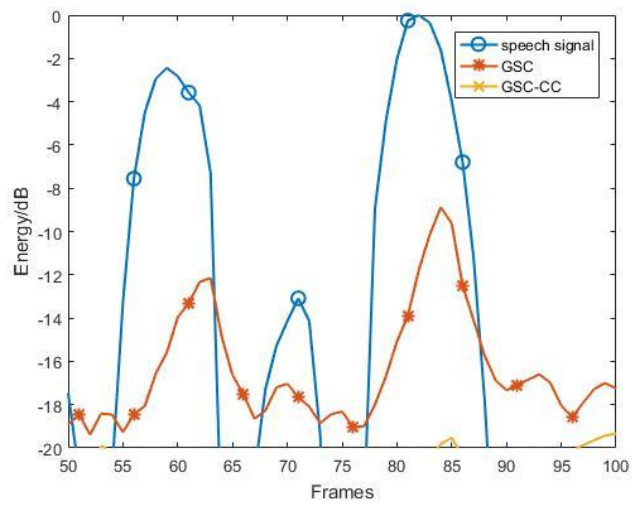
(a3) SNR = -5dB



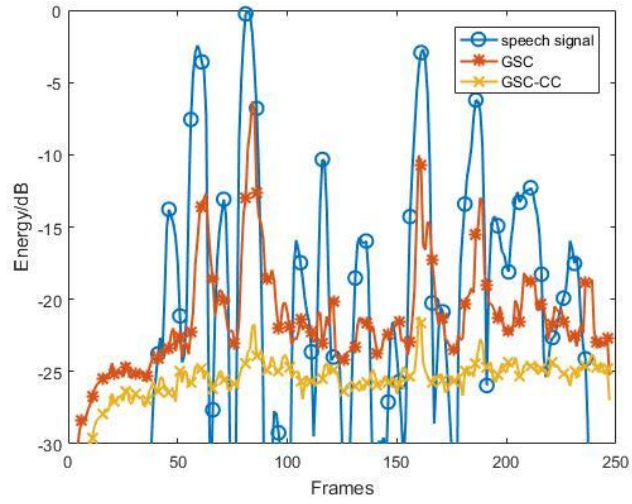
(b3) SNR = -5dB



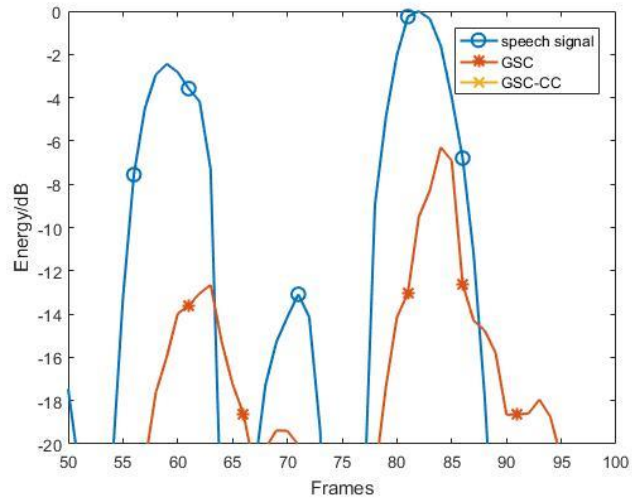
(a4) SNR = 0dB



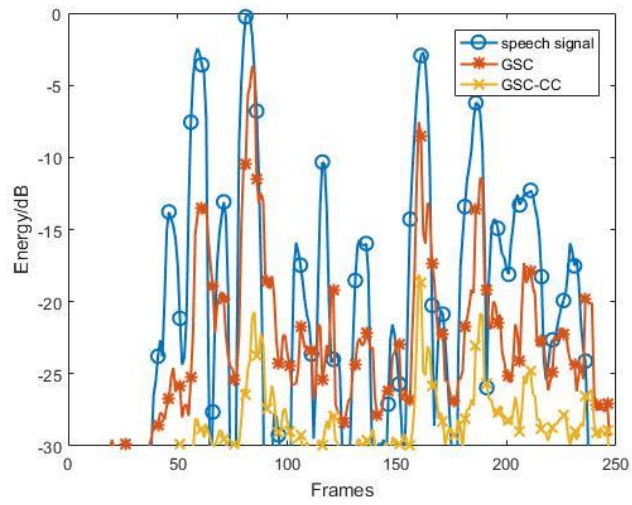
(b4) SNR = 0dB



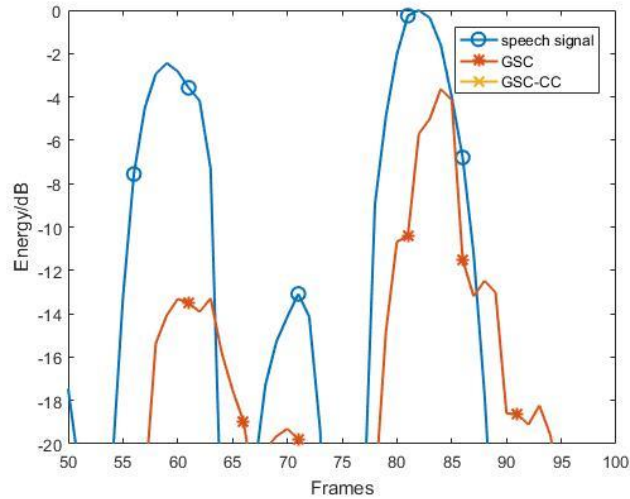
(a5) SNR = +5dB



(b5) SNR = +5dB



(a6) SNR = +10dB



(b6) SNR = +10dB

Figure 3-7 Comparison of the energy of the estimated noise frame with different methods on different SNR conditions: (a1~a6) all frames; (b1~b6) 50th to 100th frame.

The part (b1~b6) of the figures displays the partially enlarged view of the corresponding part (a1~a6) of the figures from the 50th to 100th frame. When the SNR is low, as SNR is -15 dB, the speech signal is almost submerged by the background noise. The estimated noise by the adaptive algorithm is stable and barely affected by the variation of the speech signal. With the increase of the SNR, the speech signal becomes emerging from the background noise and the estimated noise by the adaptive algorithm is increased rapidly though the background noise is stable.

The energy of the estimated noise frame with the conventional GSC method would increase when the energy of the speech frame is much more than the actual noise frame. As shown in Figure 3-7 (a4, b4, SNR=0dB), the frame energy of the estimated noise is raised when the frame energy of the speech signal is noticeable than other frames. When the SNR of the received signal is as high as +10dB, the frame energy of the estimated noise is changed rapidly with the frame energy of the speech signal even though the actual noise is still stable.

The figures illustrate the GSC-CC method could flatten the energy curve of the

estimated noise frame, which is consistent with the trend of the energy curve of the original white noise signal on different SNR conditions. This means that the distortion of the enhanced speech signal is reduced since the desired signal is damaged less when the amplitude of the speech signal is changed.

The GSC-CC method could decrease the overestimation of the noise efficiently when the desired speech signal amplitude is increased suddenly (Figure 3-7). On the other hand, the figures also show that the energy curve of the estimated noise frames with the GSC-CC method is lower than the conventional GSC method. This is because the correlation coefficient is less than or equal to one, the step size of the adaptive filter of the GSC-CC method is always smaller than the GSC method so that the convergence speech of the LMS algorithm becomes slow. Consequently, the energy of the residual noise signal with the GSC-CC method is higher than the conventional GSC method.

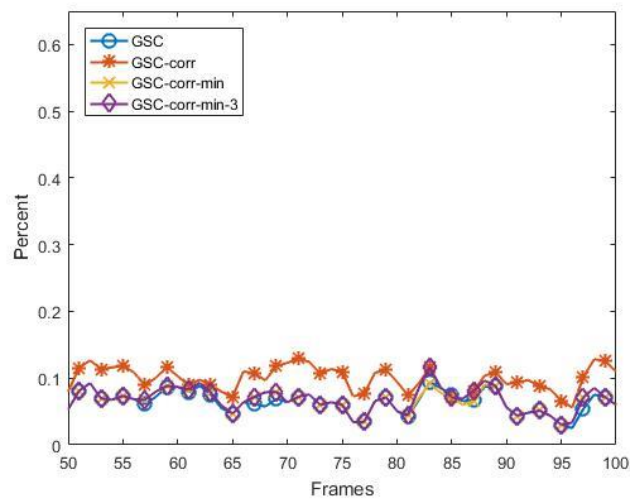
The GSC method with minimum cross-correlation coefficient (GSC-MCC) method is proposed by combining the conventional GSC method and the GSC-CC method to utilize the advantages of the two methods. The conventional GSC method and the GSC-CC method are run synchronously. One frame of the output signal frames of these two methods, which has the smaller cross-correlation coefficient (between the estimated noise signal and the corresponding output signal), is adopted to synthesize the final output signal.

Considering the overlap of speech frames and the nonstationary nature of a speech signal, the principle of the GSC-MCC method is modified as Equation 3-15 to choose the frame to synthesize the final output signal.

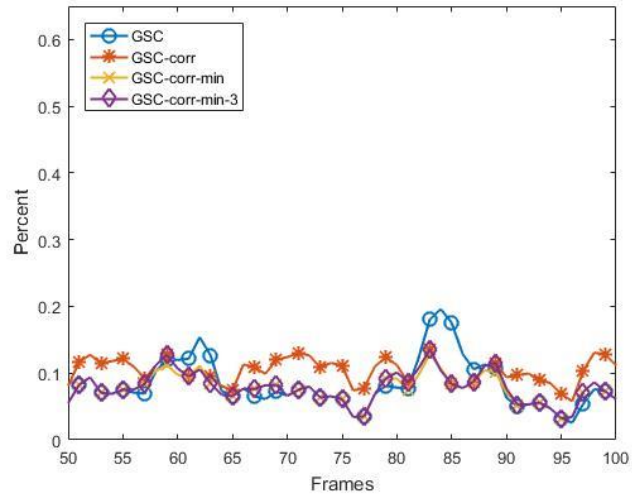
$$y_{out-final}(k) = \begin{cases} y_{out-corr}(k), & \sum_{i=k-1}^{k+1} \text{sign}(|\rho_{conv}(i)| - |\rho_{corr}(i)|) \geq 1, \\ y_{out-conv}(k), & \text{otherwise} \end{cases}, \quad (3-15)$$

where $y_{out-corr}(k)$ and $y_{out-conv}(k)$ represent the output signal frame of the GSC-CC method and the conventional GSC method, respectively.

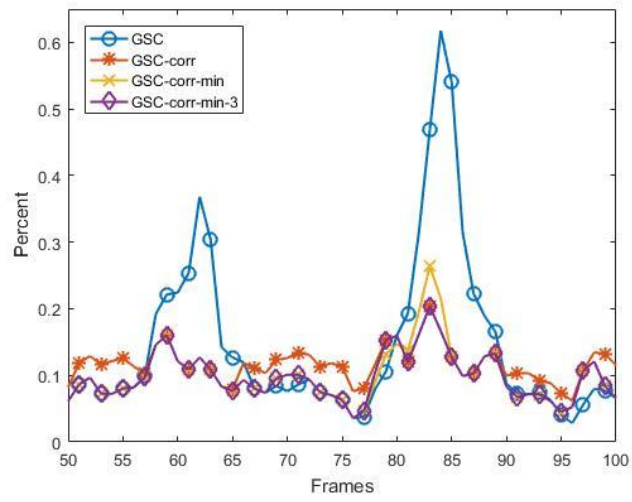
Figure 3-8 displays the distortion-noise ratio of frames with different methods from the 50th to 100th frame under different SNR conditions (-15dB, -10dB, -5dB, 0dB, +5dB, +10dB). The smaller distortion-noise ratio means lighter distortion. The distortion-noise ratio will be introduced in Chapter 4. The figures illustrate that the signal distortion processed by the GSC method is raised with the SNR increasing. The GSC-CC method could reduce the peak value of the distortion efficiently. However, the residual noise is increased since the correlation coefficient impaired the efficiency of the adaptive algorithm. The GSC-MCC method could combine the advantages of the GSC method and the GSC-CC method to reduce the distortion efficiently to adapt to nonstationary speech signals under different SNR conditions.



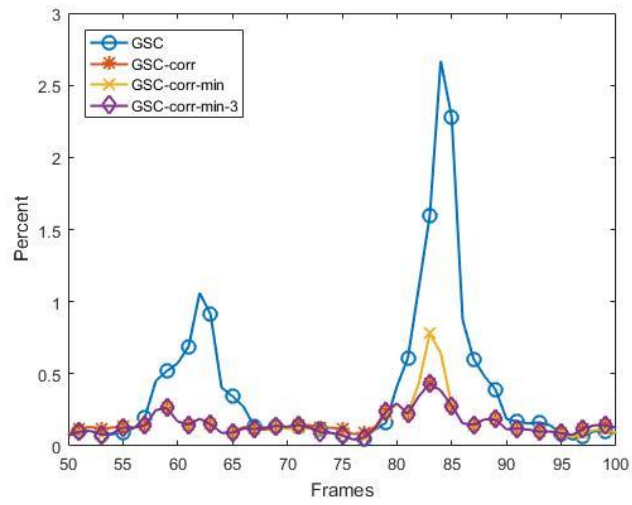
(a) SNR = -15dB



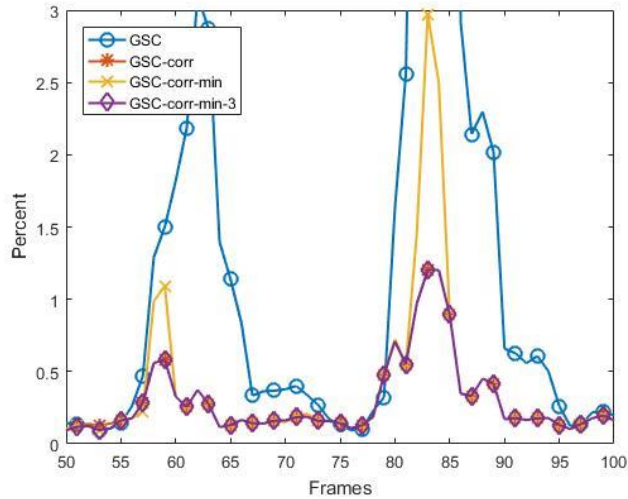
(b) SNR = -10dB



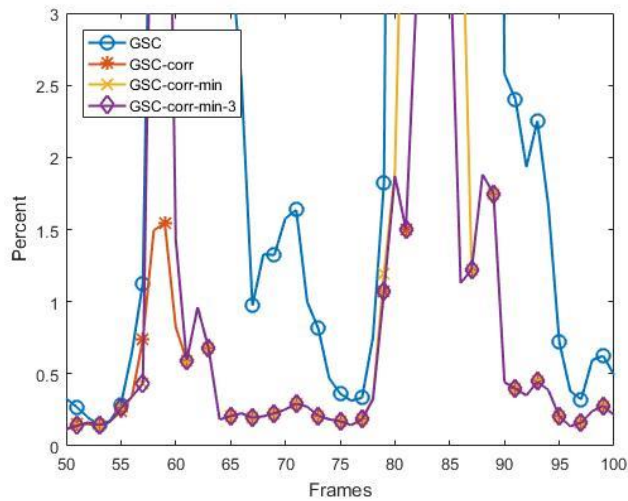
(c) SNR = -5dB



(d) SNR = 0dB



(e) SNR = +5dB



(f) SNR = +10dB

Figure 3-8. Distortion-noise ratio of frames with different methods from the 50th to 100th frame under different SNR conditions.

3.4 GSC method with sidelobe neutralization

Based on Chapter 2.1, the microphone array is determined with 6 microphones in a circle which radius is 0.1m. Because the configuration of the microphone array is determined, the sidelobe of the beamforming pattern could be calculated as Equation 2-15. If there are two sound sources s_1 and s_2 , deployed in the direction vector v_1 and v_2 , respectively, in the sound field, the sidelobe neutralization method could be derived using the following equations:

$$S = [s_1, s_2], \quad V = [v_1, v_2]. \quad (3-16)$$

The received signal of microphones x_m could be written as:

$$x_m = SV = s_1 v_{(1,m)} + s_2 v_{(2,m)}. \quad (3-17)$$

The output signal of the beamforming function to the specific direction is:

$$y_{v_1} = \sum_{m=1}^M \omega_{(s_1,m)}^* x_m = \sum_{m=1}^M v_{(1,m)}^* s_1 v_{(1,m)} + v_{(1,m)}^* s_2 v_{(2,m)} = s_1 + \sum_{m=1}^M v_{(1,m)}^* s_2 v_{(2,m)}, \quad (3-18)$$

$$y_{v_2} = \sum_{m=1}^M \omega_{(s_2,m)}^* x_m = \sum_{m=1}^M v_{(2,m)}^* s_1 v_{(1,m)} + v_{(2,m)}^* s_2 v_{(2,m)} = s_2 + \sum_{m=1}^M v_{(2,m)}^* s_1 v_{(1,m)}. \quad (3-19)$$

$$\text{Let } u_{ab} = \sum_{m=1}^M v_{(a,m)}^* v_{(b,m)} = B(\phi_a, \phi_b), \quad (3-20)$$

$$y_{v_1 s_2} = y_{v_1} - u_{12} y_{v_2} = s_1 + u_{12} s_2 - u_{12} s_2 - u_{12} u_{21} s_1 = s_1 (1 - u_{12} u_{21}) = s_1 (1 - |u_{12}|^2), \quad (3-21)$$

where u_{ab} represents the attenuation coefficient of the sidelobe when the sound source is placed at direction b and beamformed to direction a . Moreover, the attenuation coefficient will be the conjugate of u_{ab} when the cast direction is inverse.

Hence, Equation 3-21 could be used for direction v_1 to reduce the noise received from direction v_2 . This equation proves that the noise signal from the interference direction can theoretically be removed entirely from the beamforming output signal of the looking direction. When the attenuation coefficient is not equal to one, the signal amplitude from the looking direction could be recovered by multiplying a proper scale factor. On the other hand, if the attenuation coefficient of the frequency is equal to one, this frequency signal would be lost completely and cannot be compensated.

Therefore, the sidelobe neutralization method could be modified as Equations 3-23 by combining Equations 3-22. The desired signal component in the noise direction was calculated and subtracted from the beamforming output of the noise direction as Equations 3-22. Hence, the estimated noise signal becomes the pure noise without the desired signal. The output of the sidelobe neutralization method is calculated using

Equations 3-23. This equation proves that the sidelobe neutralization method would not lose the frequency signal with an attenuation coefficient of even one. Because most time u_{ab} is less than 1, the cube of u_{ab} is less than itself. Combine Equations 3-18 and Equations 3-23, it means $(u_{12})^3$ would be less than u_{12} most time. Thus, the beamforming output with less noise is obtained.

$$y_{v_2s_1} = y_{v_2} - u_{21}y_{v_1} = s_2 + u_{21}s_1 - u_{21}s_1 - u_{21}u_{12}s_2 = s_2(1 - u_{21}u_{12}) = s_2(1 - |u_{21}|^2). \quad (3-22)$$

$$y_{v_1s_2s_1} = y_{v_1} - u_{12}y_{v_2s_1} = s_1 + u_{12}s_2 - u_{12}s_2(1 - |u_{21}|^2) = s_1 + s_2(u_{12})^3. \quad (3-23)$$

However, the estimated noise direction is often inaccurate in practice. The error of the estimated noise direction would affect the performance of the sidelobe neutralization method. v_3 is assumed to be the inaccurately estimated noise direction and Equations 3-22 and 3-23 can be modified as follows:

$$y_{v_3} = \sum_{m=1}^M \omega_{(s_3,m)}^* x_m = \sum_{m=1}^M v_{(3,m)}^* s_1 v_{(1,m)} + v_{(3,m)}^* s_2 v_{(2,m)} = u_{31}s_1 + u_{32}s_2. \quad (3-24)$$

$$y_{v_3s_1} = y_{v_3} - u_{31}y_{v_1} = u_{31}s_1 + u_{32}s_2 - u_{31}s_1 - u_{31}u_{12}s_2 = s_2(u_{32} - u_{31}u_{12}). \quad (3-25)$$

$$y_{v_1s_3s_1} = y_{v_1} - u_{13}y_{v_3s_1} = s_1 + u_{12}s_2 - u_{13}s_2(u_{32} - u_{31}u_{12}) = s_1 + s_2[u_{12} - u_{13}(u_{32} - u_{31}u_{12})]. \quad (3-26)$$

$$\text{Let } r_{v_3} = \left| \frac{u_{12} - u_{13}(u_{32} - u_{31}u_{12})}{u_{12}} \right|, \quad (3-27)$$

hence, r_{v_3} is the ratio that the noise signal processed by the sidelobe neutralization method with the inaccurate noise direction to by the beamforming method. r_{v_3} can be qualitatively analyzed as follows:

if u_{12} is small or near to zero, r_{v_3} may be extremely big,

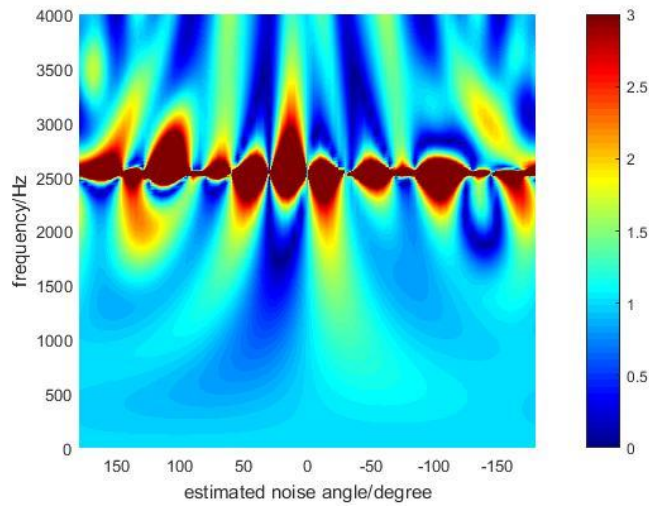
$$y_{v_1s_3s_1} = s_1 + s_2[u_{12} - u_{13}(u_{32} - u_{31}u_{12})] = s_1 + s_2(-u_{13}u_{32}). \quad (3-28)$$

If u_{12} is big, there are roughly two cases:

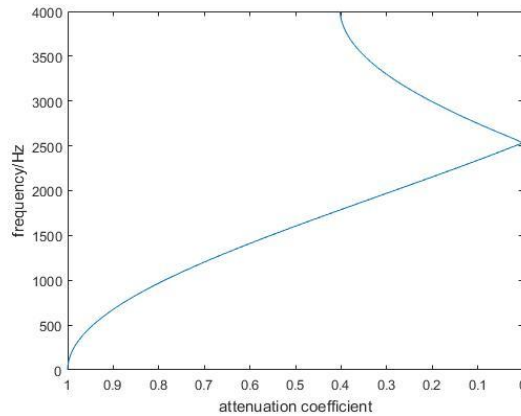
$$(1) \ v_3 \text{ is near } v_2, \text{ thus, } r_{v_3} = \left| 1 - \frac{u_{13}u_{32}}{u_{12}} + u_{13}^2 \right| \approx \left| 1 - \frac{u_{12}u_{22}}{u_{12}} + u_{13}^2 \right| = u_{13}^2 \leq 1. \quad (3-29)$$

(2) v_3 is far from v_2 , thus, u_{32} is small, $r_{v_3} = \left| 1 - \frac{u_{13}u_{32}}{u_{12}} + u_{13}^2 \right| \leq \left| 1 + u_{13}^2 \right| + \left| \frac{u_{32}}{u_{12}} \right| \leq 3. \quad (3-30)$

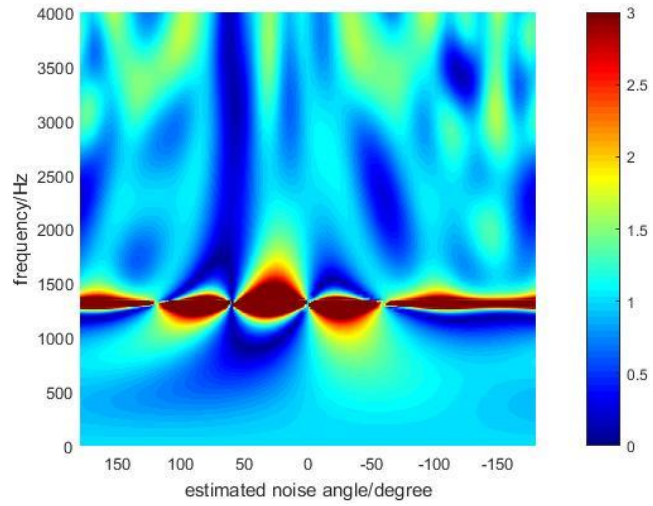
Assume an instance, which v_1 is 0° and v_2 is $30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ, 180^\circ$, respectively, to illustrate the affection of the inaccurately estimated noise direction for the residual noise. According to Equation 3-27, Figure 3-9 shows r_{v_3} in different estimated noise directions (v_3 from -180° to $+180^\circ$).



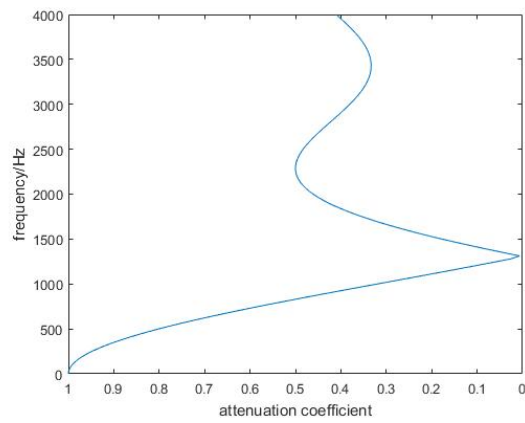
(a1) Actual noise direction: 30°



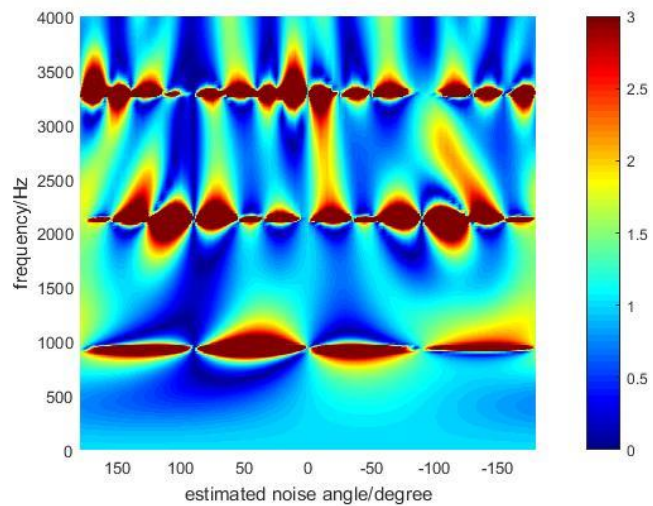
(b1) Actual noise direction: 30°



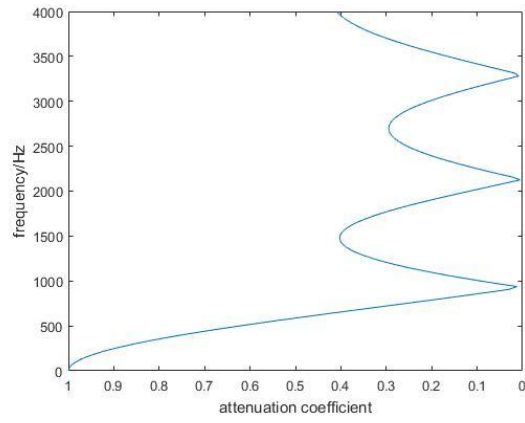
(a2) Actual noise direction: 60°



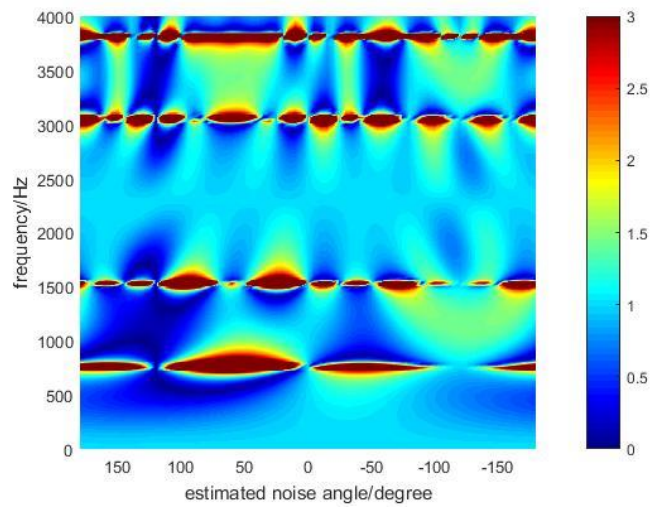
(b2) Actual noise direction: 60°



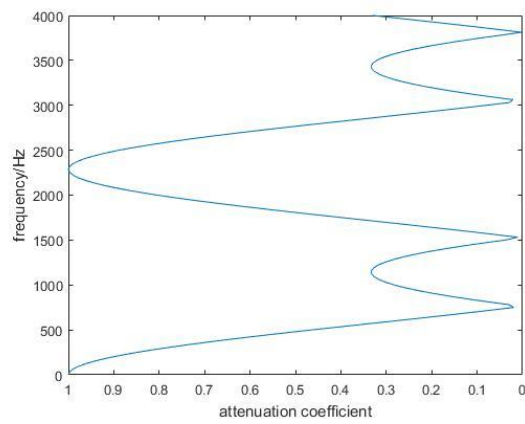
(a3) Actual noise direction: 90°



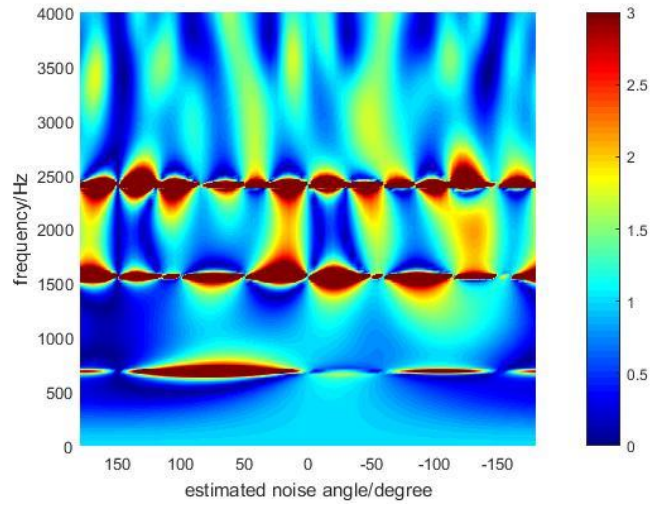
(b3) Actual noise direction: 90°



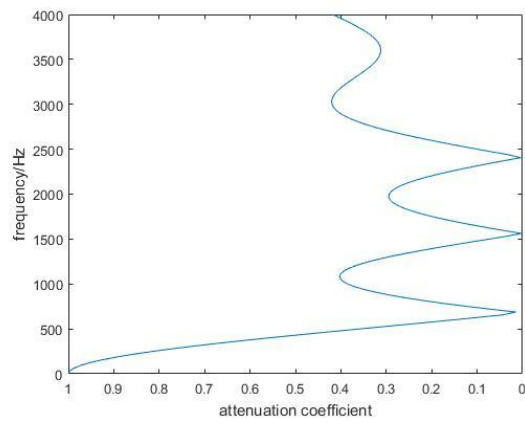
(a4) Actual noise direction: 120°



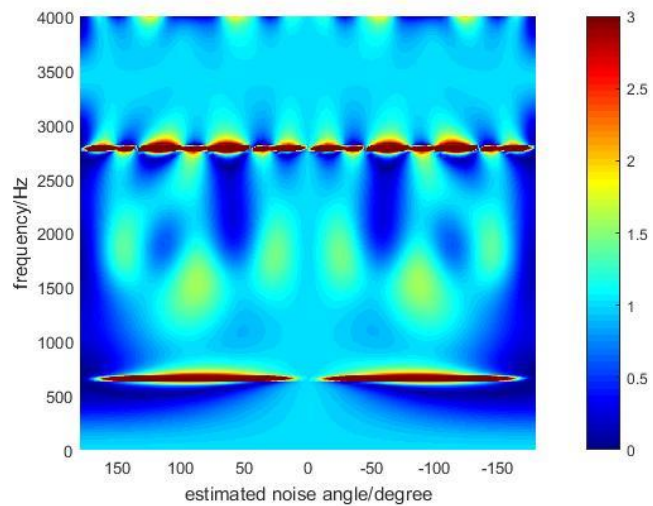
(b4) Actual noise direction: 120°



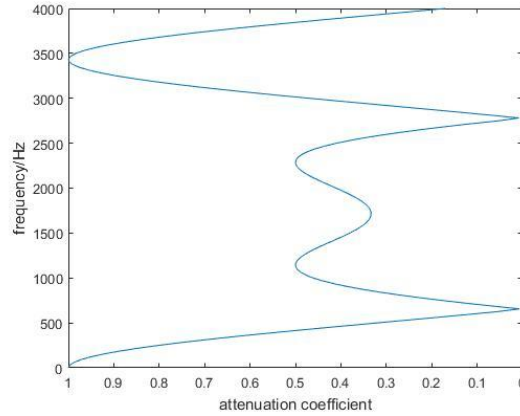
(a5) Actual noise direction: 150°



(b5) Actual noise direction: 150°



(a6) Actual noise direction: 180°



(b6) Actual noise direction: 180°

Figure 3-9. (a1~a6) Ratio of the noise processed using the sidelobe neutralization method with the different estimated noise directions and different actual noise directions. (b1~b6) Coefficient of sidelobe attenuation in the actual noise direction.

The figures show that when the actual noise direction is farther from the speech signal direction, the number of the polar point of the coefficient of the sidelobe attenuation in actual noise direction becomes more and the grating lobe may be generated.

The blue gap near the actual noise direction in Figure 3-9 (a1~a6) shows that the residual noise signal processed by the sidelobe neutralization method with an inaccurate noise direction is less than by the original beamforming method in a relatively wide direction range near the actual noise direction. When the sidelobe attenuation coefficient of the actual noise direction is close to one, it means the grating lobe is generated and the microphone array loses the ability of the spatial filter in this frequency band. Even the estimated noise direction is mismatched a lot, the obtained noise would be almost equal to the residual noise of the original beamforming method.

When the sidelobe attenuation coefficient of the actual noise direction is close to zero, the ratio of the residual noise processed using the sidelobe neutralization method would be big. It is because that the sidelobe attenuation of the actual noise direction is strong, namely the denominator of Equation 3-27 is small which leads to the ratio is

big, even though the real noise component of the noise processed by the sidelobe neutralization method is not so big. When the estimated noise direction is far away from the actual noise direction, the ratio value is close to one in figures, which means that the energy of the noise signal processed by the sidelobe neutralization method is almost equal to that by the original beamforming method. In the high-frequency band in several directions, the ratio becomes higher, but it is still less than two.

Therefore, the conclusion could be derived that the sidelobe neutralization method could work effectively when the estimated noise direction is near the actual noise direction. When the estimated noise direction is far away from the actual noise direction, the denoising performance of the sidelobe neutralization method is similar to the beamforming method except in the frequency range which the effect of the sidelobe attenuation in the actual noise direction is strong.

Thus, combining the LMS theory, in which the denoising performance of the algorithm is related to the energy of the received noise signal [46], the GSC-SN method is proposed by adding the sidelobe neutralization method to the primary path of the conventional GSC method. The block diagram of the GSC-SN method is presented in Figure 3-10. The sidelobe neutralization method is used to reduce the output noise of the beamforming method to improve the performance of the GSC method. When the noise weight vector is estimated to be ω_n , the noise signal $y_n(n)$ in the noise direction could be estimated using Equation 3-19. The desired signal in the noise direction $y_{nm}(n)$ is calculated and subtracted from the estimated noise signal as Equation 3-22. Therefore, the pure estimated noise signal $y_{nsm}(n)$ is obtained without the desired signal. The pure estimated noise signal is subtracted from the beamforming output signal of the desired signal direction $y_m(n)$. Thus, the beamforming output of the primary path with less noise

is obtained as Equation 3-23. The denoising performance of the GSC method could be improved when the beamforming output signal with less noise.

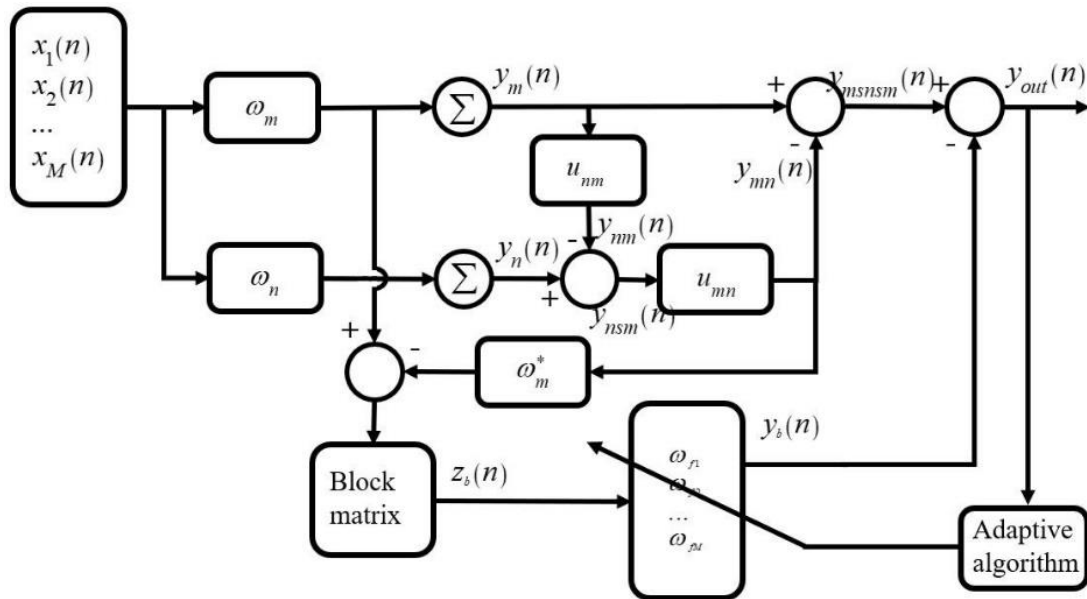


Figure 3-10. Block diagram of the GSC method with sidelobe neutralization.

3.5 Proposed GSC-SN-MCC method

Finally, the GSC-SN-MCC method was proposed by combining the GSC-MCC method and the GSC-SN method to improve the performance of the GSC method. The GSC-MCC method was used to reduce the distortion of the enhanced signal caused by the varying desired signal. The sidelobe neutralization method was placed in the primary path of the GSC-MCC method to generate the beamforming output signal with less noise to improve the denoising ability of the proposed method.

As shown in Figure 3-11, the received signal of the microphone array is beamformed in the desired signal direction ω_m and the noise signal direction ω_n , respectively. Based on Equation 3-18, 3-19, 3-22, 3-23, the beamforming output signal with less noise than the original beamforming method is obtained. Then the system is run as usual GSC method until updating the adaptive filter coefficient. The cross-correlation coefficient ρ (between the estimated noise signal and the corresponding

output signal) is added to the adaptive filter coefficient update equation to control the overestimating of the noise component, as Equation 3-10.

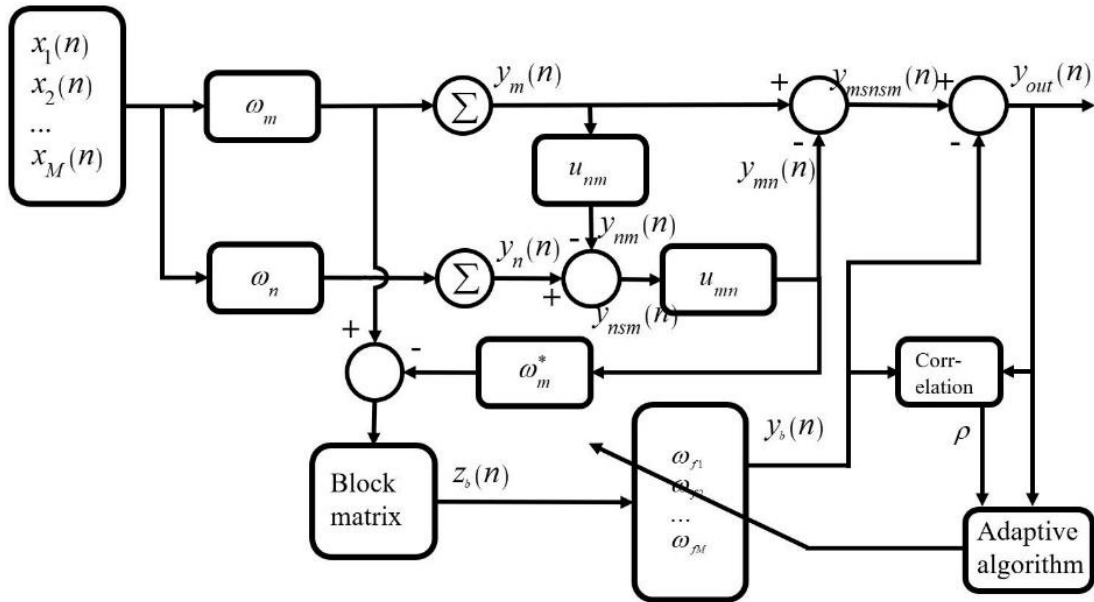


Figure 3-11. Block diagram of the GSC-SN-MCC method.

The GSC-SN-MCC method would perform like the following steps:

(1) The received signals are aligned in the looking direction and the estimated noise direction, respectively. Sum the aligned signal to obtain the corresponding output signal ($y_m(n)$), ($y_n(n)$) of the fixed beamforming method.

(2) The sum signal in the looking direction would be mapped to the estimated noise direction based on the beamforming pattern equation. The sidelobe value ($y_{nm}(n)$) is obtained.

(3) The sidelobe value of the sum signal in the looking direction ($y_{nm}(n)$) would be subtracted from the sum signal ($y_n(n)$) in the estimated noise direction, the estimated noise in the estimated noise direction without the desired signal component ($y_{nsm}(n)$) is obtained.

(4) The estimated noise in the estimated noise direction without the desired signal component ($y_{nsm}(n)$) would be mapped to the looking direction based on the

beamforming pattern equation. The sidelobe value ($y_{mn}(n)$) is obtained.

(5) The sidelobe value of the sum signal in the estimated noise direction ($y_{mn}(n)$) would be subtracted from the sum signal ($y_m(n)$) in the looking direction. The summed signal in the looking direction with less noise component and undamaged desired signal ($y_{msnsm}(n)$) is obtained.

(6) The estimated noise in the estimated noise direction without the desired signal component ($y_{nsm}(n)$) would be projected to every microphone in the array. The projected signal would be subtracted from the aligned received signals (in the looking direction, $y_m(n)$) to reduce the noise component in the aligned received signals.

The aligned received signals in the looking direction with less noise component would be input to the blocking matrix to get the pure reference noise without the desired signal component.

(7) The pure reference noise signal would be put into the adaptive filter to generate the estimated noise signal ($y_b(n)$).

(8) The estimated noise signal ($y_b(n)$) would be subtracted from the output signal of the beamformer with less noise ($y_{msnsm}(n)$) to get the error signal ($y_{out}(n)$). The error signal ($y_{out}(n)$) is the final output signal of the GSC-SN-MCC method.

(9) The cross-correlation coefficient ρ between the estimated noise signal ($y_b(n)$) and the error signal ($y_{out}(n)$) is used to control the step size of the LMS algorithm to update the coefficient of the adaptive filter.

3.6 Conclusion

This chapter introduced the conventional GSC method and proposed a modified GSC method to reduce the distortion of the enhanced signal. Distortion of the enhanced

audio signal consists of two parts: the residual acoustic noise and the distortion of the desired audio signal, which means that the desired audio signal is damaged. A modified GSC method using cross-correlation and sidelobe neutralization was proposed to reduce both kinds of distortion if the desired signal was a nonstationary speech signal.

First, the GSC-MCC method was proposed that the cross-correlation coefficient between the canceling signal and the error signal of the LMS algorithm was added to the adaptive process of the GSC method. The cross-correlation coefficient was used to control the step size of the update process of the LMS algorithm. It meant that if the cross-correlation coefficient was small, the step size of the updating would be small and the weight coefficient will be changed only slightly. Otherwise, the weight coefficient would be changed considerably to converge rapidly.

Then the sidelobe neutralization method was proposed to reduce the residual noise component in the output signal of the fixed beamforming method. When the estimated noise direction was correct, the residual noise component could be reduced effectively. The formula demonstrates that even the estimated noise direction was not inaccurate, the amount of the noise component in the output signal of the beamforming method processed by the sidelobe neutralization method was still similar to the original noise component. The figures of the simulation result were displayed with the different estimated noise directions. The noise component of the beamforming output signal could be decreased by subtracting the estimated noise signal to improve the denoising performance of the GSC method, which was referred to as the GSC-SN method.

Finally, the GSC-SN-MCC method was proposed by merging the above GSC-MCC and GSC-SN methods. The detailed performed steps were provided.

4. Experiment and Analysis

4.1 Experiment implementation

The laboratory experiments were conducted in the anechoic chamber to validate the proposed method. The experimental layout is shown in Figure 4-1. The laptop was the control center to manage the audio system for simulating a real environment. Two omnidirectional speakers were connected to the laptop as the speech and noise source, respectively, using an INTERM L-2400 power amplifier. The microphone array consisted of 6 MEMS microphones distributed uniformly on a circle with a radius of 0.1 meters. The printed circuit board (PCB) of the SCIEN company was connected to the microphones to realize the signal amplification function, A/D conversion function and transfer the digital signal to the laptop for recording. The signal was processed by MATLAB software. The speakers and the microphone array were deployed in the same horizontal plane to simplify the experiment. This meant that the elevation angles of the incident signals (desired signal and noise signal) were 90° .

Based on the common far-field equation as Equation 4-1, the minimum distance between the microphone array and the sound source was calculated to ensure that the position of the speakers satisfied the far-field condition. Assuming that the sound speed c in air was 343m/s, the highest frequency f was 4000Hz and the diameter of microphone array L was 0.2m. Hence, the minimum distance d was approximately 0.933m. Thus, the speakers were deployed about 2.5m away from the microphone array considering the space of the anechoic chamber and the minimum distance. The azimuth angles of the speech signal source and the noise signal source were 0° and 90° , respectively.

$$d = 2 * L^2 / \lambda = 2 * L^2 * f / c . \quad (4-1)$$

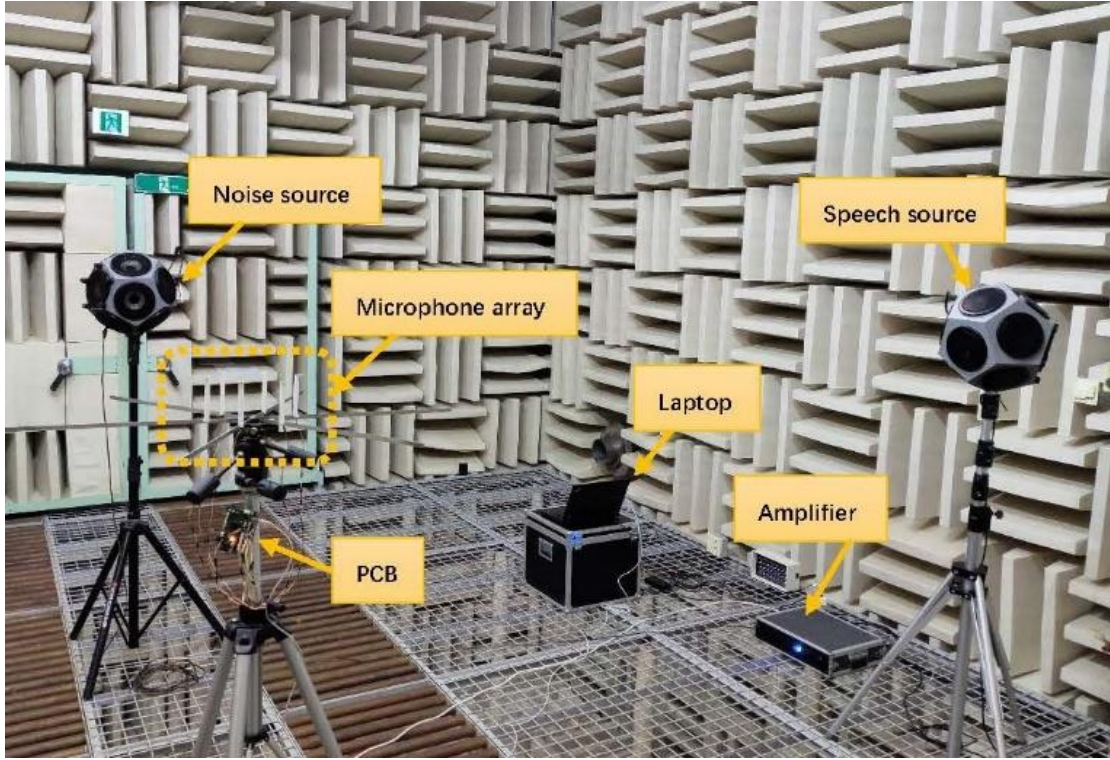


Figure 4-1. Experiment layout in the anechoic chamber.

The speech of a female was selected randomly from the TIMIT database [72] as the desired signal. White noise in the NOISEX-92 database [73] was chosen as the noise signal. The power of the noise signal was adjusted to simulate different SNR conditions. The sampling rate of the microphone array was 16000Hz and the received signal was filtered using a lowpass filter which cutoff frequency is 4000 Hz. The high sampling rate could increase the time resolution of the system and exhibit the phase difference between microphones more obviously. Because the most energy of speech is concentrated in the low and medium frequency bands, the received signal was filtered by a lowpass filter to reduce the calculation burden of the hardware, while most of the speech information was preserved.

4.2 Experiment result analysis

The distortion of the enhanced speech signal was calculated using Equation 4-2,

where k denoted the k th frame and l is the length of the frame. $y_{out}(n)$ and $s_{speech}(n)$ represented the output signal of the method and the clean desired speech signal without noise, respectively. Less distortion meant better denoising performance of the method. Equation 4-3 calculated the distortion-noise ratio as the distortion divided by the pure noise energy to normalize the distortion for comparing the distortion degree under different SNR conditions.

$$distortion(k) = \sum_{n=k*l}^{(k+1)*l} [y_{out}(n) - s_{speech}(n)]^2 \quad (4-2)$$

$$p(k) = \left| \frac{distortion(k)}{noise(k)} \right| = \frac{\sum_{n=k*l}^{(k+1)*l} [y_{out}(n) - s_{speech}(n)]^2}{[s_{noise}(n)]^2} \quad (4-3)$$

For comparison, the average value of the normalized distortion-noise ratio and the normalized distortion value of the k th frame processed by the GSC method was as the unit value. The results by other methods were normalized as Equations 4-4 and 4-5.

$$r_p = \frac{avg(\sum p_{method}(k))}{avg(\sum p_{GSC}(k))} \quad (4-4)$$

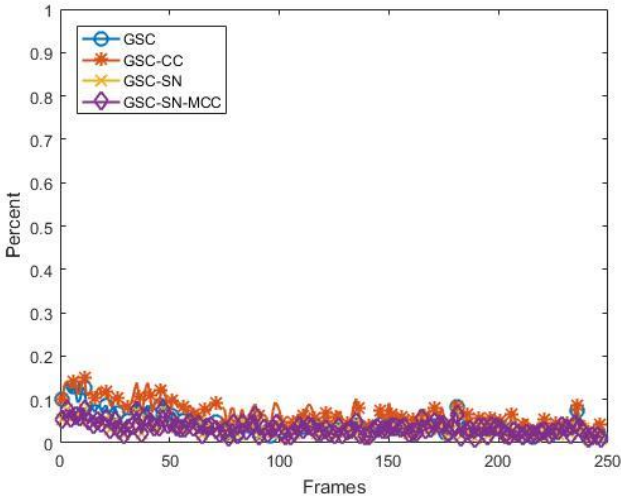
$$r_d(k) = \frac{distortion_{method}(k)}{distortion_{GSC}(k)} \quad (4-5)$$

4.2.1 Effect of the various SNR conditions

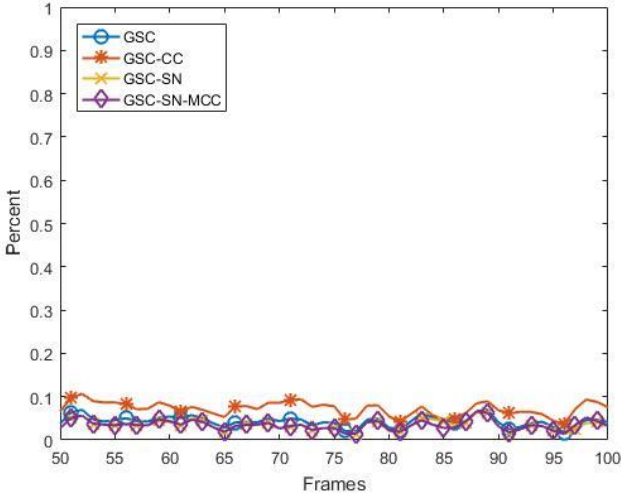
Figure 4-2 compares the distortion-noise ratio of frames by different methods while SNR is -20dB, -10dB, -5dB, 0dB, +5dB, +10dB, respectively. Figure 4-2 (b1~b6) is the enlarged views of Figure 4-2 (a1~a6) from the 50th to 100th frames on different SNR conditions, which includes two peaks of the speech frame energy. The figures show that peaks of the signal distortion are generated using the GSC method when the energy of the speech frame is increased suddenly.

When the SNR is low (-20dB), the speech signal is almost submerged by the noise signal. The residual noise is barely affected by the nonstationary speech signal. With

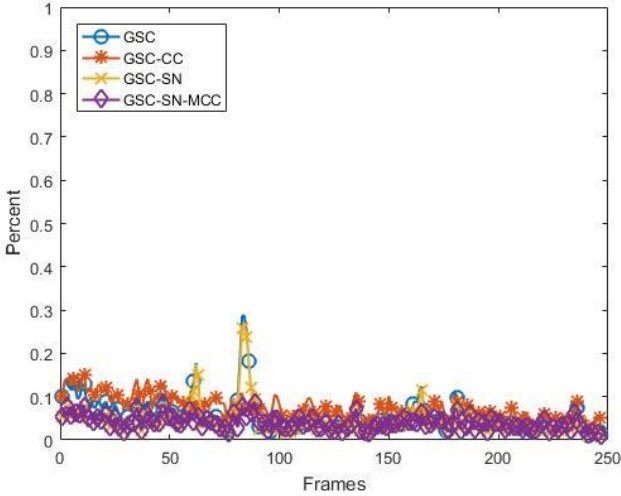
the increase of the SNR, the peaks of the signal distortion become more apparent.



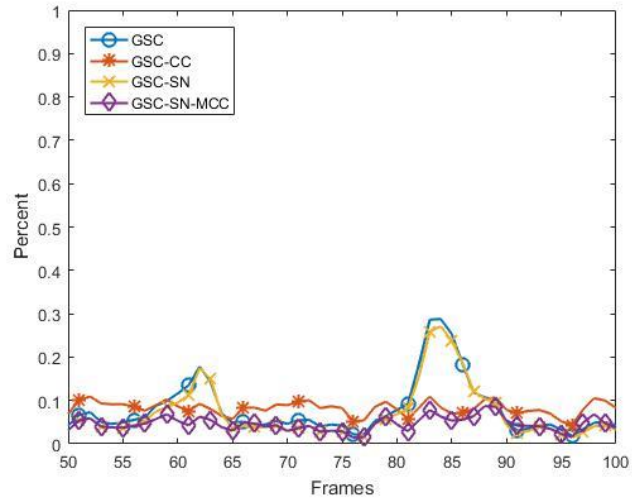
(a1) SNR = -20dB



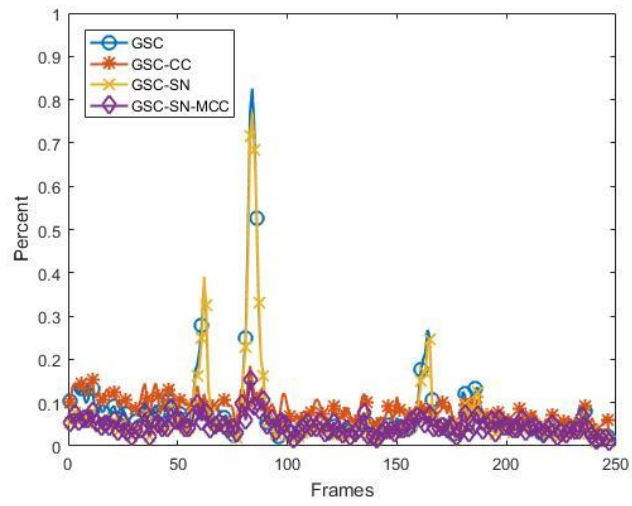
(b1) SNR = -20dB



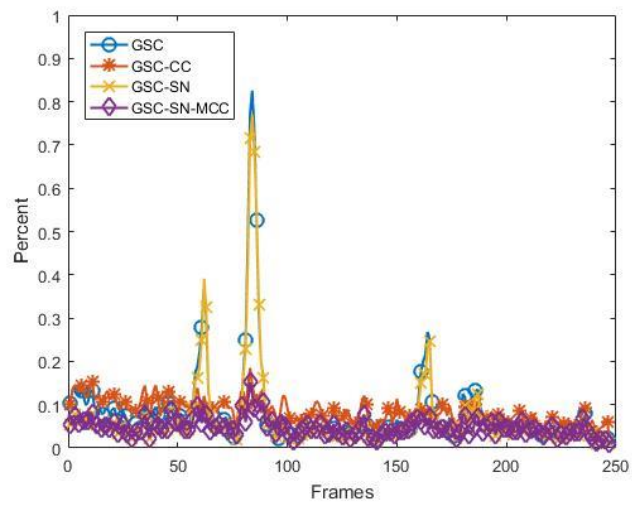
(a2) SNR = -10dB



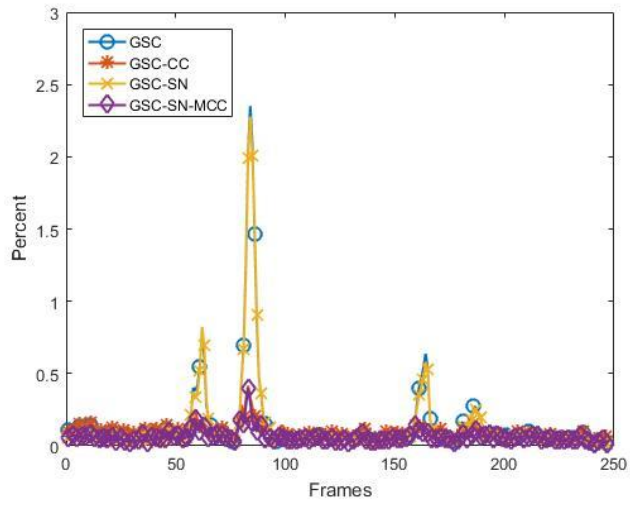
(b2) SNR = -10dB



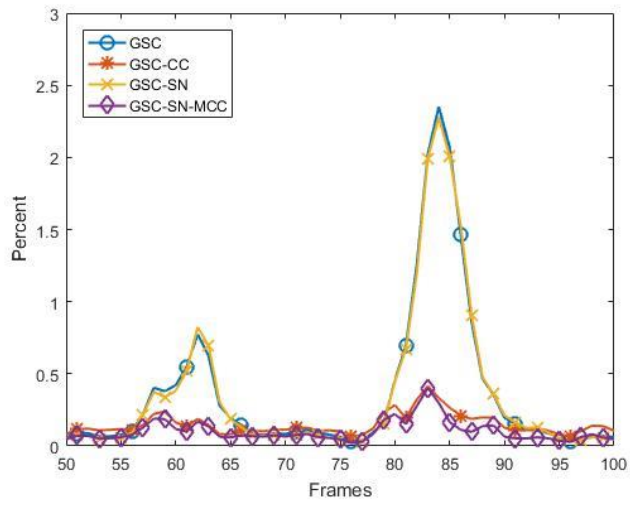
(a3) SNR = -5dB



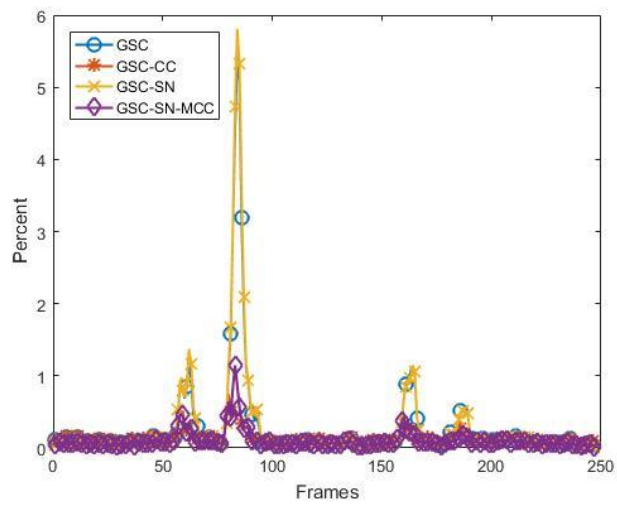
(b3) SNR = -5dB



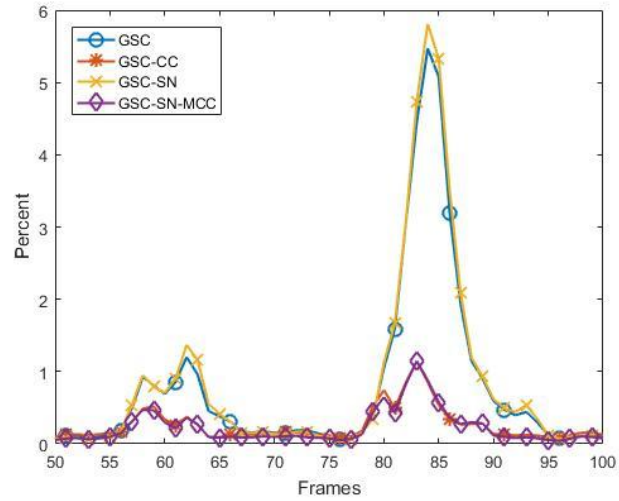
(a4) SNR = 0dB



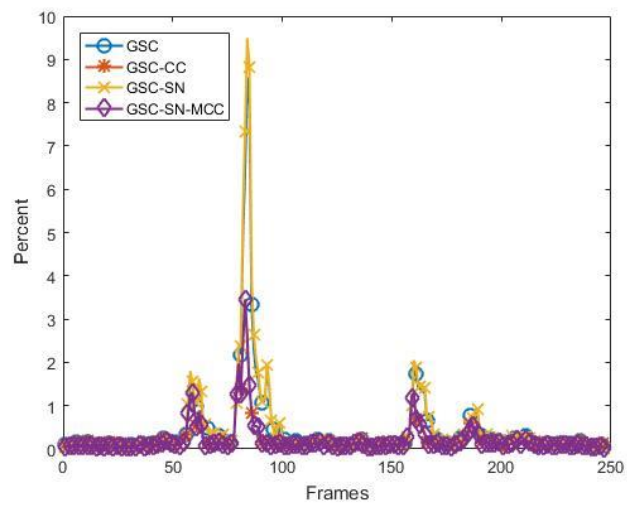
(b4) SNR = 0dB



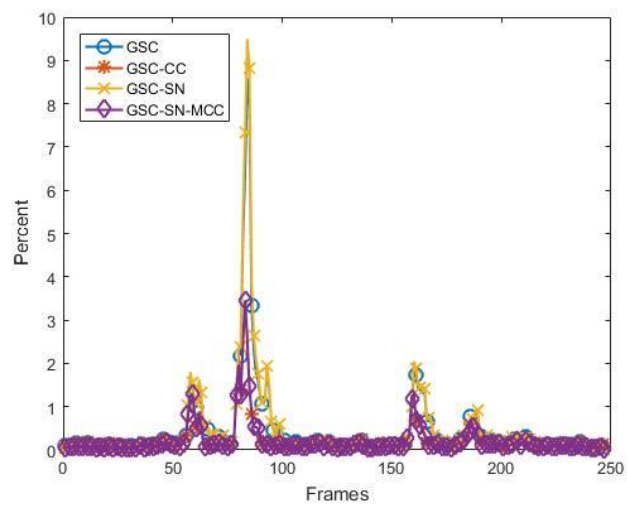
(a5) SNR = +5dB



(b5) SNR = +5dB



(a6) SNR = +10dB

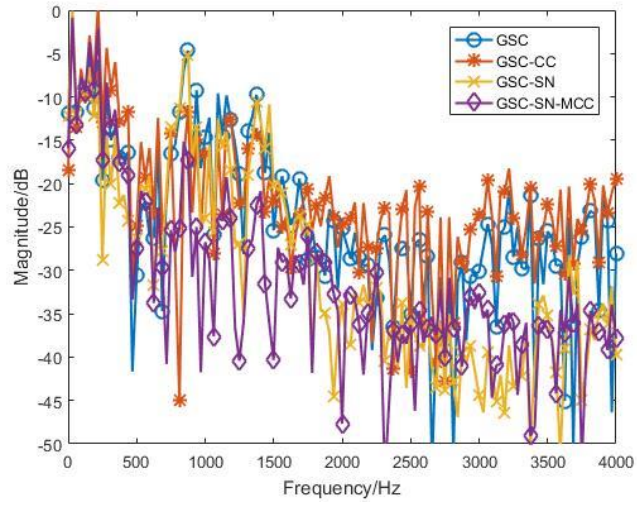


(b6) SNR = +10dB

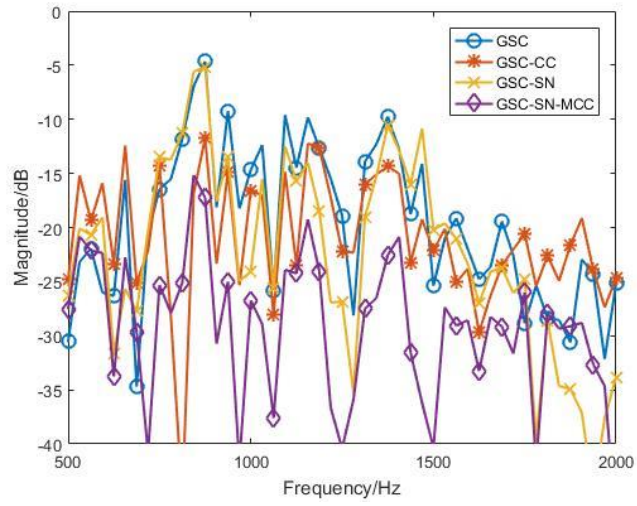
Figure 4-2. Distortion-noise ratio of frames by different methods on different SNR conditions: (a1~a6) all frames; (b1~b6) from the 50th to 100th frame.

The figure illustrates that the GSC-CC method could alleviate this problem while the residual noise will be increased at other frames that the speech energy is stable. While the residual noise of the GSC-CC method is higher than the GSC method. The GSC-SN method could reduce the distortion caused by the residual noise, but the distortion problem mentioned above that the damage of the desired signal will still exist. The GSC-SN-MCC method could take advantage of the above two methods to decrease both kinds of distortion of the enhanced signal on different SNR conditions. The residual noise of the GSC-SN-MCC method would be similar to the GSC-SN method when the speech signal is stable. The peak value of the speech distortion of the GSC-SN-MCC method would be as small as the GSC-CC method when the energy of the speech signal is increased suddenly.

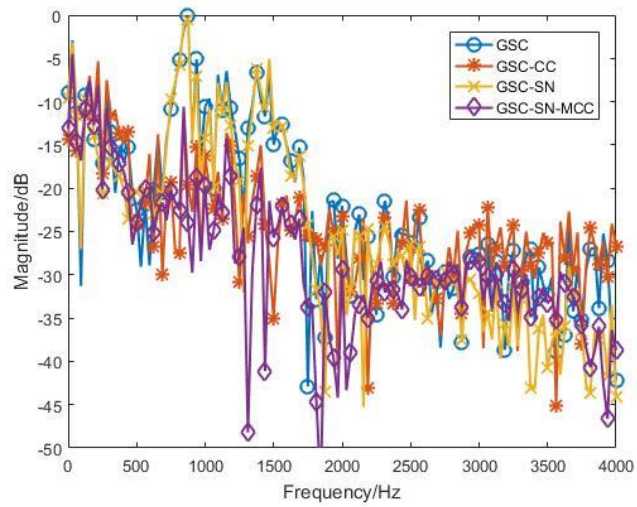
As shown in Figure 4-2, the maximum distortion occurred at the 83rd frame in this case. Figure 4-3 presents the spectrum of signal distortion of the 83rd frame on different SNR conditions where the fast Fourier transform is utilized. For convenience, the amplitude of the spectrum is normalized. Figure 4-3 (a1~a6) displays the spectrum of signal distortion of the 83rd frame over the entire frequency band with different methods on different SNR conditions. Figure 4-3 (b1~b6) shows the enlarged view of the frequency band (500Hz~2000Hz) of the corresponding Figure 4-3 (a1~a6). When the modified GSC methods are applied, there are apparent differences in the amplitude from 500Hz to 2000Hz, which is the frequency band that the speech energy is concentrated.



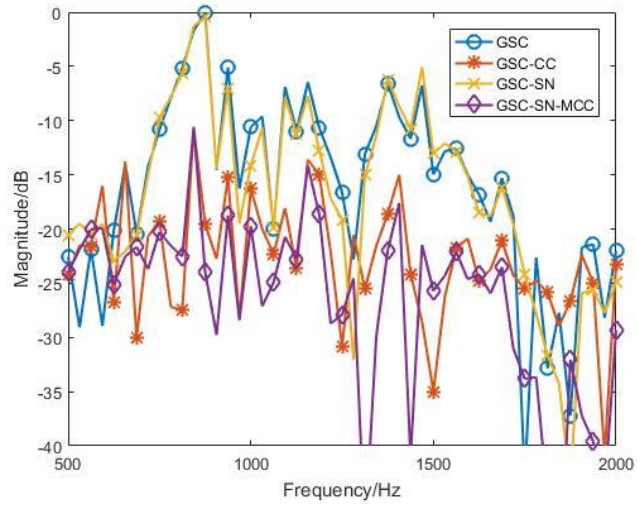
(a1) SNR = -20dB



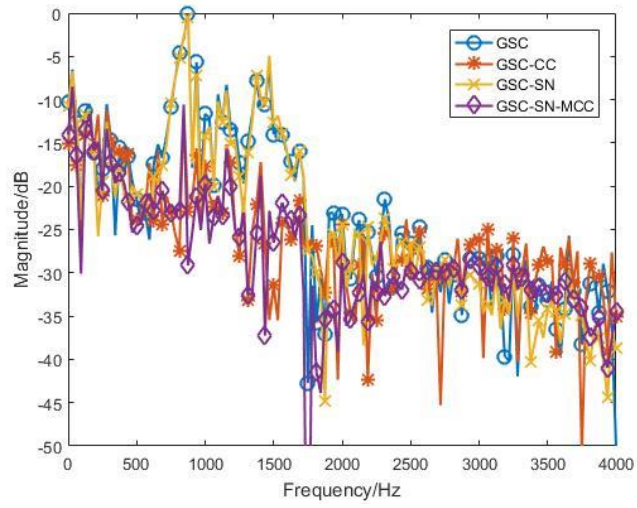
(b1) SNR = -20dB



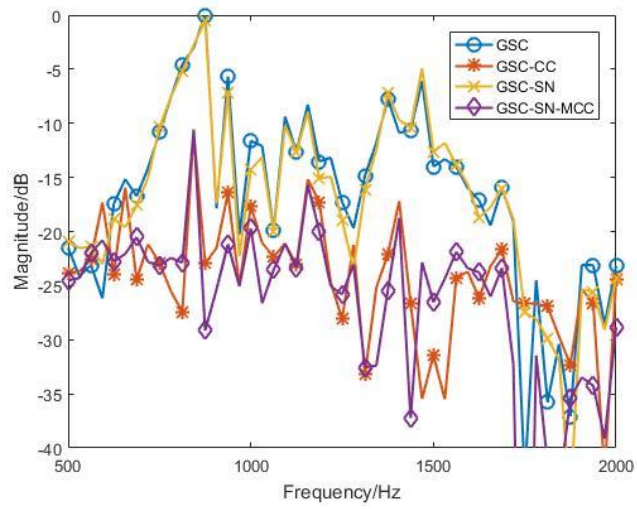
(a2) SNR = -10dB



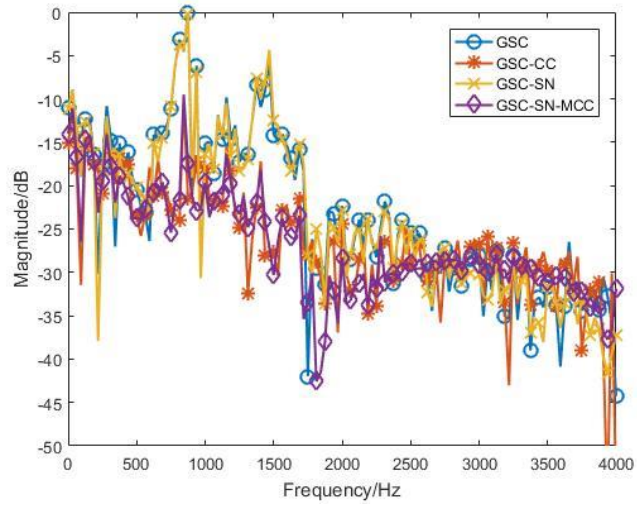
(b2) SNR = -10dB



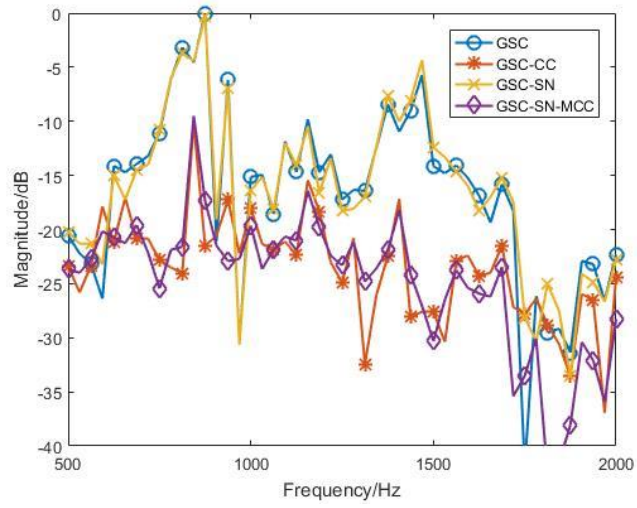
(a3) SNR = -5dB



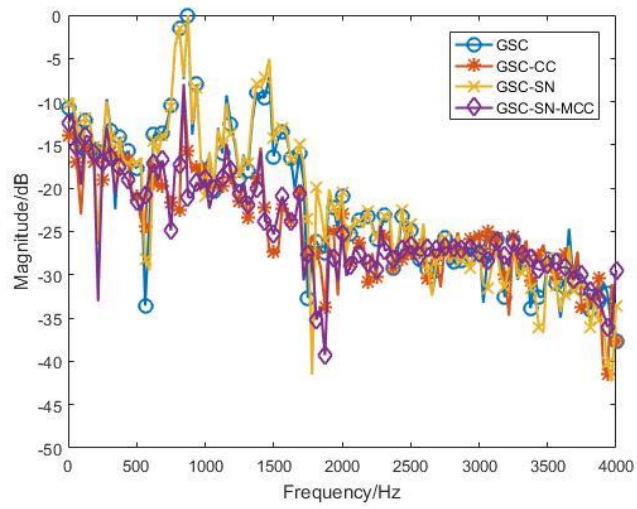
(b3) SNR = -5dB



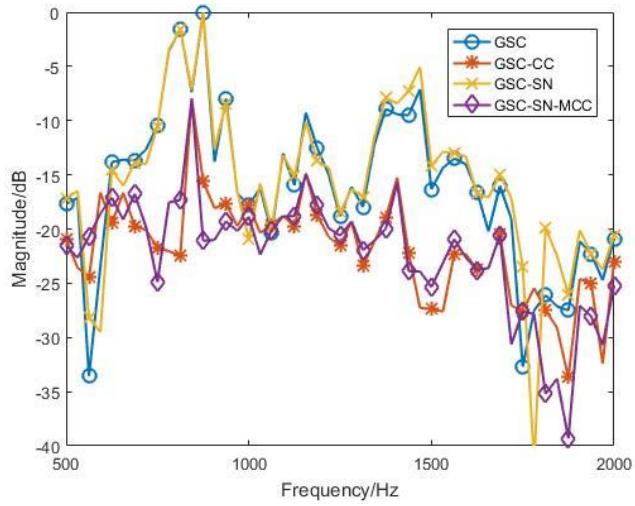
(a4) SNR = 0dB



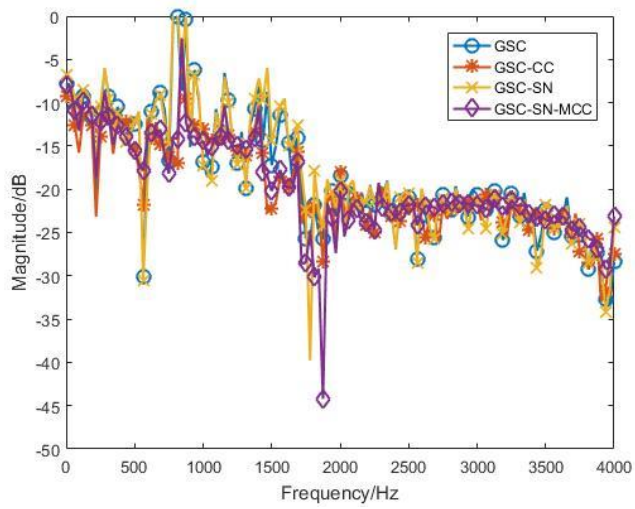
(b4) SNR = 0dB



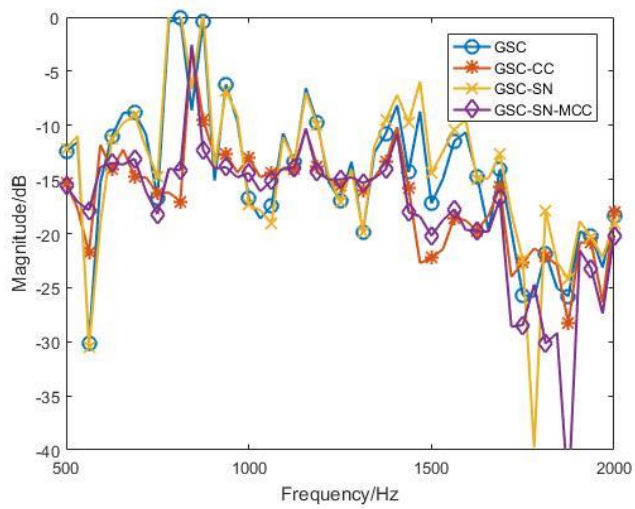
(a5) SNR = +5dB



(b5) SNR = +5dB



(a6) SNR = +10dB



(b6) SNR = +10dB

Figure 4-3. Spectrum of the signal distortion of the 83rd frame by different methods on different SNR conditions: (a1~a6) entire frequency band; (b1~b6) from 500Hz to 2000Hz.

The figures show that the trend of the distortion of the enhanced signal by the GSC method with the SNR increasing is a curve. When SNR is -20dB, the background noise is so strong that the speech signal is submerged under the background noise. The change of speech energy rarely affects the distortion of the enhanced signal. When SNR is between -10dB and 0dB, the distortion of the enhanced signal becomes more prominent with the SNR increasing because the increased speech energy raises the energy of the error signal of the LMS algorithm and influences the coefficient update process of the adaptive filter. When SNR is between 0dB and +10dB, the distortion of the enhanced signal becomes flat with the SNR increasing. This is because the distortion of the enhanced signal in other frequencies is increased lead that the spectrum of the signal distortion becomes flat.

The GSC-CC method could reduce the peak value of the spectrum of the signal distortion while the residual noise component is a litter higher than other methods. The GSC-SN method could decrease the residual noise when the speech energy is not noticeable while barely affecting the peak value of the signal distortion. The GSC-SN-MCC could reduce the peak value of the distortion and the residual noise component at the same time. The distortion spectrum of the GSC-SN-MCC method is smaller and flatter than other GSC methods. Hence, the distortion of the enhanced signal caused by damaging the desired signal and the residual noise component is reduced successfully using the proposed method.

Table 4-1 and 4-2 list the denoising performance of different methods under various SNR conditions. Table 4-1 lists the average of the normalized distortion-noise ratio and Table 4-2 presents the normalized distortion value of the 83rd frame (maximum distortion frame) under various conditions.

Table 4-1. Average of the normalized distortion-noise ratio by different methods under various SNR conditions.

SNR/dB	GSC	GSC-CC	GSC-SN	GSC-SN-MCC
+20	100%	65.98%	104.03%	75.75%
+10	100%	57.93%	100.99%	46.29%
+5	100%	54.46%	97.92%	38.51%
0	100%	72.66%	89.87%	44.14%
-5	100%	105.34%	83.98%	59.01%
-10	100%	134.85%	80.22%	72.41%
-20	100%	157.90%	77.21%	78.84%

Normalized to result of the GSC method.

Table 4-2. Normalized distortion ratio of the 83rd frame by different methods under various SNR conditions.

SNR/dB	GSC	GSC-CC	GSC-SN	GSC-SN-MCC
+20	100%	54.09%	109.34%	82.73%
+10	100%	39.69%	110.58%	40.31%
+5	100%	20.91%	106.16%	20.94%
0	100%	17.68%	96.88%	17.11%
-5	100%	22.38%	93.10%	18.44%
-10	100%	38.14%	90.32%	27.24%
-20	100%	124.24%	80.42%	72.87%

Normalized to result of the GSC method.

Table 4-1 shows that the GSC-SN method has better performance under low SNR conditions. The GSC-SN method could reduce the residual noise component effectively when SNR is below 5dB. Because when the SNR is high, the SN method will only reduce trivial noise that may not influence the average speed of the adaptive filter.

Table 4-2 indicates that the GSC-CC method could reduce the peak value of the enhanced signal distortion efficiently under most SNR conditions, except the case that the noise is too heavy that the speech energy barely affects the estimation of the residual

noise.

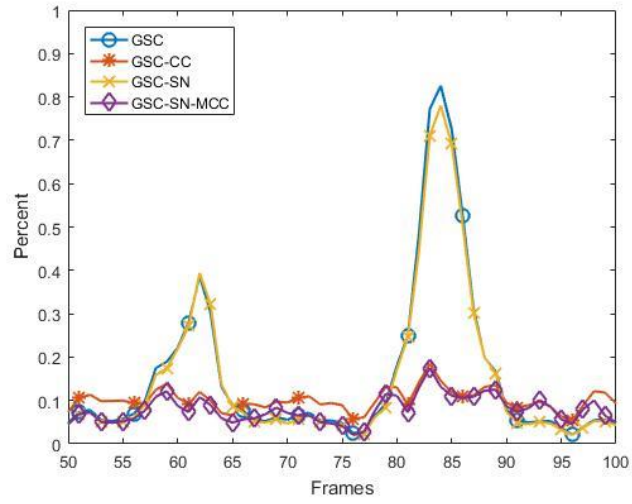
When the SNR is as high as 20 dB, the average distortion-ratio of the GSC-CC method is better compared to the GSC-SN-MCC. It is because that the noise level is extremely low in this case, the denoising effect of the SN method of GSC-SN-MCC method is less than the over subtraction effect caused by the SN method itself, even though the cross-correlation coefficient is used to control this problem.

The proposed GSC-SN-MCC method is a combination of the above two methods. The tables show that the proposed method can work efficiently under both high and low SNR conditions. The performance trend of the GSC-SN-MCC method is a parabolic curve. When the SNR is too high or too low, the performance of the proposed method would be decreased. And the best performance of the proposed method occurs when the energy of speech is almost equal to the energy of noise.

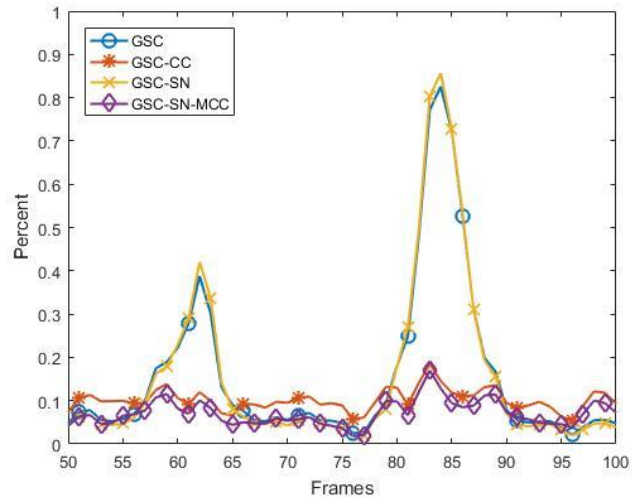
4.2.2 Effect of the inaccurate estimated noise direction

The simulation calculation was conducted based on the previous experiment data to show the influence on the performance of the proposed method when the estimated noise direction was not accurate. Therefore, the incident direction of the desired signal was 0° and the actual noise direction was 90° in this subsection. The SNR condition was performed as -5dB. Twelve angles were chosen as the inaccurately estimated noise directions, which were distributed uniformly on a circle. The meaning of the coordinate axes in the figures and the calculation equations of the result were the same as in the previous subsection.

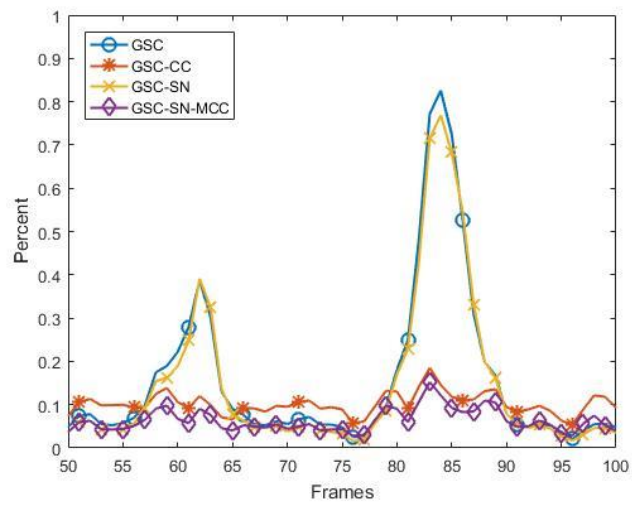
Figure 4-4 presents the distortion-noise ratio of the frames (from 50th to 100th frame) with different methods when the estimated noise direction is -60° , -30° , 0° , $+30^\circ$, $+60^\circ$, $+90^\circ$, $+120^\circ$, $+150^\circ$, $+180^\circ$, -150° , -120° and -90° .



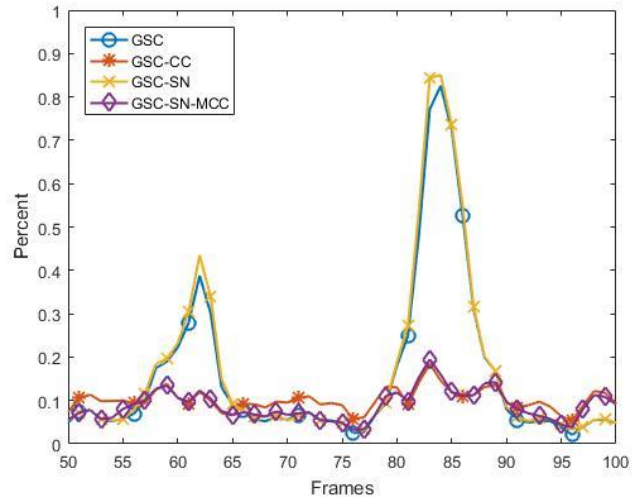
(a) estimated noise direction is -60°



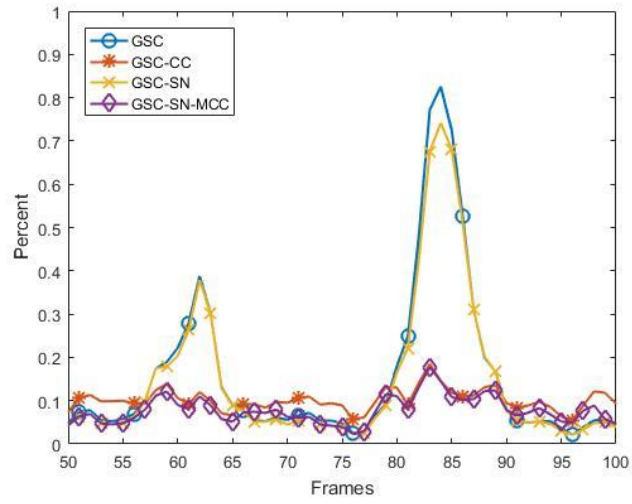
(b) estimated noise direction is -30°



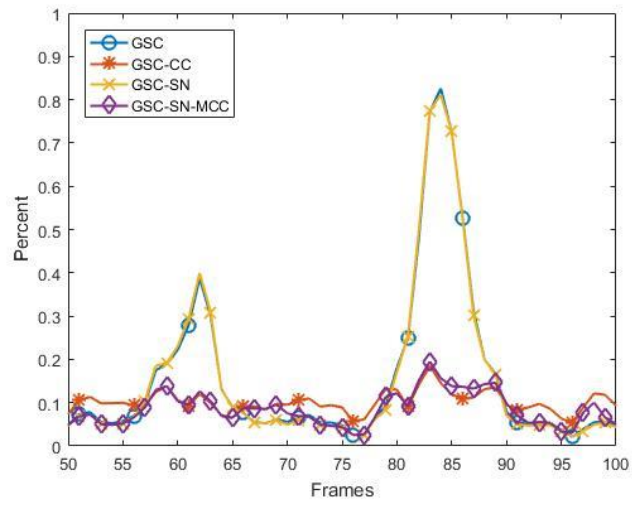
(c) estimated noise direction is 0°



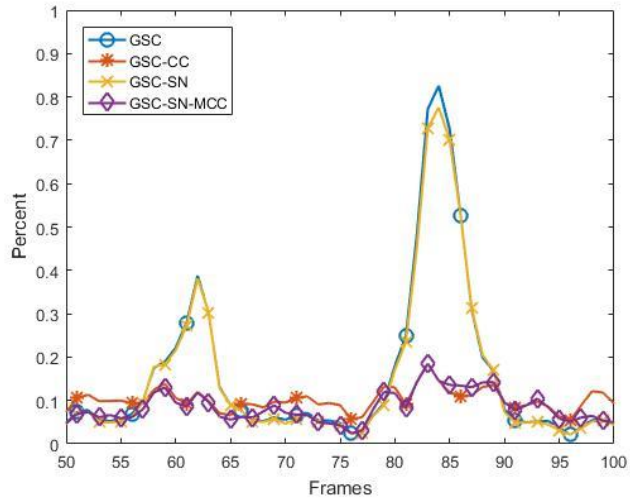
(d) estimated noise direction is $+30^\circ$



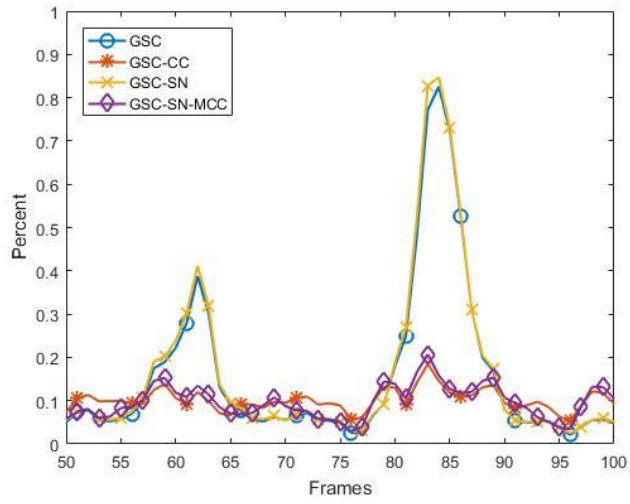
(e) estimated noise direction is $+60^\circ$



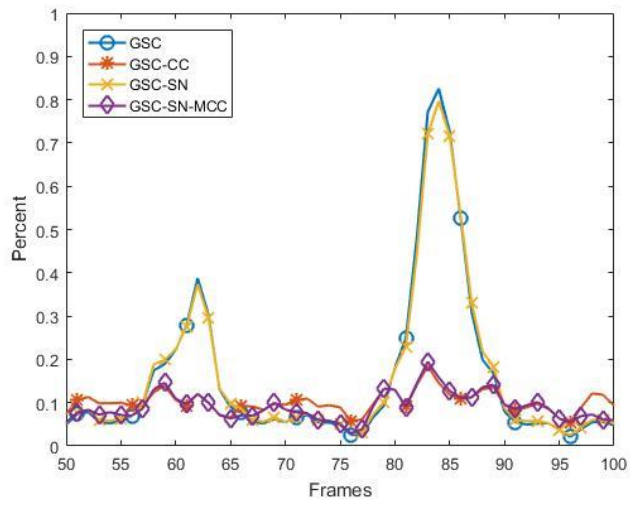
(f) estimated noise direction is $+90^\circ$



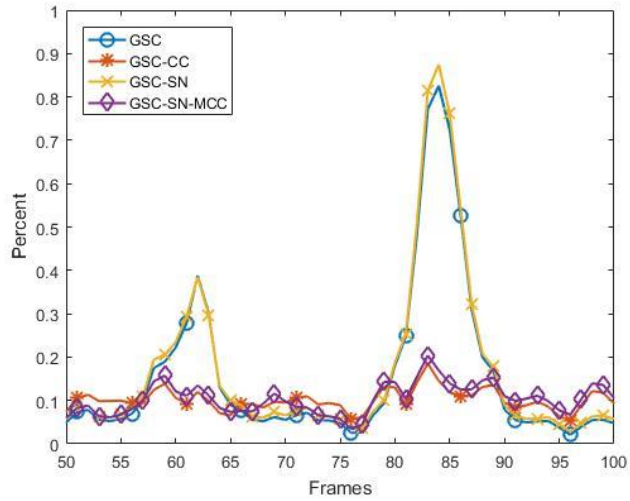
(g) estimated noise direction is +120°



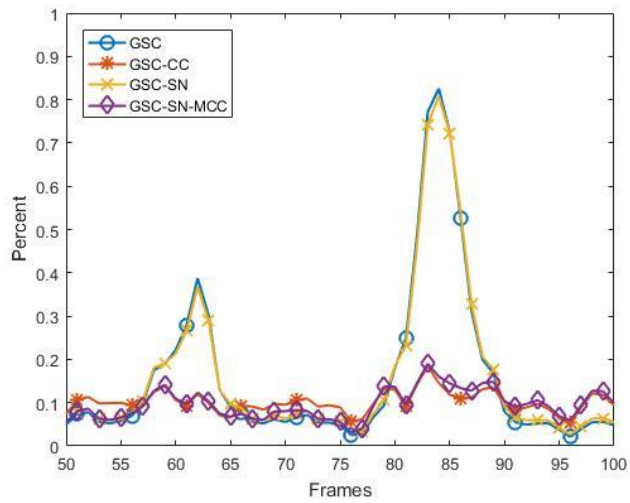
(h) estimated noise direction is +150°



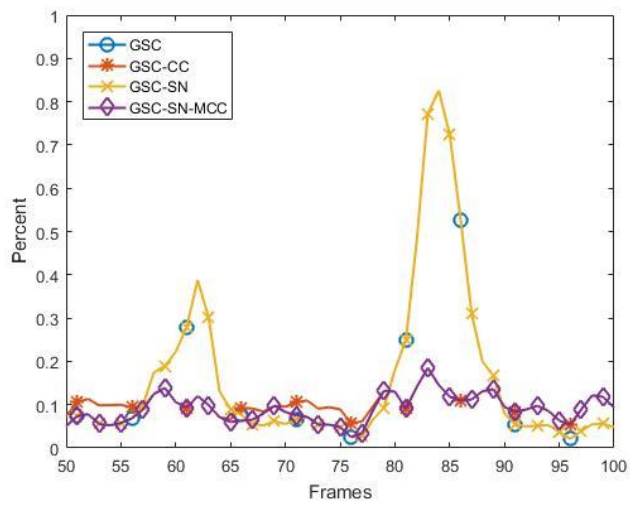
(i) estimated noise direction is +180°



(j) estimated noise direction is -150°



(k) estimated noise direction is -120°



(l) estimated noise direction is -90°

Figure 4-4. Distortion-noise ratio of frames (from 50th to 100th frame) by different methods when the estimated noise direction is different.

The figures indicate even the estimated noise direction is far away from the actual noise direction, the distortion of the proposed method is still lower and flatter than the conventional GSC method, just a little higher than the GSC-MCC method. The figures demonstrate that no matter the error of the estimated noise direction, the proposed method still could effectively reduce both the peak value and the average value of the signal distortion.

Table 4-3 and 4-4 compare the performance of the different methods in various estimated noise directions. Table 4-3 lists the average of the distortion-noise ratio when the estimated noise direction is different and Table 4-4 presents the distortion value of the 83rd frame (maximum distortion frame) under different estimated noise direction conditions.

Table 4-3. Average of the normalized distortion-noise ratio by different methods in various estimated noise directions.

Estimated noise direction/degree	GSC	GSC-CC	GSC-SN	GSC-SN-MCC
-60	100%	105.34%	112.55%	100.87%
-30	100%	105.34%	106.33%	91.93%
0	100%	105.34%	100%	88.07%
30	100%	105.34%	92.92%	76.27%
60	100%	105.34%	92.80%	68.32%
90	100%	105.34%	83.98%	59.01%
120	100%	105.34%	95.71%	80.68%
150	100%	105.34%	90.56%	76.15%
180	100%	105.34%	97.76%	82.98%
-150	100%	105.34%	94.29%	79.13%
-120	100%	105.34%	105.59%	93.54%
-90	100%	105.34%	102.98%	88.94%

Normalized to result of the GSC method.

Table 4-4. Normalized distortion ratio of the 83rd frame by different methods in various estimated noise directions.

Estimated noise direction/degree	GSC	GSC-CC	GSC-SN	GSC-SN-MCC
-60	100%	22.38%	105.90%	24.68%
-30	100%	22.38%	97.81%	23.04%
0	100%	22.38%	100%	22.38%
30	100%	22.38%	94.42%	21.15%
60	100%	22.38%	103.73%	20.55%
90	100%	22.38%	93.10%	18.44%
120	100%	22.38%	103.00%	23.60%
150	100%	22.38%	89.81%	21.52%
180	100%	22.38%	98.23%	23.62%
-150	100%	22.38%	93.95%	22.44%
-120	100%	22.38%	102.58%	25.06%
-90	100%	22.38%	96.50%	23.57%

Normalized to result of the GSC method.

Table 4-3 shows that the performance of the GSC-CC method is not related to the estimated noise direction because the estimated noise direction information is not used in the GSC-CC method. The performance of the GSC-SN method would be decreased when the estimated noise direction is far away from the actual noise direction. When the estimated noise direction is identical to the actual noise signal direction, the GSC-SN method could achieve the best denoising performance. If the estimated noise direction is identical to the desired signal direction, the GSC-SN method would reduce to the conventional GSC method.

The trend of the performance of the GSC-SN-MCC method is similar to the GSC-SN method. Better performance of the proposed method could be attained if the estimated noise direction is closer to the actual noise direction. Even when the estimated noise direction is the opposite of the actual noise direction, the performance of the

proposed method is still not worse than the conventional GSC method.

Table 4-4 presents the GSC-CC method that could efficiently reduce the peak value of the signal distortion, which is independent of the estimated noise direction. Because of the high SNR in the 83rd frame, the GSC-SN method barely affects the residual noise component of the adaptive filter. Even it is a little higher than the conventional GSC method in several directions. The ability of the GSC-SN-MCC method to decrease the peak of the enhanced signal distortion is similar to the GSC-CC method. The performance of the proposed method fluctuates with the performance of the GSC-SN method while being affected by the SN method. Overall, the proposed method could reduce the peak value of the signal distortion efficiently no matter the error of the estimated noise direction.

The tables show that the proposed method could work effectively when the estimated noise direction is near the actual noise direction. When the estimated noise direction is far from the actual noise direction, the proposed method could still reduce the peak value of the enhanced signal distortion efficiently, and in that case, the average of the enhanced signal distortion using the proposed method is similar to the conventional GSC method.

4.3 Conclusion

This chapter presented the setup of the experiment in detail. The experiment was performed in an anechoic chamber to validate the proposed method in various SNR conditions. The experimental result demonstrated that the GSC-MCC method could reduce the damage to the desired signal of the enhanced signal efficiently, except in the case that the SNR was too low so the original damage was too minor. The GSC-SN

method could work more effectively when the SNR was lower. The proposed method could reduce both kinds of noise effectively in various SNR conditions. The performance of the proposed method was like a parabolic curve, the best performance would be obtained when the energy of the speech and noise was almost equivalent, like the SNR was +5dB.

Furthermore, the simulated calculation with inaccurate estimated noise directions was conducted based on the experiment data to inspect the robustness of the proposed method to the error of the estimated noise direction. The figures of the simulation result with the inaccurate estimated noise directions were displayed. The simulation result showed that the performance of the GSC-MCC method was not related because the estimated noise direction was not used in this method. The ability of the GSC-SN method would be lost when the estimated noise direction was far away from the actual noise direction. The denoising effect of the GSC-MCC-SN method would be declined when the estimated noise direction was inaccurate. However, even the estimated noise direction was opposite to the actual noise direction, the residual noise component was still similar to the original residual noise component, which implied the feasibility of the proposed method in practical cases.

Chapter 5 Conclusions

In this dissertation, Chapter 1 presented the research background and related academic achievement about the microphone array first. The microphone array method is proposed to filter the signal with the spatial direction. The fixed beamforming is the basic method to delay and sum the signals by adding the appropriate delay to different microphones. The setup of the microphone array could influence the performance of the microphone array. Then the modified methods, like LCMV, MVDR, GSC, etc., are out forward to improve the performance of the microphone array without modifying the hardware of the microphone array.

The GSC method includes three parts and can be optimized respectively based on different purposes. The adaptive filter part of the GSC method is used to estimate the noise component in the output signal of the GSC method and try to reduce it. The common adaptive algorithm of the GSC method is the LMS algorithm. The modified LMS algorithm could improve the performance of the adaptive filter in certain conditions.

The distortion of the enhanced signal could be divided into two parts: the residual noise and the damage of the desired signal. This dissertation tried to preserve the desired part of the enhanced signal from damage and reduce the residual noise more.

Chapter 2 introduced the related foundation knowledge about the microphone array. The formula of the beamforming pattern of the UCA was derived first. Then the influence factors of the beamforming pattern were displayed. The radius of the microphone array would affect the spatial resolution of the main lobe of the microphone array. The larger diameter would bring the narrower main lobe. The more microphones

in the microphone array would make the beamforming pattern more like the continuous microphone array that avoids the grating lobe being generated.

The Pearson correlation coefficient shows the linear relationship of two variables. And the rank correlation coefficient, like the Spearman correlation coefficient and Kendall rank correlation coefficient, could show the variation trend of two variables concordant or discordant.

The LMS algorithm is a common adaptive filter algorithm to estimate the transfer function between the reference noise signal and the received noise signal. The LMS algorithm could be modified to adapt to the different situations, like the FxLMS algorithm that considers the transfer function of the secondary path to improve the convergency performance of the LMS algorithm.

Chapter 3 proposed the modified GSC method. The damage of the desired signal in the enhanced signal was displayed on different SNR conditions. The figures illustrated that serious damage would occur when the SNR was high. The GSC-MCC method was proposed by adding the cross-correlation coefficient between the canceling signal and the error signal to the filter weight update path to control the damage of the enhanced desired signal when the energy of the desired signal was increased suddenly.

The sidelobe neutralization method was derived and demonstrated that it could reduce the noise component in the output signal of the beamforming method. Even the estimated noise direction was inaccurate, the amount of noise component processed by the SN method was still similar to the original noise component. Thus, the SN method was combined with the conventional GSC method (refer to the GSC-SN method) to improve the performance of the GSC method by reducing the noise component in the beamforming output signal of the GSC method.

Finally, the GSC-SN-MCC method was proposed by combining the GSC-MCC method and GSC-SN method to synthesize the advantages of them. The performed steps of the proposed method were depicted in detail.

Chapter 4 described the implementation of the experiment and analyzed the experimental data. The experiment was performed in the anechoic chamber and the experiment setup was presented. The experimental data showed that on different SNR conditions, the GSC-MCC method could decrease the damage of the desired signal of the enhanced signal significantly. The performance would be high in the high SNR condition. The GSC-SN method could reduce the residual noise component effectively. The performance of the GSC-SN method would be better when the SNR was low.

The GSC-SN-MCC method could reduce both the damage of the desired signal of the enhanced signal and the residual noise component effectivity under various SNR conditions (the SNR range is -20~20dB). The performance of the proposed method was almost a parabolic curve as the SNR increased. The performance of the proposed method was better when the SNR condition was almost +5dB.

The simulated calculation was conducted to reveal the influence of the inaccurate estimated noise directions on the denoising performance of the proposed method. The simulated result showed that the performance of the GSC-MCC method was not related to the estimated noise direction because the noise direction information was not used in the GSC-MCC method. Whereas the ability to reduce the residual noise component of the GSC-SN method would be declined with the inaccuracy estimated noise direction.

The calculation results showed that the proposed method could work effectively when the estimated noise direction was near the actual noise direction. The performance of the proposed method would be like a parabolic curve as the estimated noise direction

became far away from the actual noise direction. When the estimated noise direction was correct, the proposed method could get the best performance. Even if the estimated noise direction was far from the actual noise direction, the peak of the enhanced signal distortion would be decreased significantly, while the average of the enhanced signal distortion would still be similar to the conventional GSC method, which indicated the feasibility of the proposed method in practice.

Bibliography

- [1]. S. Boll, "Suppression of acoustic noise in speech using spectral subtraction," in IEEE Transactions on Acoustics, Speech, and Signal Processing, vol. 27, no. 2, pp. 113-120, April 1979, doi: 10.1109/TASSP.1979.1163209.
- [2]. Wiener, Norbert (1949). Extrapolation, Interpolation, and Smoothing of Stationary Time Series. New York: Wiley. ISBN 978-0-262-73005-1.
- [3]. Flanagan, J. L.; Johnston J.D.; Zahn R.; Elko G.W. Computer-steered microphone arrays for sound transduction in large rooms. J. Acoust. Soc. Am. 1985, 78(5):1508-1518.
- [4]. <http://labbookpages.co.uk/audio/beamforming/delaySum.html>
- [5]. Ning Jun. Study on Beamforming based Speech Separation and Acoustic Echo Cancellation for Microphone Array. Dalian University of Technology. 2018.
- [6]. Capon J. High-resolution frequency-wavenumber spectrum analysis. Proc IEEE 1969, 57(8):1408-1418.
- [7]. Habets E A P, Benesty J, Cohen I, et al. New insights into the MVDR beamformer in room acoustics[J]. IEEE Transactions on Audio, Speech and Language Processing, 2010, 18(1): 158-170.
- [8]. Pan C, Chen J. Performance study of the MVDR beamformer as a function of the source incidence angle[J]. IEEE/ACM Transactions on Audio, Speech and Language Processing, 2014, 22(1): 67-79.
- [9]. Tavakoli V M, Jensen J R, Christensen M G, et al. A framework for speech enhancement with ad hoc microphone arrays[J]. IEEE/ACM Transactions on Audio, Speech and Language Processing, 2016, 24(6): 1038-1051.

- [10]. E. A. P. Habets, J. Benesty and P. A. Naylor, "A Speech Distortion and Interference Rejection Constraint Beamformer," in *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 3, pp. 854-867, March 2012, doi: 10.1109/TASL.2011.2166958.
- [11]. Frost O. L. An algorithm for linearly constrained adaptive array processing. *Proc. IEEE* 1972, 60(8):926-936.
- [12]. Markovich S, Gannot S, Cohen I. Multichannel eigenspace beamforming in a reverberant noisy environment with multiple interfering speech signals[J]. *IEEE Transactions on Audio, Speech, and Language Processing*, 2009, 17(6): 1071-1086.
- [13]. Norholm S M, Jensen J R, Christensen M G. Enhancement and noise statistics estimation for non-stationary voiced speech[J]. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 2016, 24(4): 645-658.
- [14]. Griffiths L.; Jim C. W. An alternative approach to linearly constrained adaptive beamforming. *IEEE Trans. Antennas Propag.* 1982, 30(1):27-34.
- [15]. Hoshuyama O.; Sugiyama A.; Hirano A. A robust adaptive beamformer for microphone arrays with a blocking matrix using constrained adaptive filters. *IEEE T Signal Proces.* 1999, vol. 47, no. 10, pp. 2677-2684.
- [16]. Lee Y.; Wu W. R. A robust adaptive generalized sidelobe canceller with decision feedback. *IEEE Trans. Antennas Propag.* 2005, vol. 53, no. 11, pp. 3822-3832.
- [17]. Khan Z. U.; Naveed A.; Qureshi I. M.; Zaman F. Comparison of Adaptive Beamforming Algorithms Robust Against Directional of Arrival Mismatch. *Journal of Space Technology*, 2012, vol 1, No. 1.

- [18]. Gannot S.; Burshtein D.; Weinstein E. Signal enhancement using beamforming and nonstationarity with applications to speech. *IEEE T Signal Proces.* vol. 49, no. 8, pp. 1614-1626, Aug. 2001.
- [19]. Reuven G.; Gannot S.; Cohen L. Joint noise reduction and acoustic echo cancellation using the transfer-function generalized sidelobe canceller, *Speech Commun.* Volume 49, Issues 7–8, 2007, Pages 623-635.
- [20]. Reuven G.; Gannot S.; Cohen I. Dual-Source Transfer-Function Generalized Sidelobe Canceller. *IEEE Speech Audio Process.* vol. 16, no. 4, pp. 711-727, May 2008.
- [21]. Rombouts G.; Spriet A.; Moonen M. Generalized sidelobe canceller based combined acoustic feedback- and noise cancellation, *Signal Process.* Volume 88, Issue 3, 2008, Pages 571-581.
- [22]. Krueger A.; Warsitz E.; Haeb-Umbach R. Speech Enhancement With a GSC-Like Structure Employing Eigenvector-Based Transfer Function Ratios Estimation. *IEEE/ACM Trans. Audio, Speech, Language Process* vol. 19, no. 1, pp. 206-219, Jan. 2011.
- [23]. Glentis, G. O. Implementation of adaptive generalized sidelobe cancellers using efficient complex valued arithmetic. *Int. J. Appl. Math. Comput. Sci.* 13 (2003): 549-566.
- [24]. Ali R.; Bernardi G.; Waterschoot T. V.; Moonen M. Methods of Extending a Generalized Sidelobe Canceller With External Microphones. *IEEE/ACM Trans. Audio, Speech, Language Process* vol. 27, no. 9, pp. 1349-1364, Sept. 2019.

- [25]. Kim S M, Kim H K. Multi-Microphone Target Signal Enhancement Using Generalized Sidelobe Canceller Controlled by Phase Error Filter[J]. IEEE Sensors Journal, 2016, 16(21): 7566-7567.
- [26]. Clarkson P. M. Optimal and Adaptive Signal Processing; CRC Press, Boca Raton, FL, 1993.
- [27]. Widrow B.; Stearns S. D. Adaptive Signal Processing; Prentice-Hall, Englewood Cliffs, NJ, 1985.
- [28]. Widrow B.; Hoff M.E. Adaptive Switching Circuits; IRE WESCON Convention Record, 4:96-104, August 1960.
- [29]. 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.
- [30]. Zhao S. D., Han H. L., Shi W. X., Generation on mechanism and characteristics of intermittent exhaust noise. Acad J Xian Jiaotong Univ (Chinese ed.), 1999, 33(3), 95–9.
- [31]. Longman RW, Iterative learning control and repetitive control for engineering practice, Int J Control 2000, 73(10), 930–54.
- [32]. Nagumo J. I.; Noda A. A learning method for system identification. IEEE Trans. Autom. Control vol. AC-12, pp. 282-287, June 1967.
- [33]. Shan T. J.; Kailath T. Adaptive algorithms with an automatic gain control feature. IEEE Trans. Circuits Syst. vol. 35, no. 1, pp. 122-127, Jan. 1988.
- [34]. Gitlin R. D.; Meadors H. C.; Weinstein S. B. The tap-leakage algorithm: An algorithm for the stable operation of a digitally implemented, fractionally spaced

- adaptive equalizer. *The Bell System Technical Journal*, vol. 61, no. 8, pp. 1817-1839, Oct. 1982.
- [35]. Gannot S.; Cohen I. Speech enhancement based on the general transfer function GSC and postfiltering. *IEEE Speech Audio Process.* vol. 12, no. 6, pp. 561-571, Nov. 2004.
- [36]. Cohen I.; Berdugo B. Microphone array post-filtering for non-stationary noise suppression. *2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Orlando, FL, USA, 2002, pp. I-901-I-904.
- [37]. Cohen I. Analysis of two-channel generalized sidelobe canceller (GSC) with post-filtering. *IEEE Speech Audio Process.* vol. 11, no. 6, pp. 684-699, Nov. 2003.
- [38]. Asano F.; Hayamizu S. Speech enhancement using CSS-based array processing. *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Munich, Germany, 1997, pp. 1191-1194 vol.2.
- [39]. Doclo S.; Moonen M. SVD-based optimal filtering with applications to noise reduction in speech signals. *Proceedings of the 1999 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics. WASPAA'99*, New Paltz, NY, USA, 1999, pp. 143-146.
- [40]. Doclo S.; Moonen M. Multimicrophone noise reduction using recursive GSVD-based optimal filtering with ANC post-processing stage. *IEEE Speech Audio Process.* vol. 13, no. 1, pp. 53-69, Jan. 2005.
- [41]. Tong L.; Liu R. W.; Soon V. C.; Huang Y. F. Indeterminacy and identifiability of blind identification. *IEEE Trans. Circuits Syst.* vol. 38, no. 5, pp. 499-509, May 1991.

- [42]. Common P. Independent component analysis, A new concept. *Signal Process.* 1994, 36(3): 287-314.
- [43]. Dong H. Y.; Lee C. M. Speech intelligibility improvement in noisy reverberant environments based on speech enhancement and inverse filtering. *EURASIP J Audio Spee.* vol. 3, 2018.
- [44]. M. Y. Appiah, M. Sasikath, R. Makrickaite, M. Gusaite, “Robust Voice Activity Detection and Noise Reduction Mechanism”, Institute of Electronics Systems, Aalborg University.
- [45]. X. L. Liu, Y. Liang, Y. H. Lou, H. Li, B. S. Shan, Noise-Robust Voice Activity Detector Based on Hidden Semi-Markov Models, *Proc. ICPR'10*, 81–84.
- [46]. Kuo S. M.; Morgan D. R. Adaptive active noise control systems: Algorithms and digital signal processing (DSP) imple-mentations. *Digital Signal Processing Technology: A Critical Review.* Vol. 10279. International Society for Optics and Photonics, 1995. pp. 32-33.
- [47]. Bitzer J.; Simmer K. U.; Kammeyer K. D. Theoretical noise reduction limits of the generalized sidelobe canceller (GSC) for speech enhancement. 1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. *Proceedings. ICASSP99 (Cat. No.99CH36258)*, 1999, pp. 2965-2968 vol.5.
- [48]. Van Veen, B. D.; Buckley, K. M. (1988). "Beamforming: A versatile approach to spatial filtering". *IEEE ASSP Magazine.* 5 (2): 4.doi:10.1109/53.665. S2CID 22880273.
- [49]. <https://en.wikipedia.org/wiki/Beamforming>
- [50]. R. P. Dougherty. Spiral-shaped array for broadband imaging [P]. US, Utility Patent, US5838284, 1998-11-17.

- [51]. J. R Underbrink. Circularly symmetric, zero redundancy, planar array having broad frequency range applications [P]. US, Utility Patent, US6205224, 2001-3-20.
- [52]. J. C. S. Allen, W. K. Blake, R. P. Dougherty, et al. Aeroacoustic measurements[M]. Berlin: Springer Science & Business Media, 2002: 98-217.
- [53]. Yan S. F. Optimal array signal processing: Modal array processing and direction-of-arrival estimation; Science press, Beijing, 2018, pp. 13-16.
- [54]. S. Gannot and I. Cohen, "Speech enhancement based on the general transfer function GSC and postfiltering," in IEEE Transactions on Speech and Audio Processing, vol. 12, no. 6, pp. 561-571, Nov. 2004, doi: 10.1109/TSA.2004.834599.
- [55]. <https://www.investopedia.com/terms/c/correlationcoefficient.asp>
- [56]. Laerd Statistics. "Pearson Product-Moment Correlation." Accessed Oct. 5, 2021.
- [57]. Kent State University. "SPSS Tutorials: Pearson Correlation." Accessed Oct. 5, 2021.
- [58]. "Correlation Coefficient: Simple Definition, Formula, Easy Steps". Statistics How To.
- [59]. <https://www.verywellmind.com/what-is-correlation-2794986>
- [60]. List of Probability and Statistics Symbols". Math Vault. 26 April 2020. Retrieved 22 August 2020.
- [61]. <https://baike.baidu.com/item/B0/3316537>.
- [62]. Xu Xiaoyuan; Yan Zheng; Feng Donghan; Wang Yi; Cao Lu. Probabilistic power flow calculation method based on rank correlation coefficient of input variables[J]. Automation of Electric Power Systems, 2014, 12.
- [63]. https://en.wikipedia.org/wiki/Spearman%27s_rank_correlation_coefficient.

- [64]. Lehman, Ann (2005). *Jmp For Basic Univariate And Multivariate Statistics: A Step-by-step Guide*. Cary, NC: SAS Press. p. 123. ISBN 978-1-59047-576-8.
- [65]. Myers, Jerome L.; Well, Arnold D. (2003). *Research Design and Statistical Analysis* (2nd ed.). Lawrence Erlbaum. pp. 508. ISBN 978-0-8058-4037-7.
- [66]. https://en.wikipedia.org/wiki/Kendall_rank_correlation_coefficient.
- [67]. Kendall, M. (1938). "A New Measure of Rank Correlation". *Biometrika*. 30 (1–2): 81–89. doi:10.1093/biomet/30.1-2.81. JSTOR 2332226.
- [68]. Kruskal, W. H. (1958). "Ordinal Measures of Association". *Journal of the American Statistical Association*. 53 (284): 814–861.
- [69]. https://www.statsdirect.com/help/nonparametric_methods/kendall_correlation.htm
- [70]. <https://towardsdatascience.com/kendall-rank-correlation-explained-dee01d99c535>
- [71]. 72 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.
- [72]. Zue, V.; Seneff, S.; Glass J. Speech database development at MIT: TIMIT and beyond. *Speech Commun.* 9(4), 351–356, 1990.
- [73]. Varga, A.; Steeneken M. H. J. Assessment for automatic speech recognition: II. NOISEX-92: a database and an experiment to study the effect of additive noise on speech recognition systems. *Speech Commun.* 12(3), 247–251,1993.