國 立 交 通 大 學

電信工程學系

碩 士 論 文

利用共極串音消除器建立強健
三度空間音效

Robust 3D Sound based on Common-pole
Crosstalk Canceller

研究生：黃俊榮

指導教授：謝世福 博士

中 華 民 國 九十五 年 九 月

利用共極串音消除器建立強健三度空間音效

Robust 3D Sound based on Common-pole Crosstalk Canceller

研究生 ：黃俊榮   Student：C. R. Huang

指導教授：謝世福   Advisor：S. F. Hsieh

國 立 交 通 大 學
電信工程學系碩士班
碩士論文

中華民國九十五年九月

利用共極串音消除器建立強健三度空間音效

學生：黃俊榮　　　　　指導教授：謝世福

國立交通大學電信工程學系碩士班

摘要

　　在現代多媒體系統中，有越來越多的虛擬實境應用。因此三度空間音效技術便越來越重要。我們將使用雙喇叭來產生三度空間音效。然而，主要的問題是串音(crosstalk)的干擾。我們將使用 FIR 和 IIR 兩種形式的串音消除器(crosstalk canceller)來解決這個問題。在這兩種內，我們都會提出反矩陣(matrix-inverse)或是直接最小平方錯誤(LSE)的方法。在反矩陣方法中，為了避免不穩定性，我們提出最小比例錯誤(ratio error)的方法。在直接 LSE IIR 設計中，為了避免非線性帶來設計上的困難，我們提出共極串音消除器結構。接下來我們將探討另一個問題，串音消除器的強健性。我們知道假如使用一組固定的串音濾波器，在頭可以移動的狀況下，接收到的信號將會和我們想要的信號差很多。因此，我們提出一個利用區域等化觀念的設計方法來降低頭移動造成的影響。最後，我們將用錯誤能量(EN)，串音壓制係數(CSF)和等化改善係數(EIF)來量化串音消除器的效能。

Robust 3D Sound based on Common-pole Crosstalk Canceller

Student : C. R. Huang   Advisor : S. F. Hsieh

Department of Communication Engineering

National Chiao Tung University

Abstract

There are more and more virtual reality applications in the modern multimedia systems. Therefore, the 3D sound technique becomes more important. We will use a loudspeaker pair to generate the 3D sound. However, the most critical problem is the crosstalk disturbance. To overcome this problem, we investigate both FIR and IIR forms of crosstalk cancellers. In both forms, we propose matrix inverse and direct LSE methods to implement the filters. In matrix-inverse method, to avoid the un-stability, we propose to minimize the ratio error. In direct LSE IIR design, to avoid the nonlinearity, the common-pole structure is proposed. Then, we consider another problem, the robustness of the crosstalk canceller. We know that if the crosstalk canceller is fixed, and the received signals may be very different to the signals we want when the head moves. Therefore, we propose a method by using the region-equalized concept to reduce the effect of head movements. Finally, Error energy (EN), crosstalk suppression factor (CSF) and equalization improvement factor (EIF) are used to quantify the performance of crosstalk cancellers.

致謝

感謝謝世福教授的耐心指導，使得本篇論文可以順利的完成。他所強調的物理直覺讓我獲益良多，也對事情的分析有更加一步的認識。再來要感謝我的父母與家人，由於他們的支持與關懷，讓我有信心，尤其是大姐，在我疲倦的時候，幫我減輕壓力，讓我有能量可以繼續下去。再來是感謝我的朋友，懷嘉，俊煒，威諭以及其他許多好友，有他們的加油，讓我順利完成學業。最後是實驗室的夥伴，一起度過快樂的兩年，有大家的一起努力與互相加油打氣，使得論文可以完成。

Contents

摘要   I

English Abstract   II

致謝   III

Contents   IV

List of Tables   VII

List of Figures   XIII

List of Tables

List of Figures

Chapter 1

Introduction

As we know, virtual reality technique can be used to render virtual sound sources in three-dimensional (3-D) space around a listener [1]. Applications for this technique include entertainments, communications, and simulations. For example, a traditional 5-channel system requires five speakers (Left, Center, Right, Surround Left, and Surround Right). However, the measure of the room must be large enough to position each speaker properly. Besides, the system costs a lot of money. By using the virtual reality technique, we can use only two loudspeakers to realize the effect of the 5-channel system. Therefore, a lot of money can be saved, and the measure of the room where the speakers are placed will not be limited. Figure 1.1 shows virtual 5-channel system, where the five loudspeakers are not real and created by virtual reality technique.

Physical Loudspeakers

Center

Left

Right

Surround Left

Surround Right
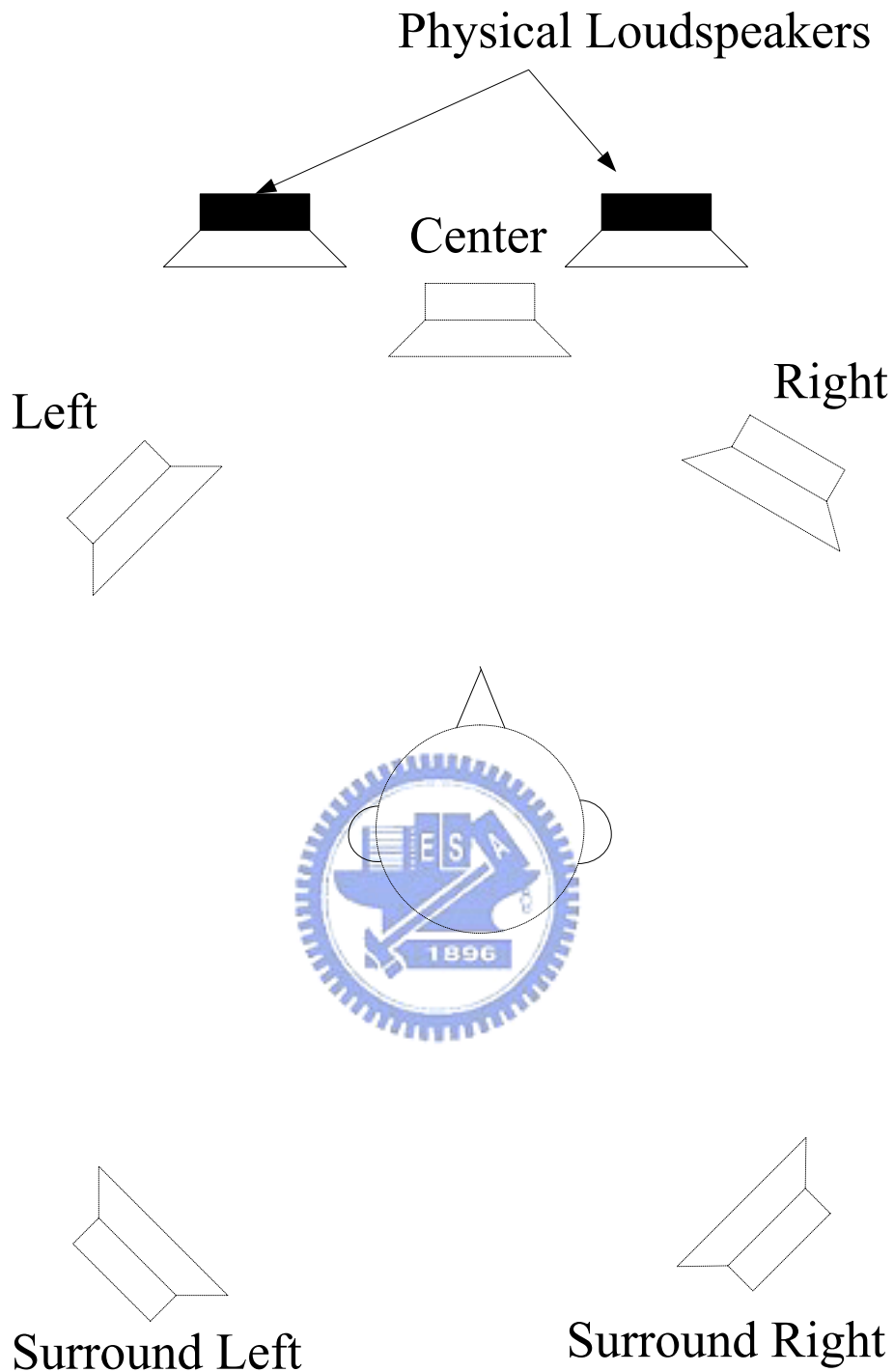
Figure 1.1: Virtual 5-channel system

To realize the 3-D sound system, we must have the sufficient database of the directional cues. It is well known that the principal cues for sound localization, are Interaural Time Difference (ITD) and Interaural Intensity Difference (IID), which represent the time and intensity differences between right and left ears [2]. However it

has some problem to localize using only the two interaural cues, so called cone of confusion [2]. Therefore, we use the head related transfer functions (HRTFs) from a database, based on MIT Media Lab. [3]. HRTFs have the spectral cues, so we can know the sound from any direction [4]. Based on these transfer functions, a virtual sound can be synthesized at any 3-D direction.

The thesis is organized as follows. Chapter 2 will explain the directional cues and how to create a 3-D sound system. Besides, the problems of sound reproduction over headphones and loudspeakers are considered in details. Chapter 3, Chapter 4 and Chapter 5 are the main parts of this thesis, and they focus on the structures and robustness of crosstalk cancellers. In Chapter 6, we will use computer simulations to compare performance of different crosstalk cancellers. In Chapter 7, we will make a conclusion to summarize the results of simulations.

Chapter 2

3-D Sound System

In Chapter 1, we mentioned that a 3-D sound system can be realized by using HRTFs. Therefore, the localization cues of HRTFs will be introduced in Section 2.1 first. We will show the ITD, IID, and spectral cues in a pair of measured HRTFs. In Section 2.2, the synthesis of a directional sound is presented. Section 2.3 will investigate the problems of the 3-D sound system over loudspeakers and the robustness of crosstalk cancellers.

2.1 Sound Localization Cues

HRTFs are frequency-domain functions which have corresponding time-domain functions called head-related impulse responses (HRIRs). HRIRs are the impulse responses measured from some specific position to left and right ears.
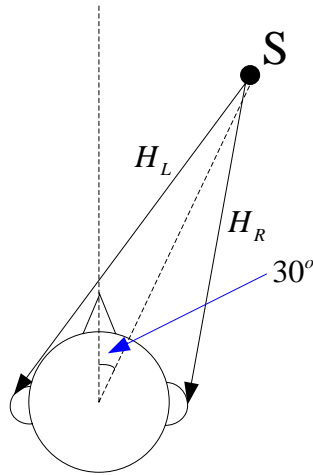
Figure 2.1: A listener with a sound source oriented on azimuth 30° and elevation 0°

A pair of HRTFs are measured in Figure 2.1, where H_R is the transfer function of source S to right ear and H_L is the transfer function to left ear, and their impulse and frequency responses are plotted in Figure 2.2 and Figure 2.3.
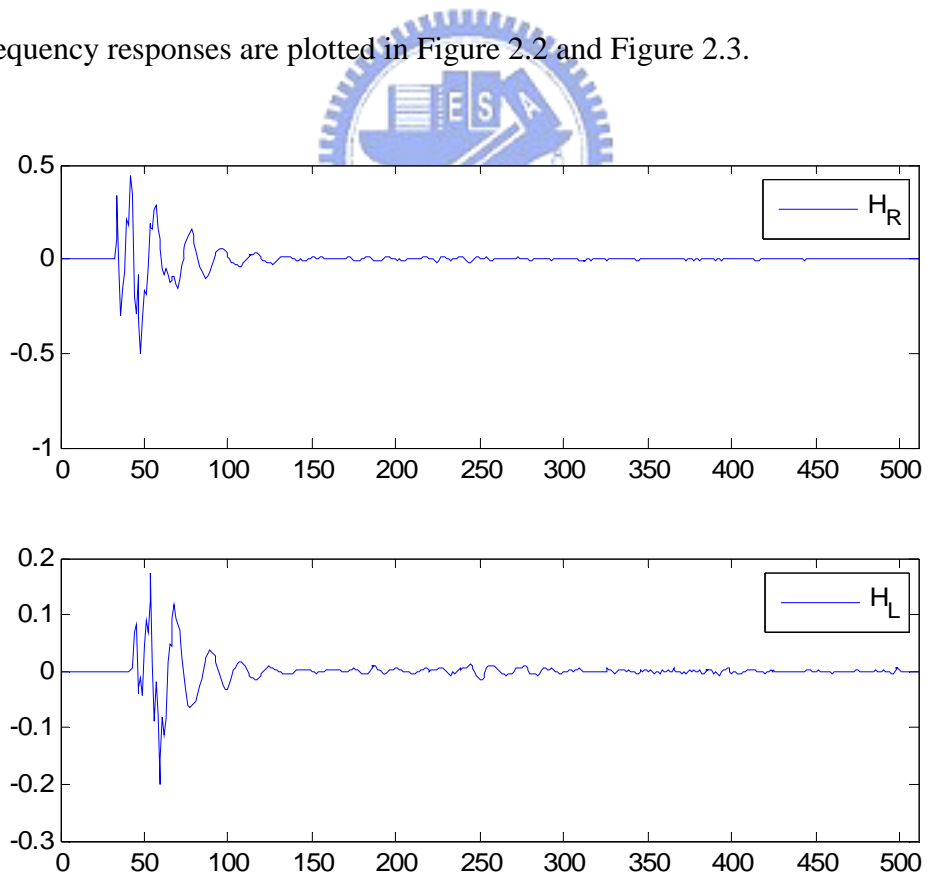
Figure 2.2: Impulse responses of two HRTFs
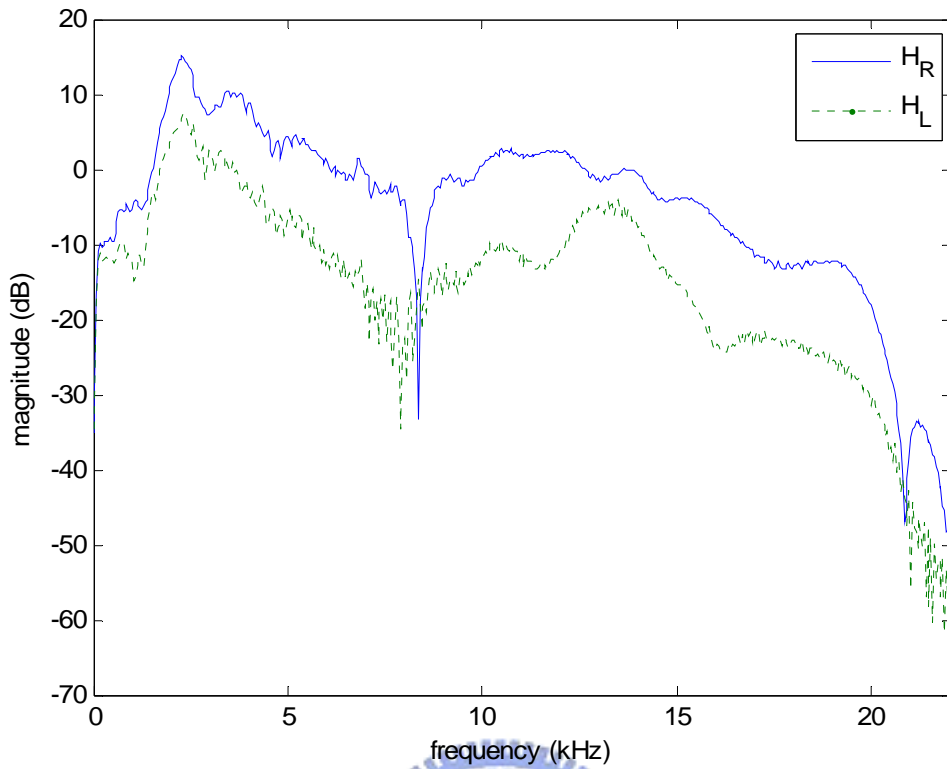
Figure 2.3: Magnitude responses of two HRTFs

In Figure 2.2, we can find the amplitude of vibration in H_R is much larger than that in H_L and H_L has more delays than H_R because of the length difference of two transmission paths. In other words, the observations in Figure 2.2 are the interaural cues of IID and ITD. In Figure 2.3, the high frequency at 8-10 kHz responses have notches caused by the concha reflection [5], and the peaks at 2-3 kHz are caused by the ear canal resonance [6]. These notched and peaks are dependent on the location of the sound source. This result suggests that spectral notches and peaks in HRTFs determine the location of sounds.

2.2 Creation of Virtual Sounds

We can create a virtual loudspeaker by using a pair of HRTFs. The binaural synthesis process is diagrammed in Figure 2.4. When a sound signal S is processed by the digital filters ($H_{\phi R}$ and $H_{\phi L}$, a pair of HRTFs) and played over headphones, the sound localization cues are reproduced and the listener should perceive the sound at the location specified by the pair of HRTFs.
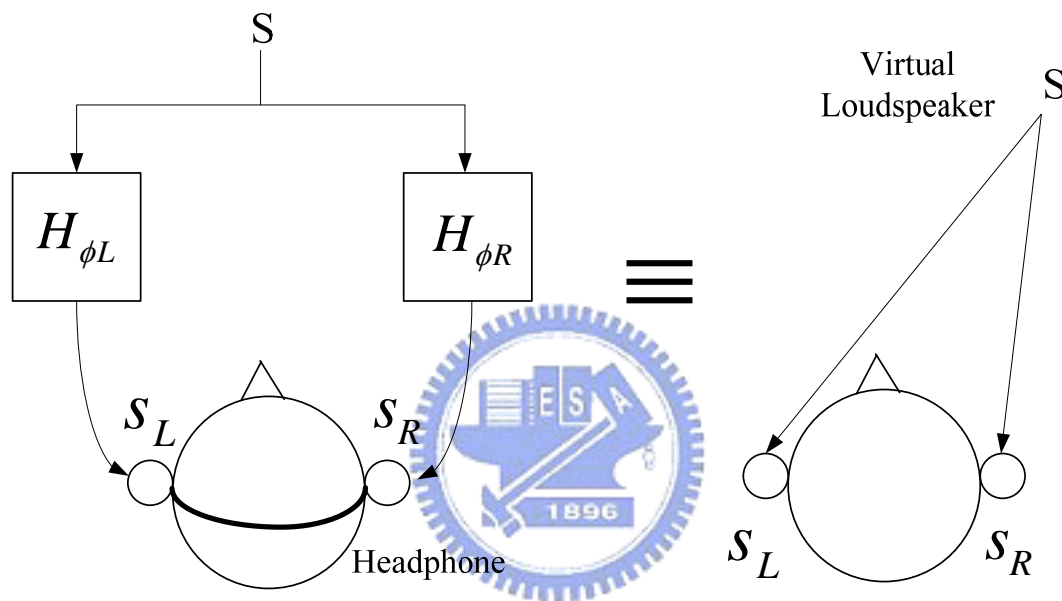
Figure 2.4: Binaural synthesis using headphones

Headphones are often used for 3-D sound audio because they have good channel separation. The directional signals (s_R and s_L) can be received directly. However, there are some drawbacks by using headphone reproduction. It often suffers from in-head localization, and is also cumbersome and inconvenient.

2.3 Virtual Sounds over Loudspeakers

2.3.1 Crosstalk Phenomenon

To avoid the drawbacks of headphones, we would replace the headphones with a pair of loudspeakers. However, the left and right loudspeakers are not coupled directly to the left and right ears. The emitted sound from the right loudspeaker goes to the left ear as well as to the right ear of a listener, and vice versa. This phenomenon is called crosstalk. If two-channel binaural sound is reproduced through a pair of loudspeakers, sound received by the listener can be severely changed from the original sound due to the crosstalk effect.

The effect of crosstalk can be cancelled if binaural signals are filtered before they are sent to the loudspeakers. The process is diagrammed is Figure 2.5.

Figure 2.5: Binaural sound reproduced with crosstalk canceller

2.3.2 Crosstalk Cancellation

Crosstalk cancellation is a technique involving cancelling the crosstalk which transits the head from each speaker to the opposite ear.

The goal of the crosstalk cancellation is that the ear signals e_R and e_L should be same as the binaural signals s_R and s_L. Essentially, the transfer functions from the loudspeakers to the ears form a system transfer function matrix. Using the matrix notations and refer to Figure 2.5, we can write:

$$\begin{bmatrix} e_R(z) \\ e_L(z) \end{bmatrix} = G(z) \begin{bmatrix} y_R(z) \\ y_L(z) \end{bmatrix}, \tag{2.3.1}$$

where $G(z) = \begin{bmatrix} g_{rr}(z) & g_{lr}(z) \\ g_{rl}(z) & g_{ll}(z) \end{bmatrix}$ is the system transfer function matrix and g_{xy} represents the channel impulse response between the x side loudspeaker and y side ear, and $x, y \in \{r, l\}$; y_R and y_L are the input signals of loudspeakers, and

$$\begin{bmatrix} y_R(z) \\ y_L(z) \end{bmatrix} = C(z) \begin{bmatrix} s_R(z) \\ s_L(z) \end{bmatrix} \tag{2.3.2}$$

Therefore, combine Equation (2.3.2) into Equation (2.3.1), we can get

$$\begin{bmatrix} e_R(z) \\ e_L(z) \end{bmatrix} = G(z) \cdot C(z) \begin{bmatrix} s_R(z) \\ s_L(z) \end{bmatrix} \tag{2.3.3}$$

Obviously, the signals of ears are same as the binaural signals if:

$$G(z) \cdot C(z) = I_{2 \times 2}(z) \tag{2.3.4}$$

Thus, our work is to find out the inverse of the system transfer function matrix $G(z)$.

$C(z)$ can be found by inversing $G(z)$ directly as follows:

$$\begin{aligned} C(z) &= G^{-1}(z) \\ &= \begin{bmatrix} g_{ll}(z) & -g_{lr}(z) \\ -g_{lr}(z) & g_{rr}(z) \end{bmatrix} \frac{1}{D(z)}, \end{aligned} \tag{2.3.5}$$

where $D(z) = g_{rr}(z)g_{ll}(z) - g_{rl}(z)g_{lr}(z)$

From Equation (2.3.5), there are some problems. First, each element of $C(z)$ is in fractional form, and the denominator $D(z)$ may be of non-minimum phase so that some poles of each term may lie outside the unit circle in the z-plane. In other words, the elements of $C(z)$ may be unstable. Second, because of the transmission delay, these elements would be non-causal; we should add some delay to make them causal.

There are many structures to implement the crosstalk canceller. The crosstalk canceller was first put into practice by Schroeder and Atal in 1963 [7], and the structure as follows:
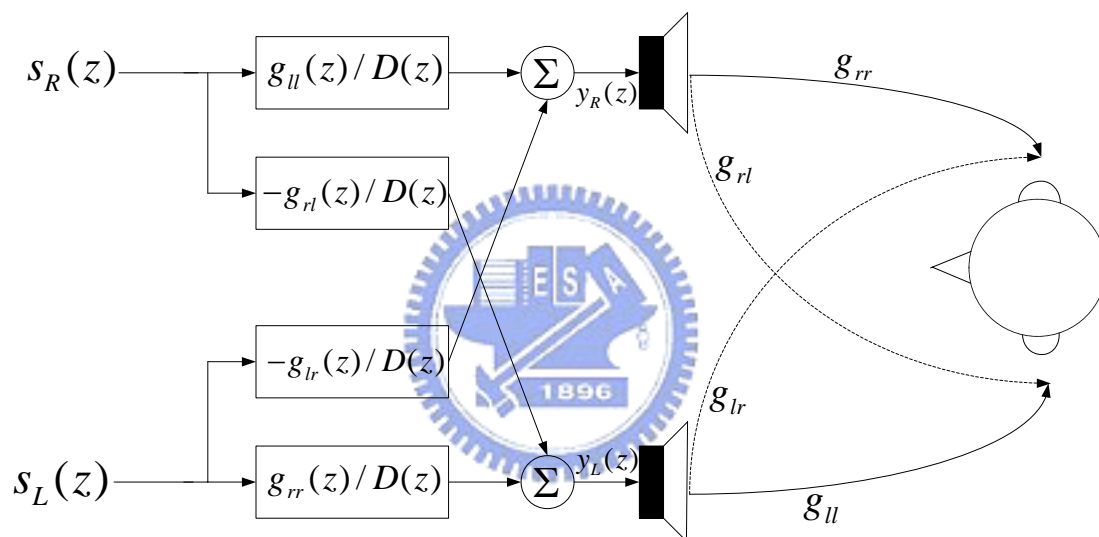
Figure 2.6: Crosstalk canceller described by Schroeder and Atal

When the listening condition is symmetry, i.e. $g_{rr} = g_{ll}$ and $g_{rl} = g_{lr}$, Cooper and Bauck in 1989 proposed the shuffler structure [2]. There are more structures and detailed discussions in [7]. In next chapter, we will use the least square error method to design the crosstalk canceller.

2.3.3 Robustness

So far in this chapter, the design of crosstalk canceller is discussed with single set of the channel transfer function matrix; in other words that means the position of a

listener does not change. In practice, it is impossible. If the position of the listener's head moves, and the crosstalk canceller does not change, the signals to ears may be very different to the original signals.

The question now is how to design a crosstalk canceller that can reduce the effect of head movements. Before designing, there are some researches and analyses about the robustness of the crosstalk cancellers. Ward and Elko et al. [9] show an evaluation of the robustness of crosstalk cancellers for various loudspeaker spacing. They noted a rule-of-thumb for the optimum loudspeaker spacing is given by $d_s = 2 \cdot \lambda \cdot d_H$, where d_H is the distance of the head from the loudspeaker center-line (see Figure 2.7), and λ is the wavelength of operation tone. Therefore, we know that optimum loudspeaker spacing is varied with frequency of sound.
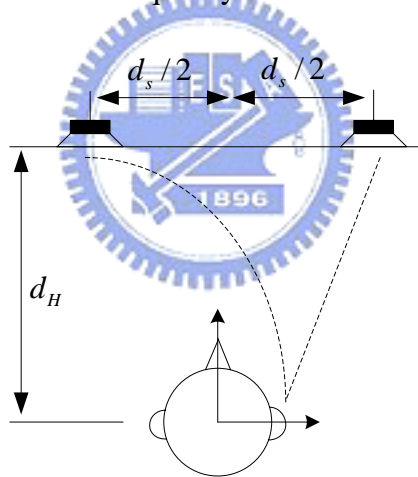
Figure 2.7: Geometry of head and loudspeakers

Later, they show another method of analysis by using the condition number [10]. Because the difference of the transfer function matrix $\Delta \mathbf{G}$ between the fixed head and the moving head can be considered as perturbations of the matrix $\mathbf{G}$. The result indicates that a small incidence angle between two loudspeakers and the head is better on the high frequency (above about 4 kHz) band, and lager incidence angle is better on the lower frequency band. Because different incidence angles are robust to

different frequency band, they analyzed the asymmetric listening condition. A crosstalk cancellation system with three loudspeakers was proposed [11]. An analysis of three loudspeakers by using condition number also shows that the robustness of three loudspeakers is better than that of two loudspeakers [12].

So far, the robustness was increased by increasing the number of loudspeakers. We will propose a new method to increase the robustness with only two loudspeakers.

2.3.4 Crosstalk Canceller Structures

In this section, we will provide an overview for the design flow of the crosstalk canceller in the thesis, and Figure 2.8 is the design flow chart.
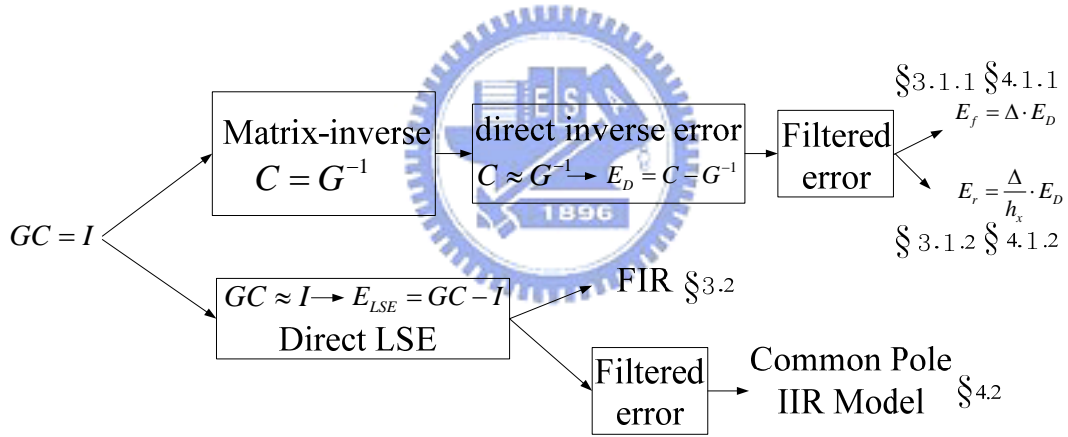
Figure 2.8: Design flow of the crosstalk canceller

Our starting point is Equation (2.3.4), and then the design can be separated into two ways, matrix-inverse and direct LSE methods. In matrix-inverse design, we can find it is hard to handle the direct error, so we handle the filtered error instead of the direct error. In filtered error, we can separate two types filtered by Δ and filtered by ratio term $\dfrac{\Delta}{h_x}$. The type filtered by Δ is used in Section (3.1.1) and Section (4.1.1),

and the type filtered by ratio term $\dfrac{\Delta}{h_x}$ is used in Section (3.1.2) and Section (4.1.2).

In direct LSE design, the FIR form is implemented in Section (3.2), and the IIR form is hard to handle. Therefore, we handle the filtered error instead of the direct LSE IIR error and propose the structure, common pole model in Section (4.2).

Chapter 3

FIR Crosstalk Canceller
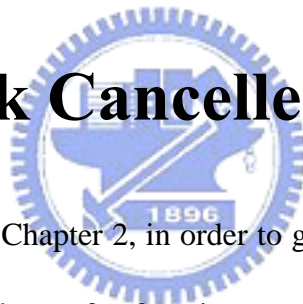
As we have discussed in Chapter 2, in order to generate 3-D sound, $C(z)$ must be the inverse of the channel transfer function matrix $G(z)$. If $G(z)$ is inversed directly, the stability and causality must be considered.

To avoid these problems, we use the matrix inverse criterion and the direct least square error (LSE) criterion to find the filter coefficients of the crosstalk canceller.

3.1 Matrix Inverse Design

3.1.1 Design in Time Domain

Referring to Equation (2.3.5) in Chapter 2, we know each theoretical element of

the crosstalk canceller is given as follows:

$$c_{11}(z) = \frac{g_{ll}(z)}{g_{rr}(z)g_{ll}(z) - g_{rl}(z)g_{lr}(z)} \quad ; \quad c_{21}(z) = \frac{-g_{rl}(z)}{g_{rr}(z)g_{ll}(z) - g_{rl}(z)g_{lr}(z)} \quad ;$$

$$c_{12}(z) = \frac{-g_{lr}(z)}{g_{rr}(z)g_{ll}(z) - g_{rl}(z)g_{lr}(z)} \quad ; \quad c_{22}(z) = \frac{g_{rr}(z)}{g_{rr}(z)g_{ll}(z) - g_{rl}(z)g_{lr}(z)} \quad ;$$

Therefore, we want to find filters of the crosstalk canceller by using these theoretical

solutions, and the block diagram is expressed in Figure 3.1.
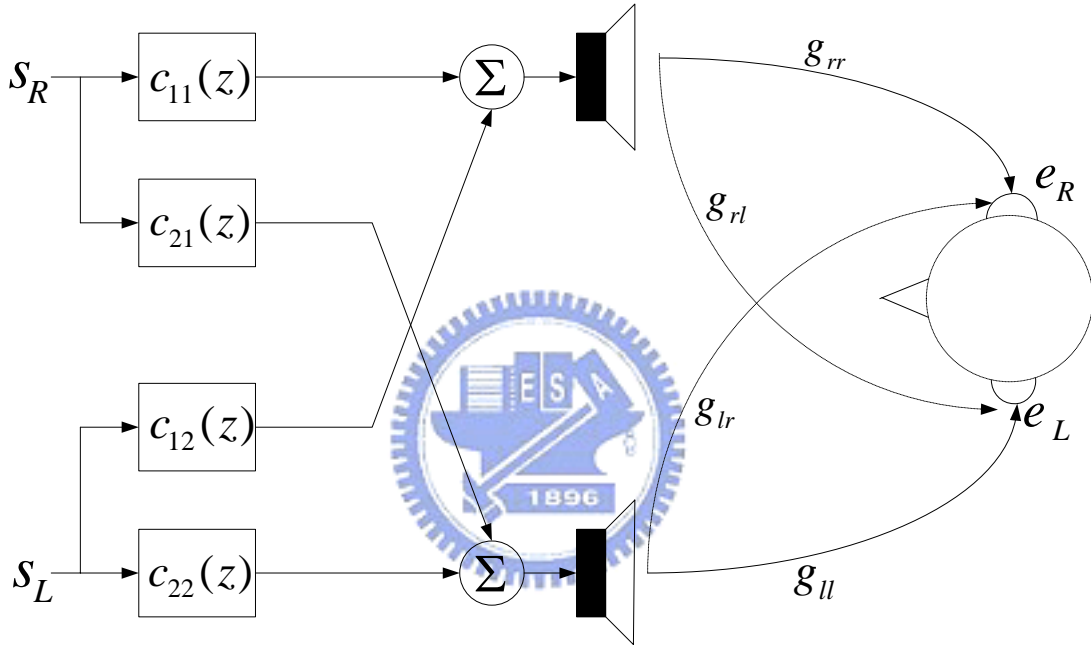
Figure 3.1: Crosstalk canceller in FIR form

From Equation (2.3.5), we know each term is in the same form, so we can

estimate each term in FIR form by using the same algorithm. The following method is

proposed to find each filter. We want to find a FIR filter $c(z)$ so that it can be

approximated to each theoretical solution of the crosstalk canceller, i.e.

$$c(z) \approx \frac{h_x(z)}{h_y(z)} \cdot z^{-\tau}, \tag{3.1.1}$$

where $h_x(z) \in \{g_{rl}(z), g_{lr}(z), g_{ll}(z), g_{rr}(z)\}$, and $h_y(z) = g_{rr}(z)g_{ll}(z) - g_{lr}(z)g_{rl}(z)$;

τ is a delay to guarantee the causality. Therefore, the criterion is to minimize the

15

direct error $error_{FIR}(z)$ as follows:

$$error_{FIR}(z) = c(z) - \frac{h_x(z)}{h_y(z)} \cdot z^{-\tau} \qquad (3.1.2)$$

$$c(z) = \arg\min\left\{ \left\| error_{FIR}(z) \right\|^2 \right\} \qquad (3.1.3)$$

The impulse response of $h_x(z) \cdot z^{-\tau} / h_y(z)$ is showed in upper Figure 3.2 and zoomed

in tap-20 to tap-40 in lower figure. From Figure 3.2, we know that it is hard to find a

FIR filter to approximate the IIR system $h_x(z) \cdot z^{-\tau} / h_y(z)$ because it diverges too

fast.
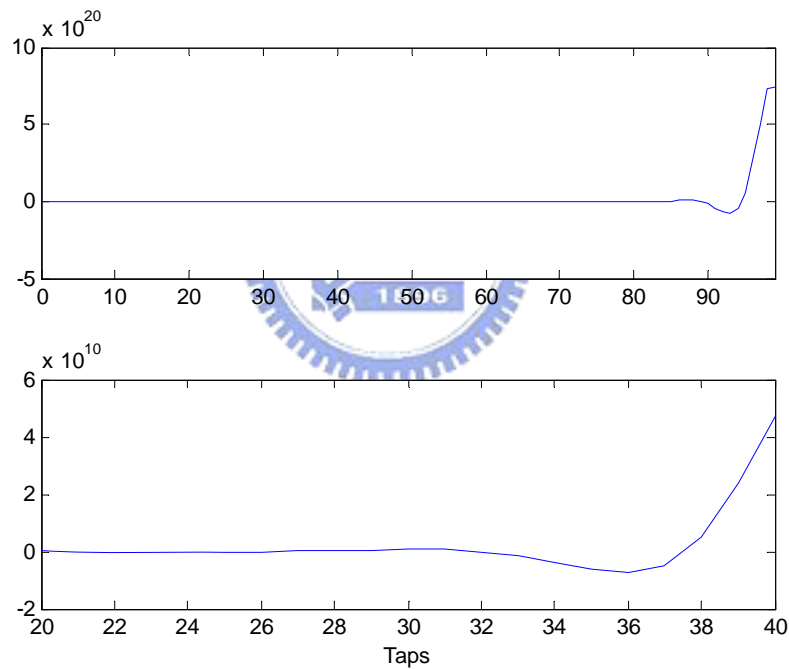
Figure 3.2: The impulse response of $h_x(z) \cdot z^{-\tau} / h_y(z)$

Therefore, Mochtaris proposes to minimize a filtered error $error_{FIR\_filtered}(z)$

expressed in Equation (3.1.4) and its block diagram is plotted in Figure 3.3 [13].

$$error_{FIR\_filtered}(z) = h_x(z) \cdot error_{FIR}(z) \qquad (3.1.4)$$

16
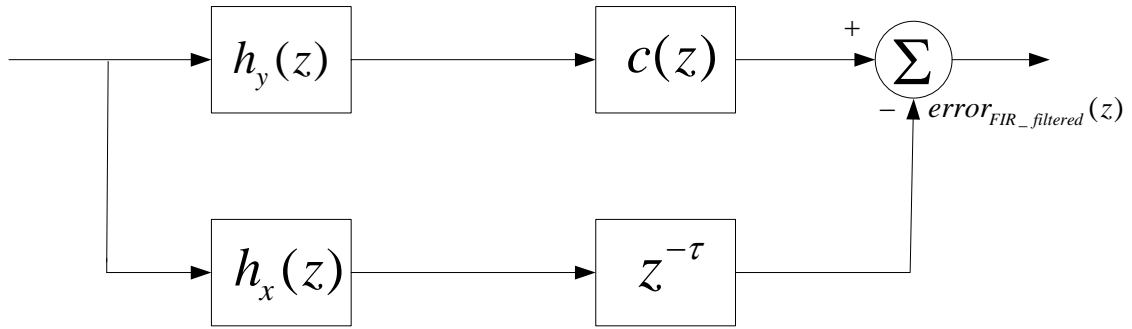
Figure 3.3: Block diagram of the filtered error $error_{FIR\_filtered}(z)$

According to Figure 3.3, the filtered error can be formulated is as follows:

$$error_{FIR\_filtered}(z) = c(z) \cdot h_y(z) - h_x(z) \cdot z^{-\tau} \qquad (3.1.5)$$

Therefore, the criterion is as follows:

$$c(z) = \arg\min\left\{\| error_{FIR\_filtered}(z) \|^2\right\} \qquad (31.6)$$

Equation (3.1.5) can be written in convolution matrix and vector forms, and expressed in Equation (3.1.7).

$$\mathbf{error}_{FIR\_filtered} = \mathbf{H}_y \cdot \mathbf{c} - \mathbf{h}_x, \qquad (3.1.7)$$

where $\mathbf{c} = \begin{bmatrix} c(0) & c(1) & \cdots & c(N-1) \end{bmatrix}^T$ is a FIR filter with N taps; $\mathbf{H}_y$ is a convolution matrix; $\mathbf{h}_x = \begin{bmatrix} \underbrace{0 \quad \cdots \quad 0}_{\tau} & h_x(0) & h_x(1) & \cdots & h_x(M-1) & \underbrace{0 \quad \cdots \quad 0}_{M+N-\tau-2} \end{bmatrix}^T$, and

M is the channel length.

We can write $\mathbf{H}_y$ in detail as follows:

$$\mathbf{H}_y = \begin{bmatrix} h_y(0) & 0 & \cdots & 0 \\ h_y(1) & h_y(0) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ h_y(2M-2) & \vdots & \cdots & h_y(0) \\ 0 & h_y(2M-2) & \cdots & h_y(1) \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & h_y(2M-2) \end{bmatrix}_{(2M+N-2) \times N}$$ (3.1.8)

One of the filters in the crosstalk canceller **c** is found as follows:

$$\mathbf{c} = \left(\mathbf{H}_y^T \mathbf{H}_y \right)^{-1} \mathbf{H}_y^T \mathbf{h}_x$$ (3. 1.9)

Referring to Equation (3.1.9) with different $h_x(z)$, we can find each filter of the crosstalk canceller. One point should be noted is that the delay in each term of the crosstalk canceller must be the same.
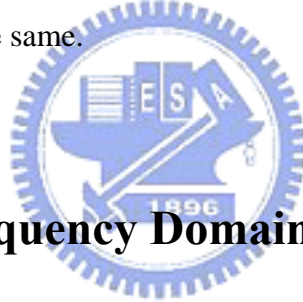
3.1.2 Design in Frequency Domain

In the previous section, the direct error is designed in the time domain, and met a divergence problem. Now we will propose a method designed in the frequency domain to avoid the problem. Equation (3.1.1) is rewritten in frequency domain as follows:

$$error_{FIR}(\omega) = c(e^{j\omega}) - \frac{H_x(e^{j\omega})}{H_y(e^{j\omega})} e^{-j\omega\tau} ,$$ (3.1.10)

where $c(e^{j\omega})$, $H_x(e^{j\omega})$ and $H_y(e^{j\omega})$ are the Fourier transforms of $c(n)$, $h_x(n)$ and $h_y(n)$. Therefore, the criterion is as follows:

18

$$c(n) = \arg\min\left\{\frac{1}{2\pi}\int_{-\pi}^{\pi}\left|error_{FIR}(\omega)\right|^2 d\omega\right\} \qquad (3.1.11)$$

In order to find the filter coefficients, we can rewrite the Fourier transform of $c(e^{j\omega})$

in vector form as follows [14].

$$\begin{aligned}
c(e^{j\omega}) &= \sum_{n=0}^{n=N-1} c(n)\cdot e^{-jn\omega} \\
&= \mathbf{c}^T \cdot \mathbf{ex}(\omega)
\end{aligned} \qquad (3.1.12)$$

where $\mathbf{ex}(\omega) = [e^{-j0\times\omega}, e^{-j1\times\omega}, \cdots, e^{-j(N-1)\times\omega}]^T$

Referring to Equation (3. 1.12), the error energy can be rewritten as follows:

$$\begin{aligned}
J_{FIR} &= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|error_{FIR}(\omega)\right|^2 d\omega \\
&= \mathbf{c}^T\cdot\mathbf{c} - 2\mathbf{c}^T\cdot\mathbf{b} + \frac{1}{2\pi}\int_{-\pi}^{\pi}\left|\frac{H_x(e^{j\omega})}{H_y(e^{j\omega})}\right|^2 d\omega,
\end{aligned} \qquad (3.1.13)$$

where $\mathbf{b} = \dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi}\mathrm{Re}\left\{\mathbf{ex}(\omega)\cdot\left(\dfrac{H_x(e^{j\omega})}{H_y(e^{j\omega})}\cdot e^{-j\cdot\tau\cdot\omega}\right)^*\right\}d\omega$

In order to minimize J_{FIR}, let $\dfrac{\partial J_{FIR}}{\partial\mathbf{c}} = \mathbf{0}$. We can get

$$\mathbf{c} = \mathbf{b} \qquad (3.1.14)$$

However, we find the performance is bad. The reason is that it is difficult to

approximate the high frequency band of $\left(H_x(e^{j\omega})/H_y(e^{j\omega})\right)$. We know that $h_y(n)$

is the convolution of two HTRFs, so $\left(H_x(e^{j\omega})/H_y(e^{j\omega})\right)$ can be viewed as the

inverse of first order of HRTF. From Figure 2.3, we can know that the high frequency

of HRTF are decayed very much, so the high frequency magnitude responses of

$\left(H_x(e^{j\omega})/H_y(e^{j\omega})\right)$ are very large. The frequency magnitude responses of

$\left(H_x(e^{j\omega})/H_y(e^{j\omega})\right)$ and $c(e^{j\omega})$ are plotted in Figure 3.4.
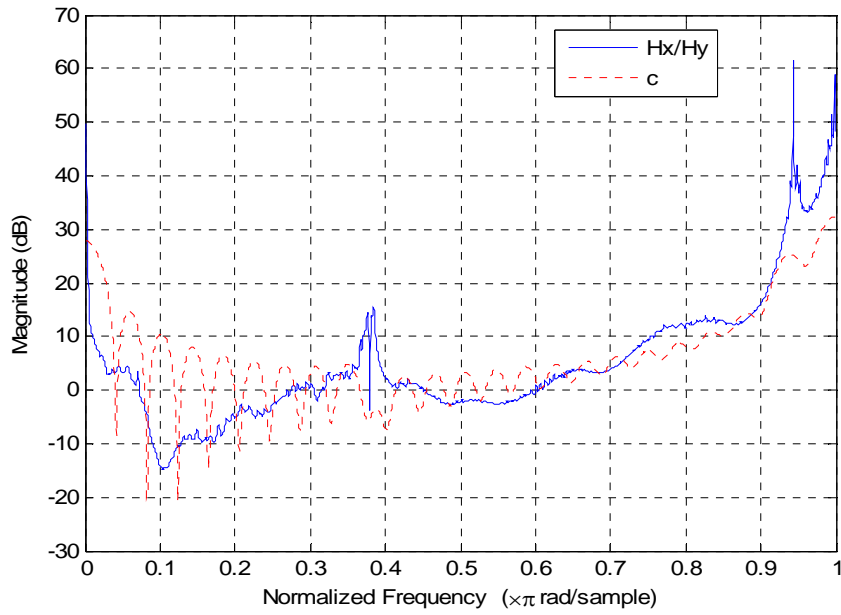
Figure 3.4: Magnitude responses of $\left(H_x(e^{j\omega})/H_y(e^{j\omega}) \right)$ and $c(e^{j\omega})$

From Figure 3.4, we know that the filter $c(e^{j\omega})$ would seek the high gain in the high

frequency band and sacrifice the low band in order to compromise the full band. The

following figure is the magnitude response of the direct error.
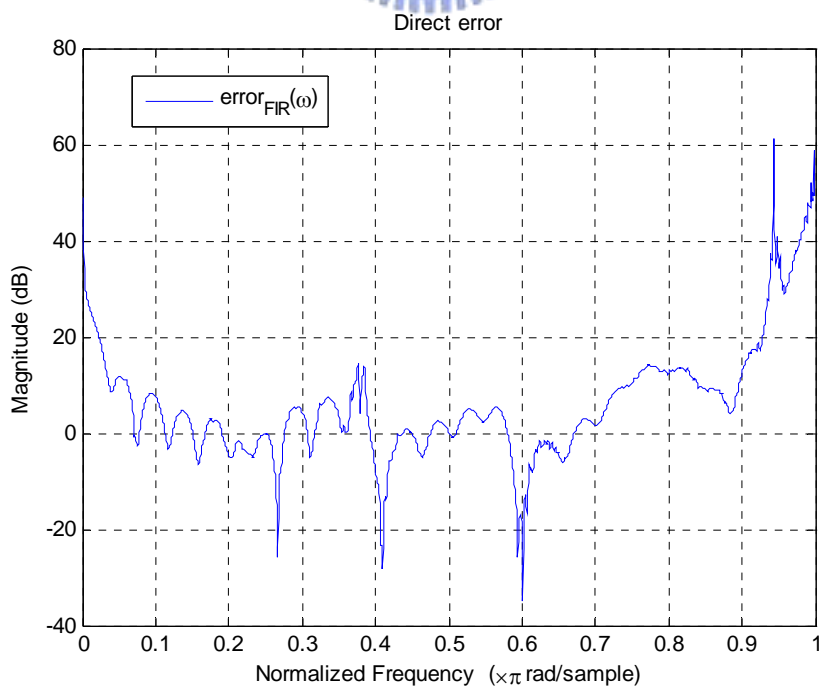
Figure 3.5: Magnitude response of the direct error $error_{FIR}(\omega)$

From Figure 3.5, we can know that the error in high frequency band is large, so the performance is bad. Therefore, we propose a method to minimize the ratio error $error_{FIR\_ratio}(\omega)$ instead of the direct error $error_{FIR}(\omega)$. The ratio error is defined as follows:

$$
\begin{aligned}
error_{FIR\_ratio}(\omega) &= \frac{H_y(e^{j\omega})}{H_x(e^{j\omega})} \cdot error_{FIR}(\omega) \\
&= \frac{H_y(e^{j\omega})}{H_x(e^{j\omega})} \cdot c(e^{j\omega}) - e^{-j\cdot\omega\cdot\tau}
\end{aligned}
\tag{3.1.15}
$$

The block diagram of the ratio error is as follows:
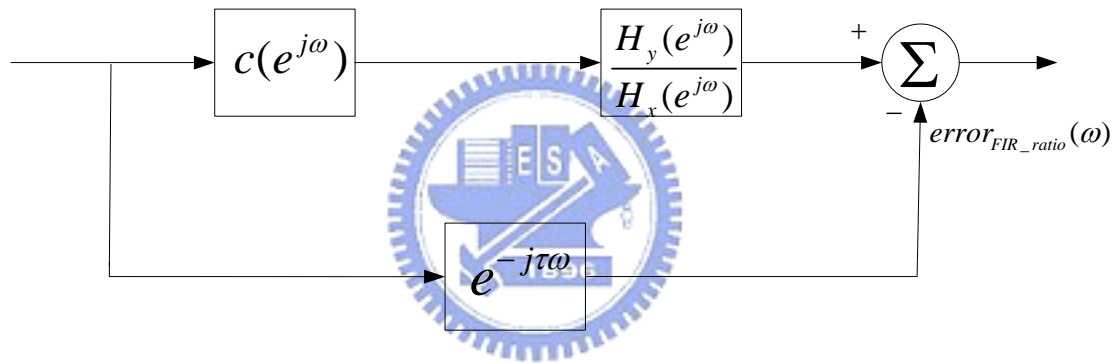
Figure 3.6: Block diagram of the ratio error $error_{FIR\_ratio}(\omega)$

Therefore, the criterion is as follows:

$$
c(n) = \arg\min\left\{ \frac{1}{2\pi}\int_{-\pi}^{\pi} |error_{FIR\_ratio}(\omega)|^2 \, d\omega \right\}
\tag{3.1.16}
$$

In the same way, we change the expression of $c(e^{j\omega})$ in the vector form as Equation (3.1.12) and the error energy can be rewritten as follows:

$$
\begin{aligned}
J_{ratio} &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |\frac{H_y(e^{j\omega})}{H_x(e^{j\omega})} \cdot c(e^{j\omega}) - e^{-j\cdot\tau\cdot\omega}|^2 \, d\omega, \\
&= \mathbf{c}^T \cdot \mathbf{A} \cdot \mathbf{c} - 2\mathbf{c}^T \cdot \mathbf{b} + 1
\end{aligned}
\tag{3.1.17}
$$

where $\mathbf{A} = \dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} \mathbf{ex}(\omega) \cdot \left| \dfrac{Hy(e^{j\omega})}{Hx(e^{j\omega})} \right|^2 \cdot \mathbf{ex}(\omega)^H \, d\omega$

$\mathbf{b} = \dfrac{1}{2\pi} \displaystyle\int_{-\pi}^{\pi} \mathrm{Re}\{\mathbf{ex}(\omega) \cdot (\dfrac{Hy(e^{j\omega})}{Hx(e^{j\omega})} \cdot e^{j\cdot\tau\cdot\omega})\} d\omega$

In order to minimize J_{FIR_ratio}, let $\dfrac{\partial J_{FIR\_ratio}}{\partial \mathbf{c}} = \mathbf{0}$. We can get

$$\mathbf{Ac} = \mathbf{b} \qquad\qquad (3.1.18)$$

Therefore, the filter can be found out as follows:

$$\mathbf{c} = \mathbf{A^{-1}b} \qquad\qquad (3.1.19)$$

The frequency magnitude responses of $\left(H_y(e^{j\omega}) / H_x(e^{j\omega}) \right)$ and $c(e^{j\omega})$ are plotted

in Figure 3.7.
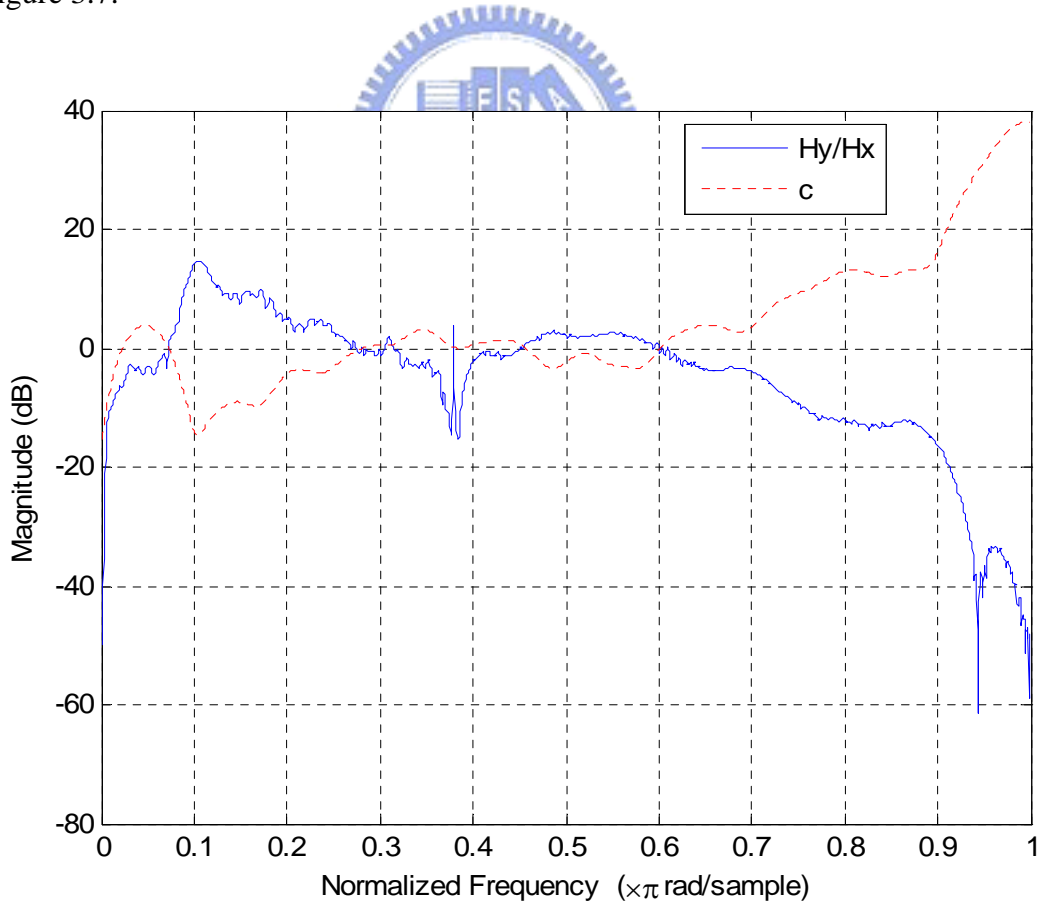
Figure 3.7: Magnitude responses of $\left(H_y(e^{j\omega}) / H_x(e^{j\omega}) \right)$ and $c(e^{j\omega})$

The high frequency band of $\left(H_y(e^{j\omega})/H_x(e^{j\omega}) \right)$ would be very small because it can be approximate as first order of HRTF. Therefore, the high frequency band of $c(e^{j\omega})$ must be very large. From Figure 3.7, although the high frequency band of $c(e^{j\omega})$ is not large enough, the error is smaller than the direct error. Therefore, the errors in high band are reduced, and the low frequency band can be done better. Figure 3.8 shows the magnitude response of the ratio error.
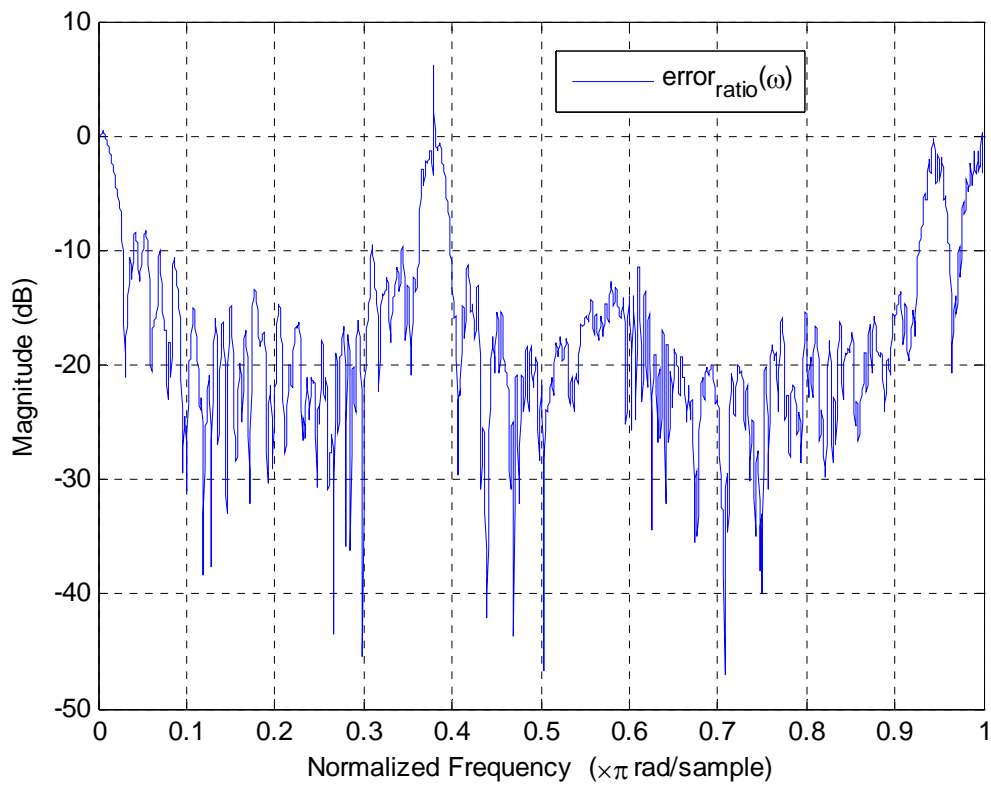
Figure 3.8: Magnitude response of the ratio error $error_{ratio}(\omega)$

From Equation (3.1.15), we know the relation between the direct error and the ratio error. Therefore, Figure 3.9 shows the results of the direct error and the ratio error filtered by $H_x(e^{j\omega})/H_y(e^{j\omega})$ in order to compare these two errors.
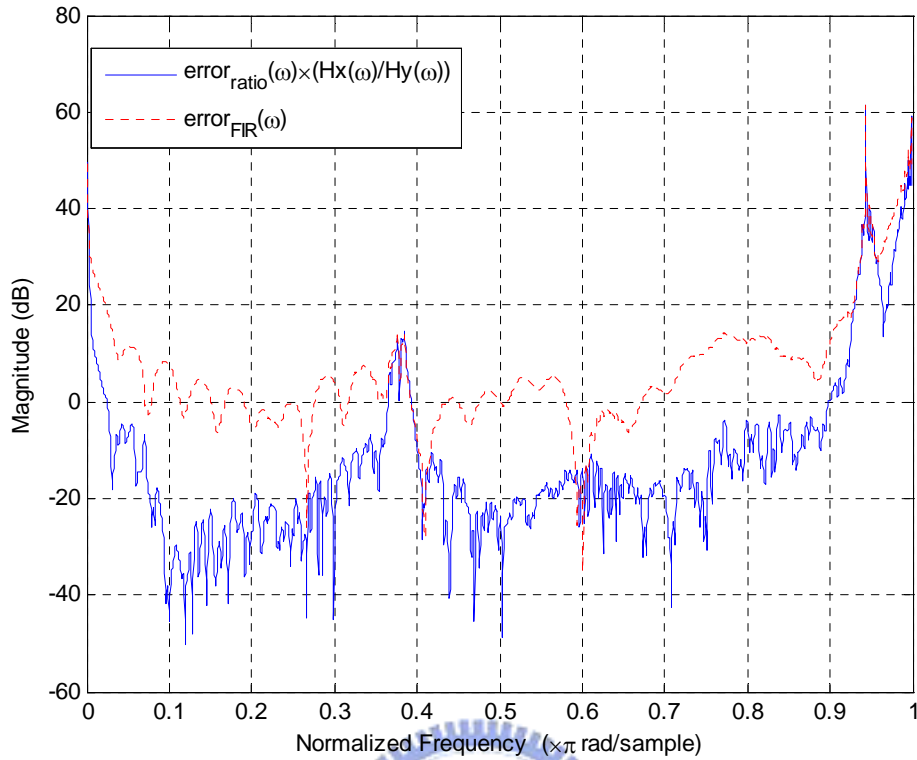
Figure 3.9: Magnitude responses of $error_{FIR}(\omega)$ and $error_{FIR\_ratio}(\omega) \cdot \dfrac{H_x(e^{j\omega})}{H_y(e^{j\omega})}$

From Figure 3.9, it is obvious that the ratio error filtered by $H_x(e^{j\omega})/H_y(e^{j\omega})$ is smaller than the direct error.

From Equation (3.1.15), we can find it is the same as the direct LSE FIR method which will be proposed in the next section.

3.2 Direct LSE FIR Design

3.2.1 Design in Time Domain

In this section, we design the crosstalk canceller by using the direct least square error (LSE) method instead of the matrix-inverse design. The structure of the

crosstalk canceller is plotted in Figure 3.1, and $C(z)$ can be represented in matrix

form as follows:

$$C(z) = \begin{bmatrix} c_{11}(z) & c_{12}(z) \\ c_{21}(z) & c_{22}(z) \end{bmatrix} \tag{3.2.1}$$

In order to show how the crosstalk canceller works, s_R is an impulse signal, and

s_L sends the zero-signal. Because of causality, we want e_R to be same as s_R with

some extra delay, and e_L to be a zero-signal. The signals to ears can be expressed as

follows:

$$e_R(z) \approx d(z) \tag{3.2.2}$$

$$e_L(z) \approx 0, \tag{3.2.3}$$

where $d(z)$ is the desired signal which is a delayed impulse signal. Therefore, the

LSE criterion to find the filters $c_{11}(z)$ and $c_{21}(z)$ can be rewritten as follows:

$$\begin{pmatrix} c_{11}(z) & c_{21}(z) \end{pmatrix} = \arg\min \left\{ \left\| G(z) \begin{bmatrix} c_{11}(z) \\ c_{21}(z) \end{bmatrix} - \begin{bmatrix} d(z) \\ 0 \end{bmatrix} \right\|^2 \right\} \tag{3.2.4}$$

In the same way, if s_L is an impulse signal, and s_R is a zero-signal such that:

$$\begin{pmatrix} c_{12}(z) & c_{22}(z) \end{pmatrix} = \arg\min \left\{ \left\| G(z) \begin{bmatrix} c_{12}(z) \\ c_{22}(z) \end{bmatrix} - \begin{bmatrix} 0 \\ d(z) \end{bmatrix} \right\|^2 \right\} \tag{3.2.5}$$

Equation (3.2.4) and Equation (3.2.5) can be derived in convolution matrixes as

follows:

$$\begin{pmatrix} \mathbf{c}_{11} & \mathbf{c}_{21} \end{pmatrix} = \arg\min \left\{ \left\| \begin{bmatrix} \mathbf{G}_{rr} & \mathbf{G}_{lr} \\ \mathbf{G}_{rl} & \mathbf{G}_{ll} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{11} \\ \mathbf{c}_{21} \end{bmatrix} - \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \right\|^2 \right\} \tag{3.2.6}$$

$$\begin{pmatrix} \mathbf{c}_{12} & \mathbf{c}_{22} \end{pmatrix} = \arg\min \left\{ \left\| \begin{bmatrix} \mathbf{G}_{rr} & \mathbf{G}_{lr} \\ \mathbf{G}_{rl} & \mathbf{G}_{ll} \end{bmatrix} \begin{bmatrix} \mathbf{c}_{12} \\ \mathbf{c}_{22} \end{bmatrix} - \begin{bmatrix} \mathbf{0} \\ \mathbf{d} \end{bmatrix} \right\|^2 \right\}, \tag{3.2.7}$$

where $\mathbf{G}_{rr}$, $\mathbf{G}_{lr}$, $\mathbf{G}_{rl}$, and $\mathbf{G}_{ll}$ are the convolution matrixes of $g_{rr}(z)$, $g_{lr}(z)$,

$g_{rl}(z)$, and $g_{ll}(z)$ in the time domain; $\mathbf{c}_{11} = \begin{bmatrix} c_{11}(0) & c_{11}(1) & \cdots & c_{11}(N-1) \end{bmatrix}^T$ and

$\mathbf{c}_{21} = \begin{bmatrix} c_{21}(0) & c_{21}(1) & \cdots & c_{21}(N-1) \end{bmatrix}^T$; the desired signal $\mathbf{d} = \begin{bmatrix} 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \end{bmatrix}^T$ and

zero-vector $\mathbf{0}$ with $L(= M + N - 1)$ taps.

Rewrite $\mathbf{G}_{ij}$ in detail, where $i, j \in \{r, l\}$ as follows:

$$\mathbf{G}_{ij} = \begin{bmatrix} g_{ij}(0) & 0 & \cdots & 0 \\ g_{ij}(1) & g_{ij}(0) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ g_{ij}(M-1) & \vdots & \cdots & g_{ij}(0) \\ 0 & g_{ij}(M-1) & \cdots & g_{ij}(1) \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & g_{ij}(M-1) \end{bmatrix}_{L \times N} \tag{3.2.8}$$

Refer to Equation (3.2.6), and let

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_{rr} & \mathbf{G}_{lr} \\ \mathbf{G}_{rl} & \mathbf{G}_{ll} \end{bmatrix}, \quad \mathbf{c}_1 = \begin{bmatrix} \mathbf{c}_{11} \\ \mathbf{c}_{21} \end{bmatrix}, \quad \mathbf{q}_1 = \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \tag{3.2.9}$$

we know the error as follows:

$$\mathbf{error} = \mathbf{Gc}_1 - \mathbf{q}_1 \tag{3.2.10}$$

The least square error criterion is such that:

$$\mathbf{c}_1 = \arg\min \left\{ \| \mathbf{error} \|^2 \right\} \tag{3.2.11}$$

The solution can be easily shown as

$$\mathbf{c}_1 = \left(\mathbf{G}^T \mathbf{G} \right)^{-1} \mathbf{G}^T \mathbf{q}_1 \tag{3.2.12}$$

Therefore,

$$\mathbf{c}_{11} = \left[c_1(0), c_1(1), \cdots, c_1(N-1) \right]^T \tag{3.2.13}$$

$$\mathbf{c}_{21} = \left[c_1(N), c_1(N+1), \cdots, c_1(2N-1) \right]^T \tag{3.2.14}$$

Refer to Equation (3.2.7), and let

$$\mathbf{c}_2 = \begin{bmatrix} \mathbf{c}_{12} \\ \mathbf{c}_{22} \end{bmatrix}, \quad \mathbf{q}_2 = \begin{bmatrix} \mathbf{0} \\ \mathbf{d} \end{bmatrix}, \tag{3.2.15}$$

where $\mathbf{c}_{12}$ and $\mathbf{c}_{22}$ with N taps, too. In the same way, we can get

$$\mathbf{c}_2 = \left(\mathbf{G}^T \mathbf{G} \right)^{-1} \mathbf{G}^T \mathbf{q}_2 \tag{3.2.16}$$

Therefore,

$$\mathbf{c}_{12} = \left[c_2(0), c_2(1), \cdots, c_2(N-1) \right]^T \tag{3.2.17}$$

$$\mathbf{c}_{22} = \left[c_2(N), c_2(N+1), \cdots, c_2(2N-1) \right]^T \tag{3.2.18}$$

If the listening situation is symmetric, Equation (3.2.15) can be reduced. The other

two filters can be found in Equation (3.2.13) and Equation (3.2.14) such that:

$$\mathbf{c}_{12} = \mathbf{c}_{21} \quad \text{and} \quad \mathbf{c}_{22} = \mathbf{c}_{11} \tag{3.2.19}$$

3.2.2 Design in Frequency Domain

In this section, we will implement the system in the frequency domain. Equation

(3. 2.4) can be written in the frequency domain as follows:

$$\left(c_{11}(e^{j\omega}) \; c_{21}(e^{j\omega}) \right) = \arg\min \left\{ \frac{1}{2\pi} \int_{-\pi}^{\pi} \left\| \begin{bmatrix} g_{rr}(e^{j\omega}) & g_{lr}(e^{j\omega}) \\ g_{rl}(e^{j\omega}) & g_{rr}(e^{j\omega}) \end{bmatrix} \begin{bmatrix} c_{11}(e^{j\omega}) \\ c_{21}(e^{j\omega}) \end{bmatrix} - \begin{bmatrix} e^{-j\omega\tau} \\ 0 \end{bmatrix} \right\|^2 d\omega \right\} \tag{3.2.20}$$

As described in Section 3.1.1.2, we can rewrite $c_{11}(e^{j\omega})$ and $c_{12}(e^{j\omega})$ as follows:

$$c_{11}(e^{j\omega}) = \mathbf{ex}_1^T \cdot \mathbf{c}_{11}, \tag{3.2.21}$$

where $\mathbf{ex}_1(\omega) = [e^{-j0\times\omega}, e^{-j1\times\omega}, \cdots, e^{-j(N-1)\times\omega}]^T$.

$$c_{21}(e^{j\omega}) = \mathbf{ex}_2^T \cdot \mathbf{c}_{21}, \tag{3.2.22}$$

where $\mathbf{ex}_2(\omega) = [e^{-j0\times\omega}, e^{-j1\times\omega}, \cdots, e^{-j(N-1)\times\omega}]^T$.

Therefore, Equation (3.2.20) can be rewritten as follows:

$$\begin{pmatrix} \mathbf{c}_{11} & \mathbf{c}_{21} \end{pmatrix} = \arg\min\left\{ \int_{-\pi}^{\pi} \left| \mathbf{P} \cdot \begin{bmatrix} \mathbf{c}_{11} \\ \mathbf{c}_{21} \end{bmatrix} - \begin{bmatrix} e^{-j\omega\tau} \\ 0 \end{bmatrix} \right|^2 d\omega \right\}, \tag{3.2.23}$$

where $\mathbf{P} = \begin{bmatrix} g_{rr}(e^{j\omega}) \cdot \mathbf{ex}_1(\omega)^T & g_{lr}(e^{j\omega}) \cdot \mathbf{ex}_2(\omega)^T \\ g_{rl}(e^{j\omega}) \cdot \mathbf{ex}_1(\omega)^T & g_{ll}(e^{j\omega}) \cdot \mathbf{ex}_2(\omega)^T \end{bmatrix}$.

The error J_{LSE_FIR} can be calculated as follows:

$$J_{LSE\_FIR} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \mathbf{P} \cdot \begin{bmatrix} \mathbf{c}_{11} \\ \mathbf{c}_{21} \end{bmatrix} - \begin{bmatrix} e^{-j\omega\tau} \\ 0 \end{bmatrix} \right|^2 d\omega \tag{3.2.24}$$

In the same way, we can rewrite the above equation as follows:

$$J_{LSE\_FIR} = \mathbf{c}_1^T \cdot \mathbf{O} \cdot \mathbf{c}_1 - 2 \cdot \mathbf{c}_1^T \cdot \mathbf{m} + 1, \tag{3.2.25}$$

where $\mathbf{c}_1 = \begin{bmatrix} \mathbf{c}_{11} & \mathbf{c}_{21} \end{bmatrix}^T$,

$$\mathbf{O} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left(\mathbf{P}^H \mathbf{P} \right) d\omega, \text{ and}$$

$$\mathbf{m} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Re}\left\{ \begin{bmatrix} e^{j\omega\tau} \cdot g_{rr}(e^{j\omega}) \cdot \mathbf{ex}_1(\omega) \\ e^{j\omega\tau} \cdot g_{lr}(e^{j\omega}) \cdot \mathbf{ex}_2(\omega) \end{bmatrix} \right\} d\omega.$$

In order to minimize the error J_{LSE_FIR}, let $\dfrac{\partial J_{LSE\_FIR}}{\partial \mathbf{c}_1} = 0$ We can get

$$\mathbf{O}\mathbf{c}_1 = \mathbf{m} \tag{3.2.26}$$

Therefore, the filter can be found out as follows:

$$\mathbf{c}_1 = \mathbf{O}^{-1}\mathbf{m} \tag{3.2.27}$$

We can find the filter coefficients as follows:

$$\mathbf{c}_{11} = \left[c_1(0), c_1(1), \cdots, c_1(N-1)\right]^T \tag{3.2.28}$$

$$\mathbf{c}_{21} = \left[c_1(N), c_1(N+1), \cdots, c_1(2N-1)\right]^T \tag{3.2.29}$$

$\mathbf{c}_{12}$ and $\mathbf{c}_{22}$ can be found out by using the same method.

3.2.3 Comparison between FIR Designs in Time and Frequency Domains

In Section (3.1.2.1) and Section (3.1.2.2), we can find out the crosstalk canceller in LSE FIR model in the time domain and in the frequency domain. The difference between these two methods is to minimize the error in time and frequency domains. According to Parseval theorem [14], they should be equal. Therefore, the results of the two methods should be the same intuitionally. We will prove that two crosstalk cancellers designed in the two methods are the same, and the process of the proof is as follows:

First, let $\mathbf{O}' = \mathbf{G}^T\mathbf{G}$ and $\mathbf{m}' = \mathbf{G}^T\mathbf{q}_1$. Equation (3.2.12) can be rewritten as follows:

$$\mathbf{c}_1 = \left(\mathbf{O}'\right)^{-1}\mathbf{m}' \tag{3.2.30}$$

We can write $\mathbf{O}'$ in more detail as follows:

$$\mathbf{O}' = \begin{bmatrix} \mathbf{O}_1' & \mathbf{O}_2' \\ \mathbf{O}_3' & \mathbf{O}_4' \end{bmatrix}, \qquad (3.2.31)$$

where

$$\mathbf{O}_1' = \mathbf{G}_{rr}{}^T \mathbf{G}_{rr} + \mathbf{G}_{rl}{}^T \mathbf{G}_{rl}$$
$$\mathbf{O}_2' = \mathbf{G}_{rr}{}^T \mathbf{G}_{lr} + \mathbf{G}_{rl}{}^T \mathbf{G}_{ll}$$
$$\mathbf{O}_3' = \mathbf{G}_{lr}{}^T \mathbf{G}_{rr} + \mathbf{G}_{ll}{}^T \mathbf{G}_{rl}$$
$$\mathbf{O}_4' = \mathbf{G}_{lr}{}^T \mathbf{G}_{lr} + \mathbf{G}_{ll}{}^T \mathbf{G}_{ll}$$

In order to compare $\mathbf{O}'$ with $\mathbf{O}$ in Equation (3.2.27), $\mathbf{O}$ is also written in detail as follows:

$$\mathbf{O} = \begin{bmatrix} \mathbf{O}_1 & \mathbf{O}_2 \\ \mathbf{O}_3 & \mathbf{O}_4 \end{bmatrix}, \qquad (3.2.32)$$

where

$$\mathbf{O}_1 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left| g_{rr}(e^{j\omega}) \right|^2 \left(\mathbf{ex}_1 \mathbf{ex}_1{}^H \right)^* d\omega + \frac{1}{2\pi}\int_{-\pi}^{\pi} \left| g_{rl}(e^{j\omega}) \right|^2 \left(\mathbf{ex}_1 \mathbf{ex}_1{}^H \right)^* d\omega$$

$$\mathbf{O}_2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} g_{lr}(e^{j\omega}) g_{rr}{}^*(e^{j\omega}) \left(\mathbf{ex}_1 \mathbf{ex}_2{}^H \right)^* d\omega + \frac{1}{2\pi}\int_{-\pi}^{\pi} g_{ll}(e^{j\omega}) g_{rl}{}^*(e^{j\omega}) \left(\mathbf{ex}_1 \mathbf{ex}_2{}^H \right)^* d\omega$$

$$\mathbf{O}_3 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left(g_{lr}(e^{j\omega}) g_{rr}{}^*(e^{j\omega}) \right)^* \left(\mathbf{ex}_2 \mathbf{ex}_1{}^H \right)^* d\omega + \frac{1}{2\pi}\int_{-\pi}^{\pi} \left(g_{ll}(e^{j\omega}) g_{rl}{}^*(e^{j\omega}) \right)^* \left(\mathbf{ex}_2 \mathbf{ex}_1{}^H \right)^* d\omega$$

$$\mathbf{O}_4 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \left| g_{lr}(e^{j\omega}) \right|^2 \left(\mathbf{ex}_2 \mathbf{ex}_2{}^H \right)^* d\omega + \frac{1}{2\pi}\int_{-\pi}^{\pi} \left| g_{ll}(e^{j\omega}) \right|^2 \left(\mathbf{ex}_2 \mathbf{ex}_2{}^H \right)^* d\omega$$

Comparing the elements between in $\mathbf{O}$ and $\mathbf{O}'$, we can find they are the same and the proof as follows:

Take $\mathbf{O}_2$ and $\mathbf{O}_2'$ for example, and other elements are proved in the same way.

From $\mathbf{O}_2' = \mathbf{G}_{rr}{}^T \mathbf{G}_{lr} + \mathbf{G}_{rl}{}^T \mathbf{G}_{ll}$, the first right terms $\mathbf{G}_{rr}{}^T \mathbf{G}_{lr}$ is considered.

$$\mathbf{G}_{rr}{}^{T}\mathbf{G}_{lr} = \begin{bmatrix} \sum_{k=0}^{L-1} g_{rr}(k)g_{lr}(k) & \sum_{k=0}^{L-1} g_{rr}(k)g_{lr}(k-1) & \cdots & \sum_{k=0}^{L-1} g_{rr}(k)g_{lr}(k-N) \\ \sum_{k=0}^{L-1} g_{rr}(k-1)g_{lr}(k) & \ddots & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \sum_{k=0}^{L-1} g_{rr}(k-N)g_{lr}(k) & \cdots & \cdots & \sum_{k=0}^{L-1} g_{rr}(k-N)g_{lr}(k-N) \end{bmatrix}$$

From above equation, we can know each element in $\mathbf{G}_{rr}{}^{T}\mathbf{G}_{lr}$ as follows:

$$\mathbf{G}_{rr}{}^{T}\mathbf{G}_{lr}(p,q) = \sum_{k=0}^{L-1} g_{rr}(k-p)g_{lr}(k-q), \tag{3.2.33}$$

where $p = 0 \sim N-1$ and $q = 0 \sim N-1$

From $\mathbf{O}_2$, the first right term $\dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} g_{lr}(e^{j\omega})g_{rr}{}^{*}(e^{j\omega})\left(\mathbf{ex}_1\mathbf{ex}_2{}^{H}\right)^{*} d\omega$ is considered.

Let $\mathbf{T} = \dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} g_{lr}(e^{j\omega})g_{rr}{}^{*}(e^{j\omega})\left(\mathbf{ex}_1\mathbf{ex}_2{}^{H}\right)^{*} d\omega$

$$\mathbf{T} = \frac{1}{2\pi}\int_{-\pi}^{\pi} g_{lr}(e^{j\omega})g_{rr}{}^{*}(e^{j\omega}) \begin{bmatrix} e^{-j(0)\omega} & e^{-j(0-1)\omega} & \cdots & e^{-j(0-(N_{21}-1))\omega} \\ e^{-j(1)\omega} & & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ e^{-j(N_{11}-1)\omega} & \cdots & \cdots & e^{-j(N_{11}-N_{21})\omega} \end{bmatrix}^{*} d\omega$$

$$\begin{aligned}
\mathbf{T}(p,q) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} g_{lr}(e^{j\omega})g_{rr}{}^{*}(e^{j\omega})e^{j(p-q)\omega} d\omega \\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sum_{n=0}^{L-1} g_{lr}(n)e^{-jn\omega}\right)\left(\sum_{m=0}^{L-1} g_{rr}(m)e^{-jm\omega}\right)^{*} e^{j(p-q)\omega} d\omega \\
&= \sum_{n=0}^{L-1} g_{lr}(n)\sum_{m=0}^{L-1} g_{rr}(m)\frac{1}{2\pi}\int_{-\pi}^{\pi} e^{j(p+m-q-n)\omega} d\omega \\
&= \sum_{m=0}^{L-1} g_{lr}(m-q)g_{rr}(m-p)
\end{aligned} \tag{3.2.34}$$

where $p = 0 \sim N-1$ *and* $q = 0 \sim N-1$.

We can find that the results in Equation (3.2.33) and Equation (3.2.34) are the same. Therefore, $\dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} g_{lr}(e^{j\omega})g_{rr}{}^{*}(e^{j\omega})\left(\mathbf{ex}_1\mathbf{ex}_2{}^{H}\right)^{*} d\omega$ equals $\mathbf{G}_{rr}{}^{T}\mathbf{G}_{lr}$. In the same way, we can prove that $\dfrac{1}{2\pi}\displaystyle\int_{-\pi}^{\pi} g_{ll}(e^{j\omega})g_{rl}{}^{*}(e^{j\omega})\left(\mathbf{ex}_1\mathbf{ex}_2{}^{H}\right)^{*} d\omega$ equals $\mathbf{G}_{rl}{}^{T}\mathbf{G}_{ll}$. In

other words, we can prove $\mathbf{O}_2 = \mathbf{O}_2'$. In the same way, other terms in $\mathbf{O}$ and $\mathbf{O}'$ can

be proved equals. Therefore, we can say

$$\mathbf{O} = \mathbf{O}'$$ (3.2.35)

Now, we consider the terms $\mathbf{m}$ and $\mathbf{m}'$. We know that

$$\mathbf{m}' = \mathbf{G}^T \mathbf{q}_1 = \begin{bmatrix} \mathbf{G}_{rr}{}^T \mathbf{d} \\ \mathbf{G}_{lr}{}^T \mathbf{d} \end{bmatrix}$$

We first consider the upper term $\mathbf{G}_{rr}{}^T \mathbf{d}$ written as follows:

$$\mathbf{G}_{rr}{}^T \mathbf{d} = \begin{bmatrix} g_{rr}(\tau) & g_{rr}(\tau-1) & \cdots & g_{rr}\left(\tau-(N_{11}-1)\right) \end{bmatrix}^T$$ (3.2.36)

The upper term in $\mathbf{m}$ is also considered first. Each term in upper term is as follows:

$$\begin{aligned} \mathbf{m}_{upper}(s) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Re}\left\{ e^{j\omega\tau} \cdot g_{rr}(e^{j\omega}) \cdot e^{-js\omega} \right\} d\omega, \ s = 0 \sim N-1 \\ &= \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{g_{rr}(e^{j\omega}) \cdot e^{-j\omega(\tau-s)} + \left(g_{rr}(e^{j\omega}) \cdot e^{-j\omega(\tau-s)} \right)^*}{2} d\omega \\ &= g_{rr}(\tau-s) \end{aligned}$$ (3.2.37)

We can find that the results in Equation (3.2.36) and Equation (3.2.37) are the same.

In the same way, we can prove the lower terms in $\mathbf{m}$ and $\mathbf{m}'$ are the same.

Therefore, we can say that

$$\mathbf{m} = \mathbf{m}'$$ (3.2.38)

From the results in Equation (3.2.37) and Equation (3.2.38), we can know the FIR

crosstalk cancellers designed directly in the time and frequency domains are the same.

Chapter 4

IIR Crosstalk Canceller

In this chapter, the crosstalk canceller is designed in IIR form. We also use these

two criteria, matrix inverse and direct LSE IIR to design the filters.

4.1 Matrix Inverse Design

4.1.1 Design in Time Domain

Referring to Equation (2.3.5) and Equation (3.1.1), we have known the theoretical

solutions of the crosstalk canceller. We want that each term of G^{-1} can be

approximated by using IIR from as expressed in Equation (4.1.1) and the IIR structure

of the crosstalk canceller is diagramed in Figure 4.1.

$$\frac{b(z)}{a(z)} \approx \frac{h_x(z)}{h_y(z)} \cdot z^{-\tau} \qquad (4.1.1)$$
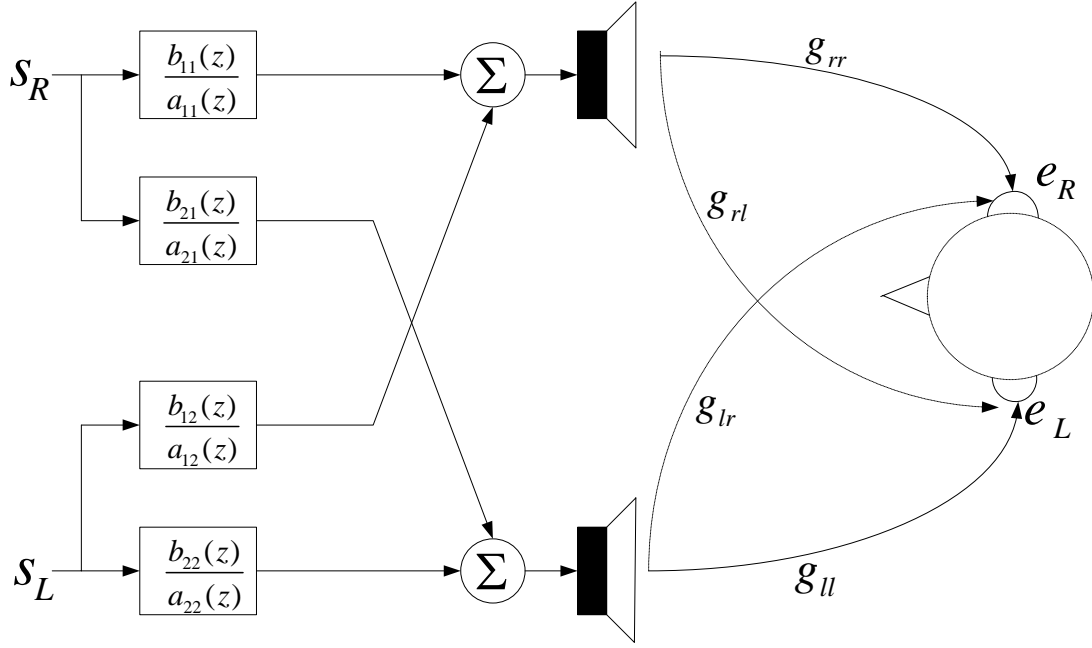
Figure 4.1: The structure of the crosstalk canceller in IIR form

Now, the criterion to design the IIR filter is to minimize the error as follows:

$$error_{IIR}(z) = \frac{b(z)}{a(z)} - \frac{h_x(z)}{h_y(z)} \cdot z^{-\tau} \tag{4.1.2}$$

$$\left(b(n), a(n)\right) = \arg\min\left\{\sum_{n=0}^{\infty} \left|error_{IIR}(n)\right|^2\right\} \tag{4.1.3}$$

Let $u(z) = \dfrac{1}{a(z)}$, and $a(0) = 1$. We can get

$$u(n) = \delta(n) - \sum_k a(k)u(n-k)$$

Therefore, $error_{IIR}(n)$ can be rewritten as follows:

$$error_{IIR}(n) = \sum_{t=0} b(t)\left(\delta(n-t) + \sum_{k=1} a(k)u(n-t-k)\right) - r(n), \tag{4.1.4}$$

where $r(n)$ is the impulse response of $h_x(z) \cdot z^{-\tau} / h_y(z)$.

Therefore, we can find that Equation (4.1.4) is a function of $a(k)$ and $b(t)$. We may want to differentiate Equation (4.1.4) with respect to $a(k)$ and $b(t)$, and set the derivatives to zeros to find $a(k)$ and $b(t)$. However, a lot of nonlinear equations

34

will appear and it will be a very tough work. Therefore, we use the Prony's Method

concept [15] to linearize the problem by multiplication (filtering) of the denominator

$a(z)$. Equation (4.1.2) can be rewritten as follows:

$$
\begin{aligned}
error_{IIR\_fitered\_1}(z) &= a(z) \cdot error_{IIR}(z) \\
&= b(z) - a(z) \cdot \frac{h_x(z)}{h_y(z)} z^{-\tau}
\end{aligned}
\tag{4.1.5}
$$

The filtered error $error_{IIR\_fitered\_1}(z)$ can be expressed as follows:

$$
error_{IIR\_filtered\_1}(n) = b(n) - \sum_{k=1} a(k)r(n-k) - r(n)
\tag{4.1.6}
$$

From the discussion in the previous section, we know that $r(n)$ will diverge.

Therefore, we multiply $h_y(z)$ to further stabilize the problem and minimize the

following filtered error $error_{IIR\_filtered\_2}(z)$.

$$
\begin{aligned}
error_{IIR\_filtered\_2}(z) &= h_y(z) \cdot error_{IIR\_filtered\_1}(z) \\
&= a(z) \cdot h_x(z) \cdot z^{-\tau} - b(z) \cdot h_y(z)
\end{aligned}
\tag{4.1.7}
$$

The block diagram of the filtered error is given in Figure 4.2.
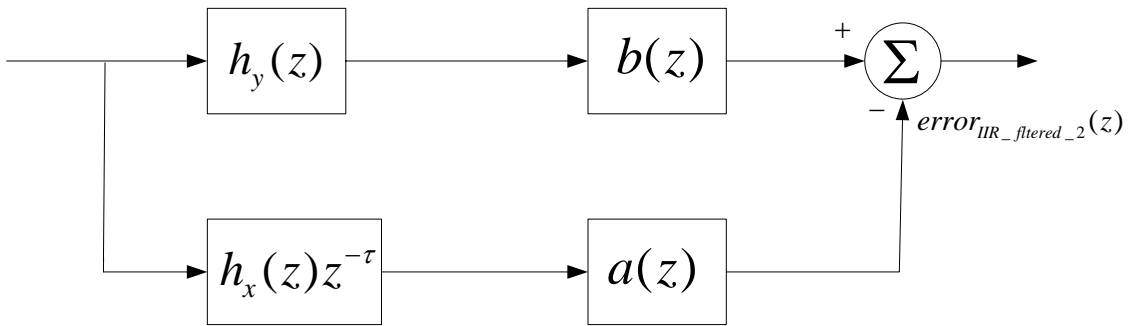
Figure 4.2: Block diagram of the filtered error $error_{IIR\_filtered\_2}(z)$

Our goal is to minimize the filtered error, $error_{IIR\_filtered\_2}(z)$, and its expression

in convolution form is as follows:

$$error_{IIR\_filtered\_2}(n) = a(n) * h_x(n-\tau) - b(n) * h_y(n) \ , \qquad (4.1.8)$$

where '$*$' means convolution; $b(n)$, and $a(n)$ are FIR filters with nb, and na

taps. Let $a_1(0) = 1$, and $u(n) = h_x(n-\tau)$. Equation (4.1.3) can be rewritten as

follows:

$$error_{IIR\_filtered\_2}(n) = u(n) + \sum_{l=1}^{na-1} u(n-l)a_1(l) + \sum_{m=0}^{nb-1} h_y(n-m)\left(-b_{11}(m)\right) \qquad (4.1.9)$$

The above equation can be rewritten in the matrix form as follows:

$$\mathbf{error}_{IIR\_filtered\_2} = \mathbf{u} - \left(\mathbf{U} \cdot (-\mathbf{a}) + \mathbf{H}_y \cdot \mathbf{b}\right), \qquad (4.1.10)$$

where vectors $\mathbf{a}$, and $\mathbf{b}$ are the filter coefficients.

$$\mathbf{a} = \begin{bmatrix} a(1) & a(2) & \cdots & a(na-1) \end{bmatrix}^T ;$$

$$\mathbf{b} = \begin{bmatrix} b(0) & b(1) & \cdots & b(nb-1) \end{bmatrix}^T ;$$

$$\mathbf{u} = \left[\underbrace{0,\cdots,0}_{\tau}, h_x(0), h_x(1), \cdots h_x(M-1), \underbrace{0,\cdots,0}_{L-M-\tau} \right]^T , \quad L = 2M + \max(nb, na) - 2$$

$\mathbf{U}$, and $\mathbf{H}_y$ are the convolution matrices given by

$$\mathbf{U} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ u(0) & 0 & \cdots & 0 \\ u(1) & u(0) & \cdots & \vdots \\ \vdots & u(1) & \cdots & u(0) \\ \vdots & \vdots & \cdots & u(1) \\ \vdots & \vdots & \cdots & \vdots \\ u(L-2) & u(L-3) & \cdots & u(L-na) \end{bmatrix}_{L \times (na-1)} \qquad (4.1.11)$$

$$\mathbf{H}_y = \begin{bmatrix} h_y(0) & 0 & \cdots & 0 \\ h_y(1) & h_y(0) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ h_y(2M-2) & \vdots & \cdots & h_y(0) \\ 0 & h_y(2M-2) & \cdots & h_y(1) \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & h_y(2M-2) \end{bmatrix}_{L \times nb} \tag{4.1.12}$$

Equation (4.1.5) can be rewritten as follows:

$$\mathbf{error}_{IIR\_filtered\_2} = \mathbf{u} - \begin{bmatrix} \mathbf{U} & \mathbf{H}_y \end{bmatrix} \begin{bmatrix} -\mathbf{a} \\ \mathbf{b} \end{bmatrix}$$

Our goal is to minimize $error_{IIR\_filtered\_2}$, so the criterion is as follows:

$$\begin{bmatrix} -\mathbf{a} \\ \mathbf{b} \end{bmatrix} = \arg\min\left\{ \| \mathbf{error}_{filtered\_2} \|^2 \right\}$$
$$= \arg\min\left\{ \| \mathbf{u} - \begin{bmatrix} \mathbf{U} & \mathbf{H}_y \end{bmatrix} \begin{bmatrix} -\mathbf{a} \\ \mathbf{b} \end{bmatrix} \|^2 \right\} \tag{4.1.13}$$

Let $\mathbf{W} = \begin{bmatrix} \mathbf{U} & \mathbf{H}_y \end{bmatrix}$, and $\mathbf{v} = \begin{bmatrix} -\mathbf{a} \\ \mathbf{b} \end{bmatrix}$. We can get the filtered least squares solution as

follows:

$$\mathbf{v} = \left(\mathbf{W}^T \cdot \mathbf{W} \right)^{-1} \cdot \mathbf{W}^T \cdot \mathbf{u} \tag{4.1.14}$$

From the above equation, we can get the filter coefficients as follows:

$$a(n) = \begin{bmatrix} 1 & v(0) & v(1) & \cdots & v(na-2) \end{bmatrix}^T \tag{4.1.15}$$

$$b(n) = \begin{bmatrix} v(na-1) & v(na) & \cdots & v(na+nb-2) \end{bmatrix}^T \tag{4.1.16}$$

4.1.2 Design in Frequency Domain

Similarly, we will try to find the filters in the frequency domain. The direct error

is given by

$$error_{IIR}(\omega) = \frac{B(e^{j\omega})}{A(e^{j\omega})} - \frac{H_x(e^{j\omega})}{H_y(e^{j\omega})} e^{-j\omega\tau} \qquad (4.1.17)$$

In the previous section, we know that the ratio error can be used for stabilization, so we would minimize the ratio error $error_{IIR\_ratio}(\omega)$ in IIR form as follows:

$$
\begin{aligned}
error_{IIR\_ratio}(\omega) &= \frac{B(e^{j\omega})}{A(e^{j\omega})} \cdot error_{IIR}(\omega) \\
&= \frac{B(e^{j\omega})}{A(e^{j\omega})} \cdot \frac{H_y(e^{j\omega})}{H_x(e^{j\omega})} - e^{-j\omega\tau}
\end{aligned}
\qquad (4.1.18)
$$

$$J_{IIR\_ratio} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| \frac{B(e^{j\omega})}{A(e^{j\omega})} \cdot \frac{H_y(e^{j\omega})}{H_x(e^{j\omega})} - e^{-j\omega\tau} \right|^2 d\omega \qquad (4.1.19)$$

In the same way, we will use the vector form to express the Fourier transforms of $B(e^{j\omega})$ and $A(e^{j\omega})$ as follows:

$$B(e^{j\omega}) = \mathbf{b}^T \cdot \mathbf{ex}_1(\omega), \text{ where } \mathbf{ex}_1(\omega) = [e^{-j0\times\omega}, e^{-j1\times\omega}, \cdots, e^{-j(nb-1)\times\omega}]^T \qquad (4.1.20)$$

$$A(e^{j\omega}) = \begin{bmatrix} 1 & \mathbf{a}^T \end{bmatrix} \cdot \begin{bmatrix} 1 \\ \mathbf{ex}_2(\omega) \end{bmatrix}, \text{ where } \mathbf{ex}_2(\omega) = [e^{-j1\times\omega}, \cdots, e^{-j(na-1)\times\omega}]^T \qquad (4.1.21)$$

However, we find it is difficult to solve Equation (4.1.19) because of the inherent nonlinearity duo to $A(e^{j\omega})$. Therefore, $error_{IIR\_ratio}(e^{j\omega})$ is multiplied by $A(e^{j\omega})$ to linearize the problem, and modified into $error_{IIR\_ratio\_filtered}(\omega)$ as

$$error_{IIR\_ratio\_filtered}(\omega) = A(e^{j\omega}) \cdot error_{IIR\_ratio}(\omega) \qquad (4.1.22)$$

Figure 4.3 is its block diagram.
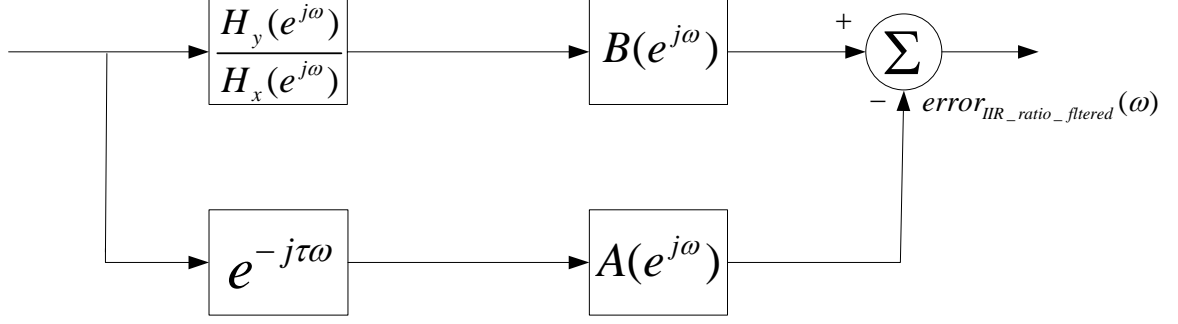
Figure 4.3: Block diagram of the ratio filtered error $error_{IIR\_ratio\_filtered}(\omega)$

Therefore, our goal is to minimize the filtered error energy:

$$J_{IIR\_ratio\_filtered} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \left| B(e^{j\omega}) \cdot \left(\frac{H_y(e^{j\omega})}{H_x(e^{j\omega})} \right) - A(e^{j\omega}) \cdot e^{-j\cdot\tau\cdot\omega} \right|^2 d\omega \qquad (4.1.23)$$

In the same way, $B(e^{j\omega})$ and $A(e^{j\omega})$ can be rewritten in vector form as described in Equation (4.1.20) and Equation (4.1.21). Combining Equation (4.1.20) and Equation (4.1.21) into Equation (4.1.23), we can rewrite Equation (4.1.23) in matrix form as follows:

$$J_{IIR\_ratio\_filtered} = \mathbf{b^T} \cdot \mathbf{Q}_1 \cdot \mathbf{b} - 2 \cdot \mathbf{b^T} \cdot \mathbf{Q}_2 \cdot \mathbf{a} + a^T \cdot \mathbf{Q}_3 \cdot \mathbf{a} - 2 \cdot \mathbf{b^T} \cdot \mathbf{Q}_4 + 2 \cdot \mathbf{a^T} \cdot \mathbf{Q}_5 + 1, \qquad (4.1.24)$$

where $\mathbf{Q}_1 = \dfrac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{ex}_1(\omega) \cdot \left| \dfrac{Hy(e^{j\omega})}{Hx(e^{j\omega})} \right|^2 \cdot \mathbf{ex}_1(\omega)^H d\omega$

$\mathbf{Q}_2 = \dfrac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Re} \left\{ \mathbf{ex}_1(\omega) \cdot \dfrac{Hy(e^{j\omega})}{Hx(e^{j\omega})} \cdot e^{j\cdot\tau\cdot\omega} \cdot \mathbf{ex}_2(\omega)^H \right\} d\omega$

$\mathbf{Q}_3 = \dfrac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{ex}_2(\omega) \cdot \mathbf{ex}_2(\omega)^H d\omega$

$\mathbf{Q}_4 = \dfrac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Re} \left\{ \mathbf{ex}_1(\omega) \cdot \dfrac{Hy(e^{j\omega})}{Hx(e^{j\omega})} \cdot e^{j\cdot\tau\cdot\omega} \right\} d\omega$

$\mathbf{Q}_5 = \dfrac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Re} \left\{ \mathbf{ex}_2(\omega) \right\} d\omega$

In order to minimize the error $J_{IIR\_ratio\_fitlered}$, let $\dfrac{\partial J_{IIR\_ratio\_filtered}}{\partial \mathbf{b}} = \mathbf{0}$ and

$\dfrac{\partial J_{IIR\_ratio\_filtered}}{\partial \mathbf{a}} = \mathbf{0}$ such that:

$$\frac{\partial J_{IIR\_ratio\_filtered}}{\partial \mathbf{b}} = 2 \cdot \left(\mathbf{Q}_1 \cdot \mathbf{b} - \mathbf{Q}_2 \cdot \mathbf{a} - \mathbf{Q}_4 \right) = \mathbf{0} \tag{4.1.24}$$

$$\frac{\partial J_{IIR\_ratio\_filtered}}{\partial \mathbf{a}} = 2 \cdot \left(\mathbf{Q}_3 \cdot \mathbf{a} - \mathbf{Q}_2^{\ T} \cdot \mathbf{b} + \mathbf{Q}_5 \right) = \mathbf{0} \tag{4.1.25}$$

Combine Equation (4.1.24) and Equation (4.1.25) into matrix form such that:

$$\begin{bmatrix} \mathbf{Q}_1 & -\mathbf{Q}_2 \\ -\mathbf{Q}_2^{\ T} & \mathbf{Q}_3 \end{bmatrix} \begin{bmatrix} \mathbf{b} \\ \mathbf{a} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_4 \\ -\mathbf{Q}_5 \end{bmatrix} \tag{4.1.26}$$

Therefore, the solution $\mathbf{a}$ and $\mathbf{b}$ can be found by

$$\begin{bmatrix} \mathbf{b} \\ \mathbf{a} \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_1 & -\mathbf{Q}_2 \\ -\mathbf{Q}_2^{\ T} & \mathbf{Q}_3 \end{bmatrix}^{-1} \cdot \begin{bmatrix} \mathbf{Q}_4 \\ -\mathbf{Q}_5 \end{bmatrix} \tag{4.1.27}$$

4.2 Common-pole Structure

4.2.1 Design in Time Domain

In Section (4.1), the filters of the crosstalk canceller are found from the matrix

inverse G^{-1}. Similar to Section (3.2), we want to find the filters by using direct LSE

method. The discussion here is same as in Section (3.2). Let s_R be an impulse signal,

and s_L be a zero signal. The direct IIR error between the ear signals and desired

signals can be expressed as follows:

$$\mathbf{error}_{IIR\_LSE}(z) = \begin{bmatrix} error_1(z) \\ error_2(z) \end{bmatrix} = G(z) \begin{bmatrix} b_{11}(z)/a_{11}(z) \\ b_{21}(z)/a_{21}(z) \end{bmatrix} - \begin{bmatrix} d(z) \\ 0 \end{bmatrix} \tag{4.2.1}$$

Therefore, the direct least square error criterion is such that:

$$\begin{bmatrix} b_{11}(z)/a_{11}(z) \\ b_{21}(z)/a_{21}(z) \end{bmatrix} = \arg\min \left\{ \| \mathbf{error}_{IIR\_LSE}(z) \|^2 \right\}, \tag{4.2.2}$$

where $h_{IIR\_1}(z) = \begin{bmatrix} \dfrac{b_{11}(z)}{a_{11}(z)} & \dfrac{b_{21}(z)}{a_{21}(z)} \end{bmatrix}^T$

Equation (4.2.2) can be rewritten as follows:

$$\begin{bmatrix} b_{11}(z)/a_{11}(z) \\ b_{21}(z)/a_{21}(z) \end{bmatrix} = \arg\min \left\{ \sum_{n=0}^{\infty} |error_1(n)|^2 + |error_2(n)|^2 \right\} \tag{4.2.3}$$

Let $q_1(z) = \dfrac{1}{a_{11}(z)}$ and $q_2(z) = \dfrac{1}{a_{21}(z)}$, and they can express in the time domain as follows:

$$q_1(n) = \delta(n) - \sum_{k=1}^{na_1-1} a_{11}(k)q_1(n-k)$$

$$q_2(n) = \delta(n) - \sum_{k=1}^{na_1-1} a_{21}(k)q_2(n-k)$$

Therefore, $error_1(n)$ and $error_2(n)$ can be written as follows:

$$error_1(n) = g_{rr}(n) * b_{11}(n) * q_1(n) + g_{lr}(n) * b_{21}(n) * q_2(n) - d(n)$$
$$= \sum_{t=0}^{nb_{11}-1} b_{11}(t) \cdot \left(g_{rr}(n-t) + \sum_{l=0}^{\infty} (\sum_{k=1}^{na_{11}-1} a_{11}(k)q_1(l-k)) \cdot g_{rr}(n-t-l) \right) +$$
$$\sum_{t=0}^{nb_{21}-1} b_{21}(t) \cdot \left(g_{lr}(n-t) + \sum_{l=0}^{\infty} (\sum_{k=1}^{na_{21}-1} a_{21}(k)q_2(l-k)) \cdot g_{lr}(n-t-l) \right) - d(n)$$

$$error_2(n) = g_{rl}(n) * b_{11}(n) * q_1(n) + g_{ll}(n) * b_{21}(n) * q_2(n)$$
$$= \sum_{t=0}^{nb_{11}-1} b_{11}(t) \cdot \left(g_{rl}(n-t) + \sum_{l=0}^{\infty} (\sum_{k=1}^{na_{11}-1} a_{11}(k)q_1(l-k)) \cdot g_{rl}(n-t-l) \right) +$$
$$\sum_{t=0}^{nb_{21}-1} b_{21}(t) \cdot \left(g_{ll}(n-t) + \sum_{l=0}^{\infty} (\sum_{k=1}^{na_{21}-1} a_{21}(k)q_2(l-k)) \cdot g_{ll}(n-t-l) \right)$$

We can find Equation (4.2.3) is a function of $a_{11}(k)$, $a_{21}(k)$, $b_{11}(t)$, and $b_{21}(t)$. To minimize it, we should differentiate it with respect to $a_{11}(k)$, $a_{21}(k)$, $b_{11}(t)$ and $b_{21}(t)$, and set the derivatives to zeros to find $a_{11}(k)$, $a_{21}(k)$, $b_{11}(t)$ and $b_{21}(t)$. Similarly, a lot of nonlinear equations will appear. It would be a very tough work to

solve these equations. Therefore, minimization of a filtered error can be formulated as

$$\mathbf{error}_{IIR\_LSE\_filtered1}(z) = a_{11}(z) \cdot a_{21}(z) \cdot \mathbf{error}_{IIR\_LSE}(z)$$
$$= G(z) \cdot \begin{bmatrix} b_{11}(z) \cdot a_{21}(z) \\ b_{21}(z) \cdot a_{11}(z) \end{bmatrix} - a_{11}(z) \cdot a_{21}(z) \cdot \begin{bmatrix} d(z) \\ 0 \end{bmatrix} \qquad (4.2.4)$$

Equation (4.2.2) can be rewritten as follows:

$$h_{IIR\_LSE\_filtered1}(z) = \arg\min\left\{ \| G(z) \cdot \begin{bmatrix} b_{11}(z) \cdot a_{21}(z) \\ b_{21}(z) \cdot a_{11}(z) \end{bmatrix} - a_{11}(z) \cdot a_{21}(z) \cdot \begin{bmatrix} d(z) \\ 0 \end{bmatrix} \|^2 \right\} \quad (4.2.5)$$

However, it is also difficult to solve the solution of Equation (4.2.5) because of

nonlinearity due to multiplication of $a_{11}(z)$ and $a_{21}(z)$. Therefore, we propose a

new structure to linearize the problem. Our approach is to impose a common-pole

constraint:

$$a_{11}(z) = a_{21}(z) = a_1(z) \qquad (4.2.6)$$

$$a_{12}(z) = a_{22}(z) = a_2(z) \qquad (4.2.7)$$

The new structure is called common-pole model and its block diagram is plotted

in Figure 4.4.
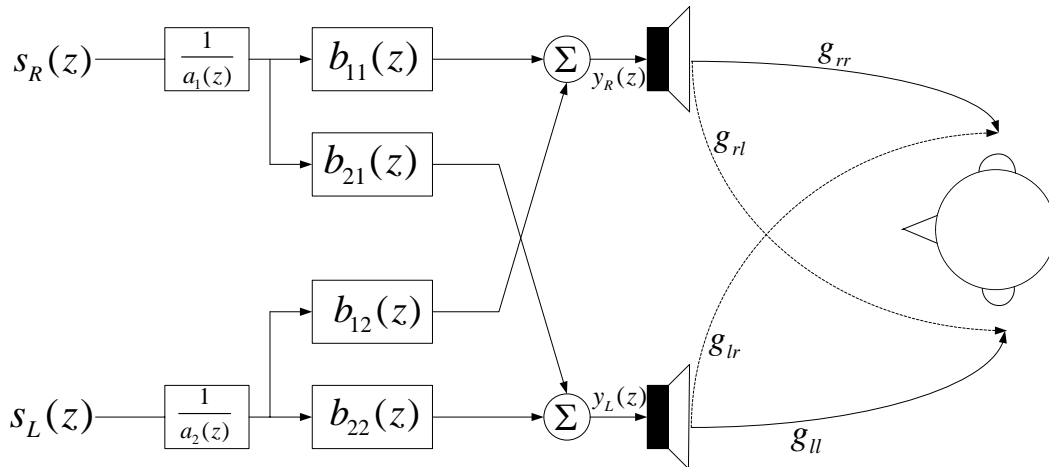
Figure 4.4: Common-pole model of the crosstalk canceller

After combining Equation (4.2.6), the error which we minimize is as follows:

$$\mathbf{error}_{IIR\_LSE\_filtered\,2} = \begin{bmatrix} error_{IIR\_filtered\,2\_1}(z) \\ error_{IIR\_filtered\,2\_2}(z) \end{bmatrix}$$

$$= G(z) \cdot \begin{bmatrix} b_{11}(z) \\ b_{21}(z) \end{bmatrix} - a_1(z) \cdot \begin{bmatrix} d(z) \\ 0 \end{bmatrix} \qquad (4.2.8)$$

Equation (4.2.3) can be rewritten as follows:

$$\left(\frac{b_{11}(z)}{a_1(z)} \quad \frac{b_{21}(z)}{a_1(z)} \right) = \arg\min \left\{ \| \mathbf{error}_{LSE\_IIR\_filtered\,2}(z) \|^2 \right\} \qquad (4.2.9)$$

The above equation can be described as follows:

$$error_{IIR\_filtered\,2\_1}(z) = g_{rr}(z)b_{11}(z) + g_{lr}(z)b_{21}(z) - a_1(z)d(z) \qquad (4.2.10)$$

$$error_{IIR\_filtered\,2\_2}(z) = g_{rl}(z)b_{11}(z) + g_{ll}(z)b_{21}(z) \qquad (4.2.11)$$

Then we rewrite the above two equations in time domain. We modify Equation (4.2.10) first as follows:

$$error_{IIR\_filtered\,2\_1}(n) = a_1(n) * d(n) - (g_{rr}(n) * b_{11}(n) + g_{lr}(n) * b_{21}(n)) \quad , \qquad (4.2.12)$$

where b_{11}, b_{21}, and a_1 are FIR filters with nb_{11}, nb_{21}, and na_1 taps. Let $a_1(0) = 1$,

and Equation (4.2.10) can be rewritten as follows:

$$error_{IIR\_filtered\,2\_1}(n) = d(n) - \left(\sum_{l=1}^{na_1-1} d(n-l)\left(-a_1(l)\right) + \sum_{m=0}^{nb_{11}-1} g_{rr}(n-m)b_{11}(m) + \sum_{q=0}^{nb_{21}-1} g_{lr}(n-q)b_{21}(q) \right)$$

(4.2.13)

In the matrix form we have

$$\mathbf{error}_{IIR\_filtered\,2\_1} = \mathbf{d} - \left(\mathbf{D} \cdot \left(-\mathbf{a}_1 \right) + \mathbf{G}_{rr} \cdot \mathbf{b}_{11} + \mathbf{G}_{lr} \cdot \mathbf{b}_{21} \right), \qquad (4.2.14)$$

where vectors $\mathbf{a}_1$, $\mathbf{b}_{11}$, and $\mathbf{b}_{21}$ are the filter coefficients

43

$$\mathbf{a}_1 = \begin{bmatrix} a_1(1) & a_1(2) & \cdots & a_1(na_1 - 1) \end{bmatrix}^T;$$

$$\mathbf{b}_{11} = \begin{bmatrix} b_{11}(0) & b_{11}(1) & \cdots & b_{11}(nb_{11} - 1) \end{bmatrix}^T;$$

$$\mathbf{b}_{21} = \begin{bmatrix} b_{21}(0) & b_{21}(1) & \cdots & b_{21}(nb_{21} - 1) \end{bmatrix}^T;$$

and $\mathbf{d} = \begin{bmatrix} d(0) & d(1) & \cdots & d(L-1) \end{bmatrix}^T$, where $L = M + \max(nb_{11}, nb_{21}, na_1) - 1$,

represents the desired signal.

$\mathbf{D}$, $\mathbf{G}_{rr}$, and $\mathbf{G}_{lr}$ are the convolution matrices given by

$$\mathbf{D} = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ d(0) & 0 & \cdots & 0 \\ d(1) & d(0) & \cdots & \vdots \\ \vdots & d(1) & \cdots & d(0) \\ \vdots & & \cdots & d(1) \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & & & \\ d(L-2) & d(L-3) & \cdots & d(L-na_1) \end{bmatrix}_{L \times (na_1 - 1)} \tag{4.2.15}$$

$$\mathbf{G}_{rr} = \begin{bmatrix} g_{rr}(0) & 0 & \cdots & 0 \\ g_{rr}(1) & g_{rr}(0) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ g_{rr}(M-1) & \vdots & \cdots & g_{rr}(0) \\ 0 & g_{rr}(M-1) & \cdots & g_{rr}(1) \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & g_{rr}(M-1) \end{bmatrix}_{L \times nb_{11}} \tag{4.2.16}$$

$$\mathbf{G}_{lr} = \begin{bmatrix} g_{lr}(0) & 0 & \cdots & 0 \\ g_{lr}(1) & g_{lr}(0) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ g_{lr}(M-1) & \vdots & \cdots & g_{lr}(0) \\ 0 & g_{lr}(M-1) & \cdots & g_{lr}(1) \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & g_{lr}(M-1) \end{bmatrix}_{L \times nb_{21}} \tag{4.2.17}$$

Equation (4.2.14) can be rewritten as

$$\mathbf{error}_{filtered2\_1} = \mathbf{d} - \begin{bmatrix} \mathbf{D} & \mathbf{G}_{rr} & \mathbf{G}_{lr} \end{bmatrix} \begin{bmatrix} -\mathbf{a}_1 \\ \mathbf{b}_{11} \\ \mathbf{b}_{21} \end{bmatrix} \tag{4.2.18}$$

$error_{IIR\_filtered2\_2}(z)$ in Equation (4.2.11) can be further modified as

$$\begin{aligned} error_{IIR\_filtered2\_2}(n) &= g_{rl}(n) * b_{11}(n) + g_{ll}(n) * b_{21}(n) \\ &= \sum_{k=0}^{nb_{11}-1} g_{rl}(n-k)b_{11}(k) + \sum_{q=0}^{nb_{21}-1} g_{ll}(n-q)b_{21}(q) \end{aligned} \tag{4.2.19}$$

In convolution matrix, it becomes

$$\mathbf{error}_{IIR\_filtered2\_2} = \mathbf{G}_{rl} \cdot \mathbf{b}_{11} + \mathbf{G}_{ll} \cdot \mathbf{b}_{21} \tag{4.2.20}$$

with $\mathbf{G}_{rl}$ and $\mathbf{G}_{ll}$ being the convolution matrices

$$\mathbf{G}_{rl} = \begin{bmatrix} g_{rl}(0) & 0 & \cdots & 0 \\ g_{rl}(1) & g_{rl}(0) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ g_{rl}(M-1) & \vdots & \cdots & g_{rl}(0) \\ 0 & g_{rl}(M-1) & \cdots & g_{rl}(1) \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & g_{rl}(M-1) \end{bmatrix}_{L \times nb_{11}} \tag{4.2.21}$$

$$\mathbf{G}_{ll} = \begin{bmatrix} g_{ll}(0) & 0 & \cdots & 0 \\ g_{ll}(1) & g_{ll}(0) & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ g_{ll}(M-1) & \vdots & \cdots & g_{ll}(0) \\ 0 & g_{ll}(M-1) & \cdots & g_{ll}(1) \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & g_{ll}(M-1) \end{bmatrix}_{L \times nb_{21}} \tag{4.2.22}$$

Equation (4.2.20) can be written as

$$\mathbf{error}_{IIR\_filtered2\_2} = \begin{bmatrix} \mathbf{0} & \mathbf{G}_{rl} & \mathbf{G}_{ll} \end{bmatrix} \begin{bmatrix} -\mathbf{a}_1 \\ \mathbf{b}_{11} \\ \mathbf{b}_{21} \end{bmatrix} \qquad (4.2.23)$$

Our goal is to minimize $error_{IIR\_filtered2\_1}$ and $error_{IIR\_filtered2\_2}$, so the combined criterion is given by

$$\left(\frac{b_{11}(z)}{a_1(z)} \quad \frac{b_{21}(z)}{a_1(z)} \right) = \arg\min \left\{ \left\| \begin{matrix} \mathbf{error}_{IIR\_filtered2\_1} \\ \mathbf{error}_{IIR\_filtered2\_2} \end{matrix} \right\|^2 \right\} \qquad (4.2.24)$$

Using Equation (4.2.18) and Equation (4.2.23), Equation (4.2.24) becomes

$$\left(\frac{b_{11}(z)}{a_1(z)} \quad \frac{b_{21}(z)}{a_1(z)} \right) = \arg\min \left\{ \left\| \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{D} & \mathbf{G}_{rr} & \mathbf{G}_{lr} \\ \mathbf{0} & \mathbf{G}_{rl} & \mathbf{G}_{ll} \end{bmatrix} \begin{bmatrix} -\mathbf{a}_1 \\ \mathbf{b}_{11} \\ \mathbf{b}_{21} \end{bmatrix} \right\|^2 \right\} \qquad (4.2.25)$$

Let $\mathbf{G}_1 = \begin{bmatrix} \mathbf{D} & \mathbf{G}_{rr} & \mathbf{G}_{lr} \\ \mathbf{0} & \mathbf{G}_{rl} & \mathbf{G}_{ll} \end{bmatrix}$ and $\mathbf{v}_1 = \begin{bmatrix} -\mathbf{a}_1 & \mathbf{b}_{11} & \mathbf{b}_{21} \end{bmatrix}^T$

Equation (4.2.25) can be solved as follows:

$$\mathbf{v}_1 = \left(\mathbf{G}_1^T \mathbf{G}_1 \right)^{-1} \mathbf{G}_1^T \begin{bmatrix} \mathbf{d} \\ \mathbf{0} \end{bmatrix} \qquad (4.2.26)$$

We can get the filter coefficients in Equation (4.2.26) as follows:

$$a_1 = \begin{bmatrix} 1 & -v_1(0) & -v_1(1) & \cdots & -v_1(na_1 - 2) \end{bmatrix} \qquad (4.2.27)$$

$$b_{11} = \begin{bmatrix} v_1(na_1 - 1) & v_1(na_1) & \cdots & v_1(nb_{11} + na_1 - 2) \end{bmatrix} \qquad (4.2.28)$$

$$b_{21} = \begin{bmatrix} v_1(nb_{11} + na_1 - 1) & v_1(nb_{11} + na_1) & \cdots & v_1(nb_{21} + nb_{11} + na_1 - 2) \end{bmatrix} \qquad (4.2.29)$$

In the same way, we can find out a_2, b_{12}, and b_{22} with na_2, nb_{12}, and nb_{22} taps.

$$\mathbf{v}_2 = \left(\mathbf{G}_2^T \mathbf{G}_2 \right)^{-1} \mathbf{G}_2^T \begin{bmatrix} \mathbf{0} \\ \mathbf{d} \end{bmatrix}, \qquad (4.2.30)$$

where $\mathbf{G}_2 = \begin{bmatrix} \mathbf{0} & \mathbf{G}_{rr2} & \mathbf{G}_{lr2} \\ \mathbf{D}_2 & \mathbf{G}_{rl2} & \mathbf{G}_{ll2} \end{bmatrix}$ and $\mathbf{v}_2 = \begin{bmatrix} -\mathbf{a}_2 & \mathbf{b}_{12} & \mathbf{b}_{22} \end{bmatrix}^T$. The terms in $\mathbf{G}_2$ can

be found by using similar method obtaining $\mathbf{G}_1$.

$$a_2 = \begin{bmatrix} 1 & -v_2(0) & -v_2(1) & \cdots & -v_2(na_2 - 2) \end{bmatrix} \qquad (4.2.31)$$

$$b_{12} = \begin{bmatrix} v_2(na_2 - 1) & v_2(na_2) & \cdots & v_2(nb_{12} + na_2 - 2) \end{bmatrix} \qquad (4.2.32)$$

$$b_{22} = \begin{bmatrix} v_2(nb_{12} + na_2 - 1) & v_2(nb_{12} + na_2) & \cdots & v_2(nb_{22} + nb_{12} + na_2 - 2) \end{bmatrix} \qquad (4.2.33)$$

From Equation (4.2.24), if the order of the common pole $\mathbf{a}_1$ is zero, we can find

it is the same as the LSE FIR design. The LSE FIR is the special case of the

common-pole IIR model.

4.2.2 Design in Frequency Domain

Because of nonlinearity, it is also difficult to handle $error_{IIR\_LSE}$ and

$error_{IIR\_LSE\_filtered1}$ in the frequency domain. We minimize the error $error_{IIR\_LSE\_filtered2}$

in the frequency. We replace Equation (4.2.8) in the frequency domain, and the error

is expressed as follows:

$$error_{IIR\_LSE\_filtered2}(\omega) = \begin{bmatrix} g_{rr}(e^{j\omega}) & g_{lr}(e^{j\omega}) \\ g_{rl}(e^{j\omega}) & g_{ll}(e^{j\omega}) \end{bmatrix} \begin{bmatrix} B_{11}(e^{j\omega}) \\ B_{21}(e^{j\omega}) \end{bmatrix} - A_1(e^{j\omega}) \begin{bmatrix} e^{-j\tau\omega} \\ 0 \end{bmatrix}, \qquad (4.2.34)$$

where $B_{11}(e^{j\omega})$, $B_{21}(e^{j\omega})$ and $A_1(e^{j\omega})$ are the Fourier transforms of $b_{11}(n)$,

$b_{21}(n)$ and $a_1(n)$. Our goal is to minimize the error norm and its formulation is as

follows:

$$J_{IIR\_LSE\_filtered2} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \| \begin{bmatrix} g_{rr}(e^{j\omega}) & g_{lr}(e^{j\omega}) \\ g_{rl}(e^{j\omega}) & g_{ll}(e^{j\omega}) \end{bmatrix} \begin{bmatrix} B_{11}(e^{j\omega}) \\ B_{21}(e^{j\omega}) \end{bmatrix} - A_1(e^{j\omega}) \begin{bmatrix} e^{-j\tau\omega} \\ 0 \end{bmatrix} \|^2 \, d\omega \qquad (4.2.35)$$

In the same way, the filters of Fourier transforms are replaced with vector forms as follows:

$$B_{11}(e^{j\omega}) = \mathbf{b}_{11}^{T} \cdot \mathbf{ex}_1(\omega), \text{ where } \mathbf{ex}_1(\omega) = [e^{-j0\times\omega}, e^{-j1\times\omega}, \cdots, e^{-j(nb_{11}-1)\times\omega}]^{T} \qquad (4.2.36)$$

$$B_{21}(e^{j\omega}) = \mathbf{b}_{21}^{T} \cdot \mathbf{ex}_2(\omega), \text{ where } \mathbf{ex}_2(\omega) = [e^{-j0\times\omega}, e^{-j1\times\omega}, \cdots, e^{-j(nb_{21}-1)\times\omega}]^{T} \qquad (4.2.37)$$

Here, the first coefficient of filter $a(n)$ is also set to 1, and its Fourier transform expressed in vector form is as follows:

$$A_1(e^{j\omega}) = \begin{bmatrix} 1 & \mathbf{a}_1^{T} \end{bmatrix} \cdot \begin{bmatrix} 1 \\ \mathbf{ex}_3(\omega) \end{bmatrix}, \text{ where } \mathbf{ex}_3(\omega) = [e^{-j1\times\omega}, \cdots, e^{-j(na_1-1)\times\omega}]^{T} \qquad (4.2.38)$$

Combining Equation (4.2.36) to Equation (4.2.38) into Equation (4.2.35), the error norm can be rewritten as follows:

$$\begin{aligned} J_{IIR\_LSE\_filtered2} &= \mathbf{a}_1^{T} \cdot \mathbf{M}_1 \cdot \mathbf{a}_1 - 2 \cdot \mathbf{b}_{11}^{T} \cdot \mathbf{M}_2 \cdot \mathbf{a}_1 - 2 \cdot \mathbf{b}_{21}^{T} \cdot \mathbf{M}_3 \cdot \mathbf{a}_1 + \mathbf{b}_{11}^{T} \cdot \mathbf{M}_4 \cdot \mathbf{b}_{11} + \\ &\quad 2 \cdot \mathbf{b}_{11}^{T} \cdot \mathbf{M}_5 \cdot \mathbf{b}_{11} + \mathbf{b}_{21} \cdot \mathbf{M}_6 \cdot \mathbf{b}_{21} + 2 \cdot \mathbf{a}_1^{T} \cdot \mathbf{M}_7 - 2 \cdot \mathbf{b}_{11}^{T} \cdot \mathbf{M}_8 - 2 \cdot \mathbf{b}_{21}^{T} \cdot \mathbf{M}_9 + 1 \end{aligned} \qquad (4.2.39)$$

Where

$$\mathbf{M}_1 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{ex}_3 \mathbf{ex}_3^{H} d\omega$$

$$\mathbf{M}_2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Re}\left\{ \mathbf{ex}_1^{*} \cdot e^{-j\cdot\tau\cdot\omega} \cdot g_{rr}^{*}(e^{j\omega}) \cdot \mathbf{ex}_3^{T} \right\} d\omega$$

$$\mathbf{M}_3 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Re}\left\{ \mathbf{ex}_2^{*} \cdot e^{-j\cdot\tau\cdot\omega} \cdot g_{lr}^{*}(e^{j\omega}) \cdot \mathbf{ex}_3^{T} \right\} d\omega$$

$$\mathbf{M}_4 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{ex}_1^{*} \cdot (\left| g_{rr}(e^{j\omega}) \right|^2 + \left| g_{rl}(e^{j\omega}) \right|^2) \cdot \mathbf{ex}_1^{T} d\omega$$

$$\mathbf{M}_5 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Re}\left\{ \mathbf{ex}_1^{*} \cdot \left(g_{rr}^{*}(e^{j\omega}) \cdot g_{lr}(e^{j\omega}) + g_{rl}^{*}(e^{j\omega}) \cdot g_{ll}(e^{j\omega}) \right) \cdot \mathbf{ex}_2^{T} \right\} d\omega$$

$$\mathbf{M}_6 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathbf{ex}_2^{*} \cdot (\left| g_{lr}(e^{j\omega}) \right|^2 + \left| g_{ll}(e^{j\omega}) \right|^2) \cdot \mathbf{ex}_2^{T} d\omega$$

$$\mathbf{M}_7 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Re}\left\{ \mathbf{ex}_3 \right\} d\omega$$

48

$$\mathbf{M}_8 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Re}\left\{ e^{-j\tau\omega} \cdot g_{rr}^{\ *}(e^{j\omega}) \cdot \mathbf{ex}_1^{\ *} \right\} d\omega$$

$$\mathbf{M}_9 = \frac{1}{2\pi} \int_{-\pi}^{\pi} \mathrm{Re}\left\{ e^{-j\tau\omega} \cdot g_{lr}^{\ *}(e^{j\omega}) \cdot \mathbf{ex}_2^{\ *} \right\} d\omega$$

In order to minimize the error $J_{IIR\_LSE\_filtered2}$, let $\dfrac{\partial J_{IIR\_LSE\_filtered2}}{\partial \mathbf{b}_{11}} = \mathbf{0}$,

$\dfrac{\partial J_{IIR\_LSE\_filtered2}}{\partial \mathbf{b}_{21}} = \mathbf{0}$ and $\dfrac{\partial J_{IIR\_LSE\_filtered2}}{\partial \mathbf{a}_1} = \mathbf{0}$ such that:

$$\frac{\partial J_{IIR\_LSE\_filtered2}}{\partial \mathbf{a}_1} = 2 \cdot \left(\mathbf{M}_1 \cdot \mathbf{a}_1 - \mathbf{M}_2^{\ T} \cdot \mathbf{b}_{11} - \mathbf{M}_3^{\ T} \cdot \mathbf{b}_{21} + \mathbf{M}_7 \right) = \mathbf{0} \qquad (4.2.40)$$

$$\frac{\partial J_{IIR\_LSE\_filtered2}}{\partial \mathbf{b}_{11}} = 2 \cdot \left(-\mathbf{M}_2 \cdot \mathbf{a}_1 + \mathbf{M}_4 \cdot \mathbf{b}_{11} + \mathbf{M}_5 \cdot \mathbf{b}_{21} - \mathbf{M}_8 \right) = \mathbf{0} \qquad (4.2.41)$$

$$\frac{\partial J_{IIR\_LSE\_filtered2}}{\partial \mathbf{b}_{21}} = 2 \cdot \left(-\mathbf{M}_3 \cdot \mathbf{a}_1 + \mathbf{M}_5^{\ T} \cdot \mathbf{b}_{11} + \mathbf{M}_6^{\ T} \cdot \mathbf{b}_{21} - \mathbf{M}_9 \right) = \mathbf{0} \qquad (4.2.42)$$

Combining the above three equations, we can write them as follows:

$$\begin{bmatrix} \mathbf{M}_1 & \mathbf{M}_2^{\ T} & \mathbf{M}_3^{\ T} \\ \mathbf{M}_2 & \mathbf{M}_4 & \mathbf{M}_5 \\ \mathbf{M}_3 & \mathbf{M}_5^{\ T} & \mathbf{M}_6 \end{bmatrix} \begin{bmatrix} -\mathbf{a}_1 \\ \mathbf{b}_{11} \\ \mathbf{b}_{21} \end{bmatrix} = \begin{bmatrix} \mathbf{M}_7 \\ \mathbf{M}_8 \\ \mathbf{M}_9 \end{bmatrix} \qquad (4.2.43)$$

Therefore, the coefficients of the IIR filters can be solved as the following equation.

$$\mathbf{v}_1 = \begin{bmatrix} -\mathbf{a}_1 \\ \mathbf{b}_{11} \\ \mathbf{b}_{21} \end{bmatrix} = \begin{bmatrix} \mathbf{M}_1 & \mathbf{M}_2^{\ T} & \mathbf{M}_3^{\ T} \\ \mathbf{M}_2 & \mathbf{M}_4 & \mathbf{M}_5 \\ \mathbf{M}_3 & \mathbf{M}_5^{\ T} & \mathbf{M}_6 \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{M}_7 \\ \mathbf{M}_8 \\ \mathbf{M}_9 \end{bmatrix} \qquad (4.2.44)$$

We can find each filter as follows:

$$a_1 = \begin{bmatrix} 1 & -v_1(0) & -v_1(1) & \cdots & -v_1(na_1 - 2) \end{bmatrix} \qquad (4.2.45)$$

$$b_{11} = \begin{bmatrix} v_1(na_1 - 1) & v_1(na_1) & \cdots & v_1(nb_{11} + na_1 - 2) \end{bmatrix} \qquad (4.2.46)$$

$$b_{21} = \begin{bmatrix} v_1(nb_{11} + na_1 - 1) & v_1(nb_{11} + na_1) & \cdots & v_1(nb_{11} + nb_{21} + na_1 - 2) \end{bmatrix} \qquad (4.2.47)$$

4.2.3 Comparison between IIR Designs in Time and Frequency Domain

In the same way, the error is minimized just in different domains. Therefore, the results of these two methods are the same. The proof is as follows:

First, Equation (4.2.26) can be decomposed two parts, $\mathbf{G}^T\mathbf{G}$ and $\mathbf{G}^T\begin{bmatrix}\mathbf{d}\\\mathbf{0}\end{bmatrix}$, and they are rewritten as follows:

$$\mathbf{G}^T\mathbf{G} = \begin{bmatrix} \mathbf{M}_1' & \mathbf{M}_2'^T & \mathbf{M}_3'^T \\ \mathbf{M}_2' & \mathbf{M}_4' & \mathbf{M}_5' \\ \mathbf{M}_3' & \mathbf{M}_5'^T & \mathbf{M}_6' \end{bmatrix} \quad \text{and} \quad \mathbf{G}^T\begin{bmatrix}\mathbf{d}\\\mathbf{0}\end{bmatrix} = \begin{bmatrix} \mathbf{M}_7' \\ \mathbf{M}_8' \\ \mathbf{M}_9' \end{bmatrix}, \qquad (4.2.48)$$

where

$$\mathbf{M}_1' = \mathbf{D}^T\mathbf{D} ; \quad \mathbf{M}_2' = \mathbf{G}_{rr}^T \cdot \mathbf{D} ; \quad \mathbf{M}_3' = \mathbf{G}_{lr}^T \cdot \mathbf{D} ; \quad \mathbf{M}_4' = \mathbf{G}_{rr}^T\mathbf{G}_{rr} + \mathbf{G}_{rl}^T\mathbf{G}_{rl}$$

$$\mathbf{M}_5' = \mathbf{G}_{rr}^T\mathbf{G}_{lr} + \mathbf{G}_{rl}^T\mathbf{G}_{ll} ; \quad \mathbf{M}_6' = \mathbf{G}_{lr}^T\mathbf{G}_{lr} + \mathbf{G}_{ll}^T\mathbf{G}_{ll} ; \quad \mathbf{M}_7' = \mathbf{D}^T\mathbf{d}$$

$$\mathbf{M}_8' = \mathbf{G}_{rr}^T\mathbf{d} ; \quad \mathbf{M}_9' = \mathbf{G}_{lr}^T\mathbf{d}$$

We can find Equation (4.2.44) is similar to Equation (4.2.48), so we will compare each term in the two equations. First, $\mathbf{M}_1'$ is considered, and each element is calculated as follows:

$$\begin{aligned} \mathbf{M}_1'(i,j) &= \sum_{k=0}^{L-1} D(i,k)D(k,j) \\ &= \sum_{k=0}^{L-1} \delta(k-(\tau+i+1)) \cdot \delta(k-(\tau+j+1)) \\ &= \begin{cases} 1, & i = j \\ 0, & i \neq j \end{cases} \end{aligned}$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}_{(na_1-1)\times(na_1-1)} \tag{4.2.49}$$

We can find that $\mathbf{M}_1'$ is the identity matrix. Similarly, $\mathbf{M}_1$ is calculated as follows:

$$\mathbf{M}_1 = \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathbf{ex}_3\mathbf{ex}_3{}^H d\omega$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} \begin{bmatrix} e^{-j1\times\omega} \\ \vdots \\ e^{-j(na_1-1)\times\omega} \end{bmatrix} [e^{j1\times\omega}, \cdots, e^{j(na_1-1)\times\omega}]d\omega$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \vdots \\ \vdots & 0 & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{bmatrix}_{(na_1-1)\times(na_1-1)} \tag{4.2.50}$$

We can find $\mathbf{M}_1'$ and $\mathbf{M}_1$ are identity matrixes with same dimension. Then, $\mathbf{M}_2'$ is considered, and each term of $\mathbf{M}_2'$ is calculated as follows:

$$\mathbf{M}_2'(i,j) = \sum_{k=0}^{L-1}\mathbf{G}_{rr}(k,i)\cdot\mathbf{D}(k,j),\ i=0\sim nb_{11}\text{-}1 \text{ and } j=0\sim na_1-2$$

$$= \sum_{k=0}^{L-1}g_{rr}(k-i)\cdot\delta(k-j-\tau-1)$$

$$= g_{rr}(j+\tau+1-i)$$

$$\mathbf{M}_2' = \begin{bmatrix} g_{rr}(\tau+1) & g_{rr}(\tau+2) & \cdots & g_{rr}(\tau+na_1-1) \\ g_{rr}(\tau) & g_{rr}(\tau+1) & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ g_{rr}(\tau+2-nb_{11}) & \cdots & \cdots & g_{rr}(\tau+na_1-nb_{11}-1) \end{bmatrix} \tag{4.2.51}$$

$\mathbf{M}_2$ is calculated as follows:

51

$$\mathbf{M}_2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{Re}\left\{ \mathbf{ex}_1^{\,*}\cdot e^{-j\cdot\tau\cdot\omega}\cdot g_{rr}^{\,*}(e^{j\omega})\cdot \mathbf{ex}_3^{\,T}\right\} d\omega$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{Re}\left\{ e^{-j\cdot\tau\cdot\omega}\cdot g_{rr}^{\,*}(e^{j\omega})\cdot\begin{bmatrix} e^{j0\omega}\\ \vdots\\ e^{j(nb_{11}-1)\omega}\end{bmatrix}\cdot\begin{bmatrix} e^{-j\omega} & \cdots & e^{-j(na_1-1)\omega}\end{bmatrix}\right\} d\omega$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{Re}\left\{ g_{rr}^{\,*}(e^{j\omega})\cdot\begin{bmatrix} e^{-j(1+\tau)\omega} & e^{-j(2+\tau)\omega} & \cdots & e^{-j(na_1+\tau-1)\omega}\\ e^{-j\tau\omega} & e^{-j(1+\tau)\omega} & \cdots & \vdots\\ \vdots & \vdots & \ddots & \vdots\\ e^{-j(\tau-nb_{11}+2)\omega} & \cdots & \cdots & e^{-j(na_1+\tau-nb_{11}-1)\omega}\end{bmatrix}\right\} d\omega$$

$$= \begin{bmatrix} g_{rr}(\tau+1) & g_{rr}(\tau+2) & \cdots & g_{rr}(\tau+na_1-1)\\ g_{rr}(\tau) & g_{rr}(\tau+1) & \cdots & \vdots\\ \vdots & \vdots & \ddots & \vdots\\ g_{rr}(\tau+2-nb_{11}) & \cdots & \cdots & g_{rr}(\tau+na_1-nb_{11}-1)\end{bmatrix} \tag{4.2.52}$$
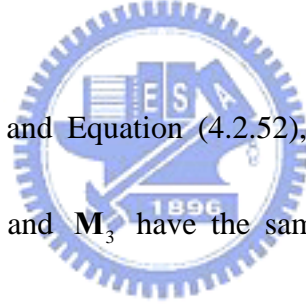
From Equation (4.2.51) and Equation (4.2.52), it is proved that $\mathbf{M}_2{}'$ equals $\mathbf{M}_2$. We can find that $\mathbf{M}_3{}'$ and $\mathbf{M}_3$ have the same forms as $\mathbf{M}_2{}'$ and $\mathbf{M}_2$, so they can also be proved to be equal in the same way. Next, the forms of $\mathbf{M}_4{}'$, $\mathbf{M}_5{}'$, $\mathbf{M}_6{}'$, $\mathbf{M}_4$, $\mathbf{M}_5$ and $\mathbf{M}_6$ are the same as those expressed in Section (4.2.3). Therefore, the same technique can be applied to prove that they are equal. We consider $\mathbf{M}_7{}'$ and $\mathbf{M}_7$, and each element is expressed as follows:

$$\mathbf{M}_7{}'(i) = \sum_{k=0}^{L-1}\mathbf{D}(k,i)\delta(k-\tau) = \sum_{k=0}^{L-1}\delta(k-\tau-i-1)\delta(k-\tau) = 0,\ i = 0 \sim na_1 - 2 \tag{4.2.53}$$

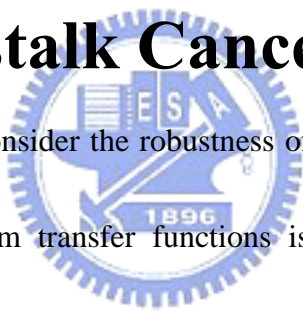We can know that $\mathbf{M}_7{}'$ is a zero vector with dimension $na_1 - 1$.

$$\mathbf{M}_7 = \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{Re}\left\{\mathbf{ex}_3\right\} d\omega = \frac{1}{2\pi}\int_{-\pi}^{\pi}\mathrm{Re}\left\{\begin{bmatrix} e^{-j\omega}\\ \vdots\\ e^{-j(na_1-1)\omega}\end{bmatrix}\right\} d\omega = \begin{bmatrix} 0\\ \vdots\\ 0\end{bmatrix}_{(na_1-1)\times 1} \tag{4.2.54}$$

From Equation (4.2.53) and Equation (4.2.54), $\mathbf{M}_7{}'$ and $\mathbf{M}_7$ are the same. The

forms of $\mathbf{M}_8{}'$, $\mathbf{M}_9{}'$, $\mathbf{M}_8$, and $\mathbf{M}_9$ are the same as those expressed in Section

(4.2.3). We can prove that they are the same. Because all the elements in Equation

(4.2.26) are the same as those in Equation (4.2.44), the IIR crosstalk cancellers

designed in the time and frequency domains are the same.

Chapter 5

Robust Crosstalk Canceller

As we know, we must consider the robustness of the crosstalk canceller. So far, only single one set of system transfer functions is used to design the crosstalk canceller. We propose a new method to design a robust crosstalk canceller. A region-based concept [16] is considered instead of single one point. The new crosstalk canceller is designed by using the transfer function matrices in the region. In the previous section, we only want one set of ear signals to approach the desired signals. However, when the head moves, we get another set of ear signals. Figure 5.1 shows that when the head rotates around right to left, the transfer function matrix also changes accordingly.
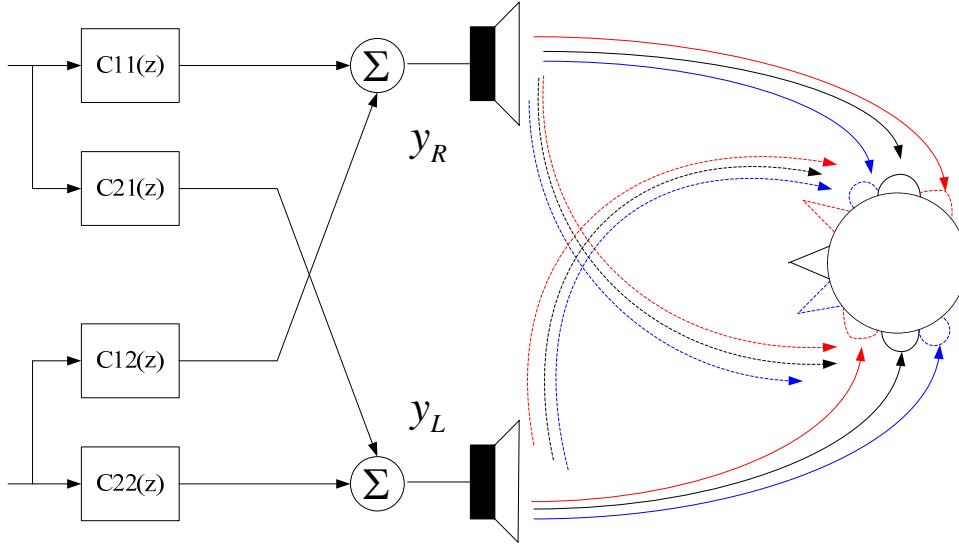
Figure 5.1: Head movements

Now, we want other sets of ear signals with head movements to approach the

desired signals. We can express the concept as follows:

$$\left(G(z)+\Delta G(z)\right)\cdot C(z)\approx\begin{bmatrix}d(z)\\0\end{bmatrix}$$ (5.1)

Referring to the concept in Equation (5.1), the problem can be formulated as

$$\begin{cases}G_o(z)\cdot h_1(z)\approx q_o(z)\\G_+(z)\cdot h_1(z)\approx q_+(z),\\G_-(z)\cdot h_1(z)\approx q_-(z)\end{cases}$$ (5.2)

where $h_1(z)=\begin{bmatrix}c_{11}(z)&c_{21}(z)\end{bmatrix}^T$ is the filters of the crosstalk canceller; $G_o(z)$ is the

transfer function matrix of fixed head, $G_+(z)$ is the one which head rotates to right,

and $G_-(z)$ is the one head rotates to left. $q_o(z)=\begin{bmatrix}d_o(z)&0\end{bmatrix}^T$, $q_+(z)=\begin{bmatrix}d_+(z)&0\end{bmatrix}^T$,

and $q_-(z)=\begin{bmatrix}d_-(z)&0\end{bmatrix}^T$ where $d_o(z)$, $d_+(z)$, and $d_-(z)$ are the desired signals

dependent on the head positions. Because of different head positions, the desired

signals must add delay compensation. We will propose the method to find the

55

compensation by using a head-centered coordinate model in the next section.

5.1 Delay Compensations

We have known that when the head moves, the distance between the source and

the ears changes. Therefore, the time and the amplitude of the sound source arrived to

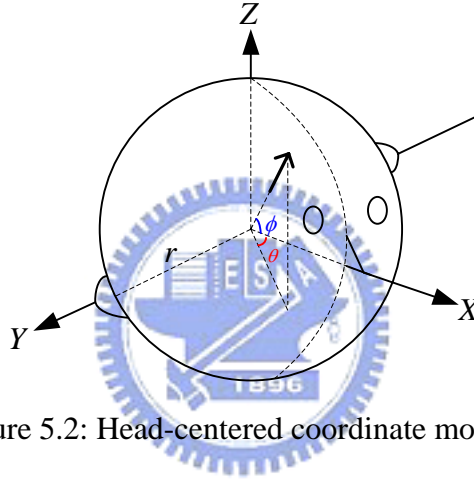the ear change. We will calculate the changes by the following figure.

Figure 5.2: Head-centered coordinate model

From Figure 5.2, any source can be presented in the head-centered coordinate

model, where the radius of the head is r. A source can be presented in three elements,

distance R, azimuth angle θ, and elevation angle ϕ. The source and the fixed right

ear can be written in vector forms as follows:

$$\vec{Rs} = R\left(\cos\phi\cos\theta, \cos\phi\sin\theta, \sin\phi\right) \quad \text{and} \quad \vec{re_o} = r(0,1,0), \tag{5.3}$$

where $\vec{Rs}$ represents the source and $\vec{re_o}$ represents the fixed right ear.

If the head rotates $\Delta\theta$ and $\Delta\phi$, the rotated right ear can be written as follows:

$$\vec{re_\Delta} = r\left(-\cos\Delta\phi\sin\Delta\theta, \cos\Delta\phi\cos\Delta\theta, \sin\Delta\phi\right) \tag{5.4}$$

If the distance from the source to the head center R is much larger than the radius of the head r. The compensated delay can be calculated by using the difference of two transmission paths such that:

$$delay\_comp(\Delta\theta, \Delta\phi) = \left(\left| R\vec{s} - r\overline{e_o} \right| - \left| R\vec{s} - r\overline{e_\Delta} \right| \right) \cdot fs/S$$
$$\approx r \cdot < \vec{s}, \overline{e_o} - \overline{e_\Delta} > \cdot fs/S, \tag{5.5}$$

where fs is the sampling frequency and S is the sound velocity; $<,>$ means vector inner product. Then the $delay\_comp(\Delta\theta, \Delta\phi)$ can be rounded off to get the integer. The amplitude compensation will be omitted if $R \gg r$.

According to Equation (5.5), the compensated desired signals can be expressed. For example, the desired signal to the fixed head is $d_o(z)$ and then the head rotates right side about $\Delta\theta_+$ and $\Delta\phi_+$. The desired signal to the rotated-right head becomes

$$d_+(z) = d_o(z) \cdot z^{-delay\_comp(\Delta\theta_+, \Delta\phi_+)} \tag{5.6}$$

In the same way, if the head turn left side about $\Delta\theta_-$ and $\Delta\phi_-$, the desired signal to the rotated-left head becomes

$$d_-(z) = d_o(z) \cdot z^{-delay\_comp(\Delta\theta_-, \Delta\phi_-)} \tag{5.7}$$

Therefore, the desired signals dependent on the head positions in Equation (5.2) can be found from Equation (5.6) and Equation (5.7).

5.2 Robust Common Pole Design

When the compensated delay has been found, it can be incorporated to implement the region-equalized concept. The LSE criterion is as follows:

$$h_1(z) = \arg\min \left\{ \left\| \begin{bmatrix} G_o(z) \\ G_+(z) \\ G_-(z) \end{bmatrix} \cdot h_1(z) - \begin{bmatrix} q_1(z) \\ q_{1+}(z) \\ q_{1-}(z) \end{bmatrix} \right\|^2 \right\} \tag{5.7}$$

If $h_1(z)$ is designed in FIR form, the solution is easy to find out.

We continue use the same concept by using IIR filters. The LSE criterion can be written as follows:

$$\begin{bmatrix} \dfrac{b_{11}(z)}{a_{11}(z)} \\ \dfrac{b_{21}(z)}{a_{21}(z)} \end{bmatrix} = \arg\min \left\{ \left\| \begin{bmatrix} G_o(z) \\ G_+(z) \\ G_-(z) \end{bmatrix} \cdot \begin{bmatrix} \dfrac{b_{11}(z)}{a_{11}(z)} \\ \dfrac{b_{21}(z)}{a_{21}(z)} \end{bmatrix} - \begin{bmatrix} q_o(z) \\ q_+(z) \\ q_-(z) \end{bmatrix} \right\|^2 \right\} \tag{5.8}$$

From the discussions in the previous section, we will meet the non-linearity because of the denominators $a_{11}(z)$ and $a_{21}(z)$. Therefore, the concept of the common pole model is used, and let

$$a_{11}(z) = a_{21}(z) = a_1(z) \tag{5.9}$$

Equation (5.8) can be written as

$$\begin{bmatrix} \dfrac{b_{11}(z)}{a_1(z)} \\ \dfrac{b_{21}(z)}{a_1(z)} \end{bmatrix} = \arg\min \left\{ \left\| \begin{bmatrix} G_o(z) \\ G_+(z) \\ G_-(z) \end{bmatrix} \cdot \begin{bmatrix} b_{11}(z) \\ b_{21}(z) \end{bmatrix} - a_1(z) \cdot \begin{bmatrix} q_o(z) \\ q_+(z) \\ q_-(z) \end{bmatrix} \right\|^2 \right\} \tag{5.10}$$

The above equation can be written as

$$\begin{bmatrix} \dfrac{b_{11}(z)}{a_1(z)} \\[2mm] \dfrac{b_{21}(z)}{a_1(z)} \end{bmatrix} = \arg\min \left\{ \left\| \begin{bmatrix} G_o(z) \cdot \begin{bmatrix} b_{11}(z) \\ b_{21}(z) \end{bmatrix} - a_1(z) \cdot q_o(z) \\[2mm] G_+(z) \cdot \begin{bmatrix} b_{11}(z) \\ b_{21}(z) \end{bmatrix} - a_1(z) \cdot q_+(z) \\[2mm] G_-(z) \cdot \begin{bmatrix} b_{11}(z) \\ b_{21}(z) \end{bmatrix} - a_1(z) \cdot q_-(z) \end{bmatrix} \right\|^2 \right\} \tag{5.11}$$

We can find that the right term in Equation (5.11) is composed by three common pole

models. Therefore, we can express the above equation in matrix and vector forms by

using the Equation (4.2.25).

$$\begin{bmatrix} \dfrac{b_{11}(z)}{a_1(z)} \\[2mm] \dfrac{b_{21}(z)}{a_1(z)} \end{bmatrix} = \arg\min \left\{ \left\| \begin{bmatrix} \begin{bmatrix} \mathbf{D}_o & \mathbf{G}_{rro} & \mathbf{G}_{lro} \\ \mathbf{0} & \mathbf{G}_{rlo} & \mathbf{G}_{llo} \end{bmatrix} \begin{bmatrix} -\mathbf{a}_1 \\ \mathbf{b}_{11} \\ \mathbf{b}_{21} \end{bmatrix} - \mathbf{q}_o \\[4mm] \begin{bmatrix} \mathbf{D}_+ & \mathbf{G}_{rr+} & \mathbf{G}_{lr+} \\ \mathbf{0} & \mathbf{G}_{rl+} & \mathbf{G}_{ll+} \end{bmatrix} \begin{bmatrix} -\mathbf{a}_1 \\ \mathbf{b}_{11} \\ \mathbf{b}_{21} \end{bmatrix} - \mathbf{q}_+ \\[4mm] \begin{bmatrix} \mathbf{D}_- & \mathbf{G}_{rr-} & \mathbf{G}_{lr-} \\ \mathbf{0} & \mathbf{G}_{rl-} & \mathbf{G}_{ll-} \end{bmatrix} \begin{bmatrix} -\mathbf{a}_1 \\ \mathbf{b}_{11} \\ \mathbf{b}_{21} \end{bmatrix} - \mathbf{q}_- \end{bmatrix} \right\|^2 \right\}$$

$$= \arg\min \left\{ \left\| \begin{bmatrix} \mathbf{D}_o & \mathbf{G}_{rro} & \mathbf{G}_{lro} \\ \mathbf{0} & \mathbf{G}_{rlo} & \mathbf{G}_{llo} \\ \mathbf{D}_+ & \mathbf{G}_{rr+} & \mathbf{G}_{lr+} \\ \mathbf{0} & \mathbf{G}_{rl+} & \mathbf{G}_{ll+} \\ \mathbf{D}_- & \mathbf{G}_{rr-} & \mathbf{G}_{lr-} \\ \mathbf{0} & \mathbf{G}_{rl-} & \mathbf{G}_{ll-} \end{bmatrix} \begin{bmatrix} -\mathbf{a}_1 \\ \mathbf{b}_{11} \\ \mathbf{b}_{21} \end{bmatrix} - \begin{bmatrix} \mathbf{q}_o \\ \mathbf{q}_+ \\ \mathbf{q}_- \end{bmatrix} \right\|^2 \right\}, \tag{5.12}$$

where $\mathbf{G}_{xyz}$ is the convolution matrix of channel g_{xyz}; $\mathbf{D}_z$ is the convolution

matrix of the desired signal; $\mathbf{q}_z$ is the vectors composed with desired signal and zero

signal. From above equation, we can find out the filters, $\mathbf{b}_{11}$, $\mathbf{b}_{21}$ and $\mathbf{a}_1$ by using

the LSE method.

Chapter 6

Computer Simulations

In this chapter, we will simulate different crosstalk cancellers to see their performance. Besides, we will compare the results of single point method and region-based method. Here, our channels used in the simulations are the HRTFs from the database of MIT Media Lab.
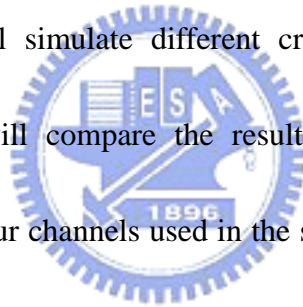
6.1 Figure of Merit

Perfect crosstalk cancellation would make the listener to receive no crosstalk and desired signal without distortion. Our simulation is that an impulse signal $\delta[n]$ is inputted at s_R and a zero signal is inputted at s_L in Figure 2.5. Figure 6.1 and Figure 6.2 show the signals received at both ears without the crosstalk canceller. The loudspeakers are placed at $\pm 30^o$.
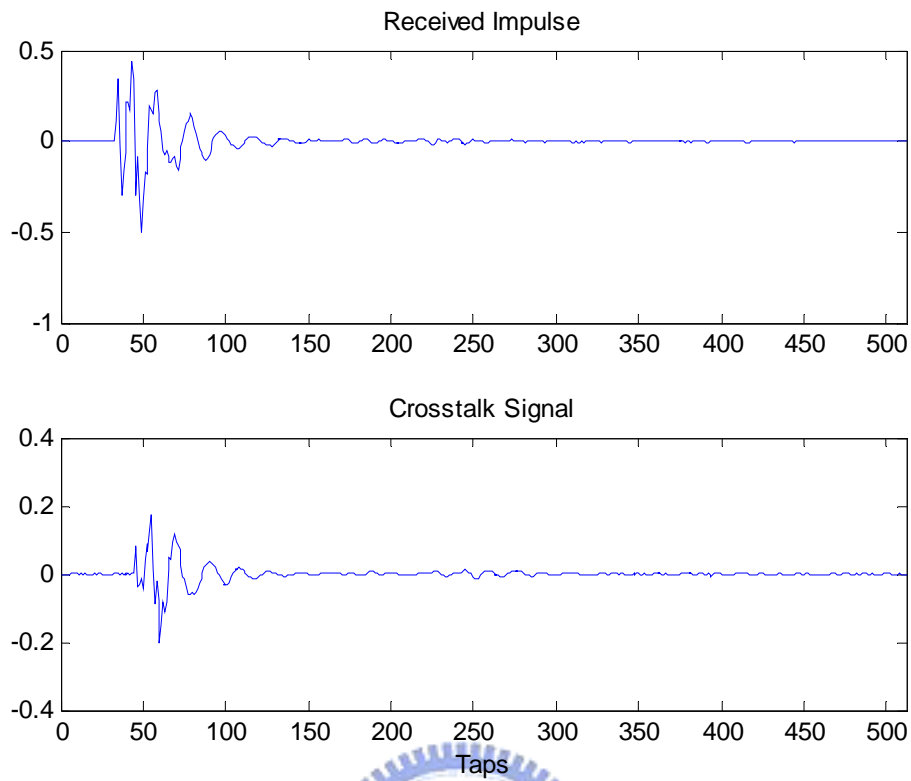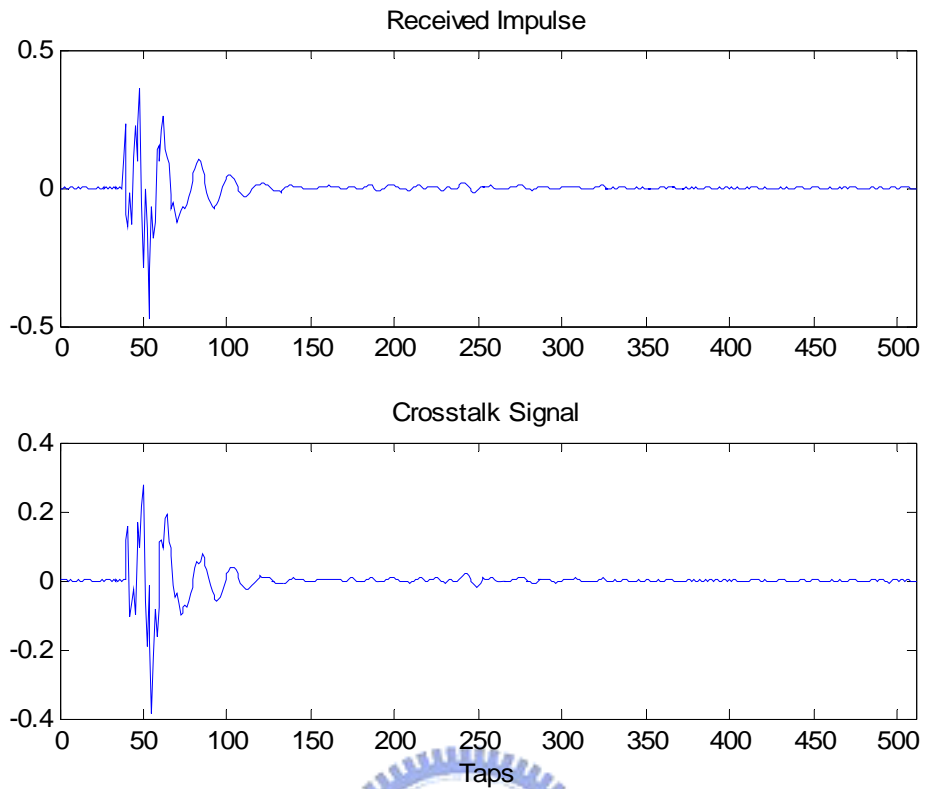
Figure 6.1: The impulse response at $\pm 30^o$ without the crosstalk canceller
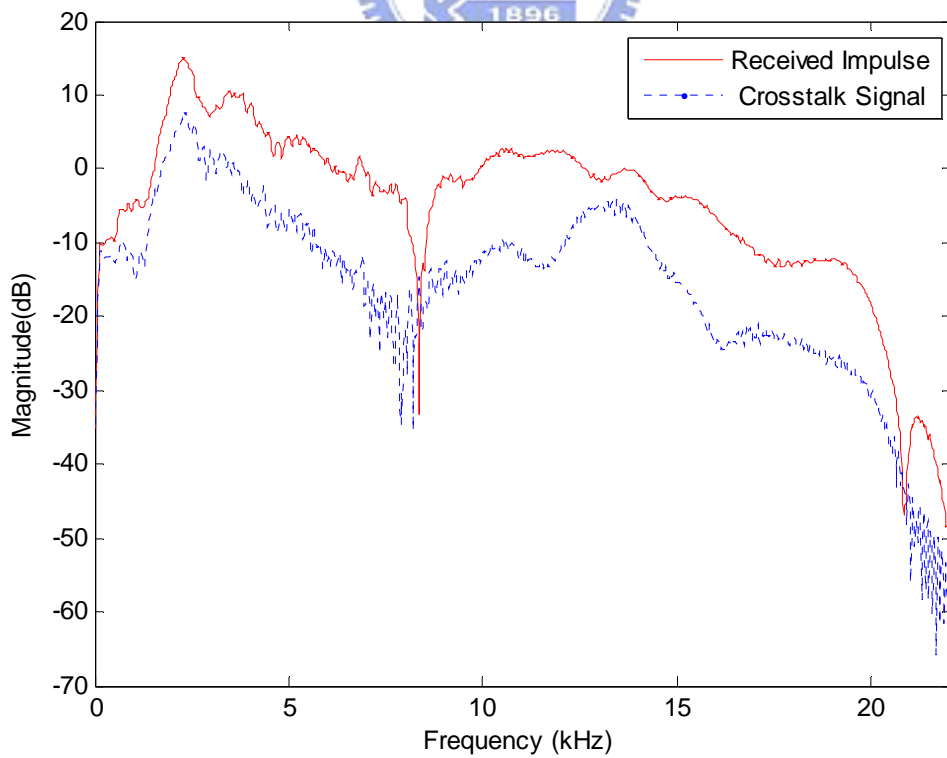
Figure 6.2: The frequency response at $\pm 30^o$ without the crosstalk canceller

The received impulses in Figure 6.1 we want should be a delayed impulse signal and the crosstalk signal should be zero. In the same way, the magnitude response of the received impulse in Figure 6.2 should be flat and the crosstalk signal should be negative infinite dB. Therefore, we define two indexes to evaluate the difference of channel attenuations and the flatness of the magnitude response. First, differential attenuation (*DA*) is to evaluate the difference of channel attenuations and defined as follows:

$$DA = 10 \log \frac{\sum\limits_{k=0}^{k=K-1} \left| E_R(k) \right|^2}{\sum\limits_{k=0}^{k=K-1} \left| E_L(k) \right|^2} \,, \qquad (6.1)$$

where $E_R(k)$ and $E_L(k)$ are the K-point DFT of $e_R(n)$ and $e_L(n)$. Second, equalization index (*EI*) is to evaluate the flatness of the received impulse signal and defined as follows:

$$EI = \left(\frac{1}{K} \sum\limits_{k=0}^{k=K-1} \left(20 \log | E_R(k) | - AV \right)^2 \right)^{1/2} , \qquad (6.2)$$

where $AV = \dfrac{1}{K} \sum\limits_{k=0}^{k=K-1} 20 \log | E_R(k) |$. It indicates the average deviation and AV is the average of the magnitude of $E_R(k)$.

In order to see the performance of the crosstalk canceller, we define three FOM (Figure of Merit) to quantify it. The crosstalk suppression factor (*CSF*) [16] is to evaluate how much the crosstalk canceller reduces crosstalk and defined as follows:

$$CSF = DA_{with\ C(z)} - DA_{without\ C(z)} \qquad (6.3)$$

The first term of the *CSF* indicates the *DA* after crosstalk cancellation, and the second term indicates the one without crosstalk cancellation.

The equalization improvement factor (*EIF*) is to see the equalization performance in the frequency domain and defined as follows:

$$EIF = EI_{without\ C(z)} - EI_{with\ C(z)},$$ (4.4)

We also want to see the total performance including the crosstalk suppression and spectral flatness of the crosstalk canceller. Therefore, we see the total energy of the error (*EN*) between the desired signals and the era signals, which is defined as

$$EN = \sum_{n=0}^{\infty} |\delta(n-delay) - e_R(n)|^2 + |e_L(n)|^2$$ (4.5)

In the simulation, the larger the *CSF* and *EIF* are, the more important in crosstalk suppression and spectral flatness the crosstalk cancellation achieve. The smaller the *EN* is, the smaller difference to desired signals achieves. In the next paragraph, we will use these factors, *CSF*, *EIF*, and *EN* to justify or compare the performance of different implements of crosstalk cancellers. From the definitions, we can calculate *DA* = 8.45 dB, and *EI* = 12 dB^2 at $\pm 30^o$ without the crosstalk canceller. We also use the small angle at $\pm 5^o$ to simulate the 3-D mobile phone environment. These results at $\pm 5^o$ are *DA* = 1.85 dB, and *EI* = 11.69 dB^2. Figure 6.3 and Figure 6.4 show the signals received at both ears at $\pm 5^o$ without the crosstalk canceller.

Figure 6.3: The impulse response at $\pm 5^o$ without the crosstalk canceller
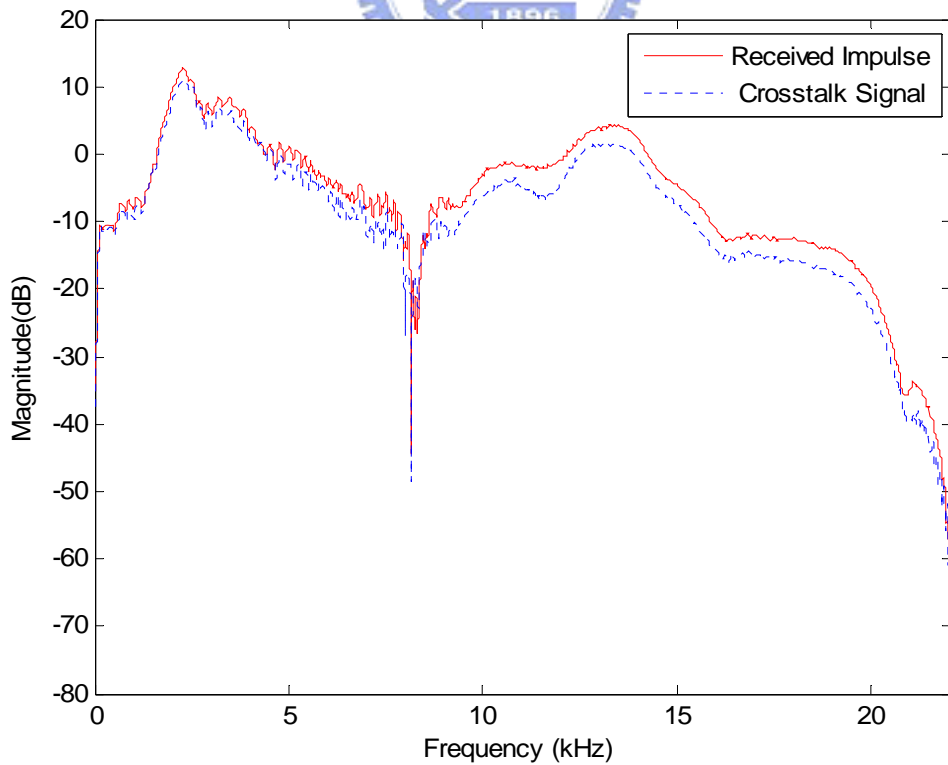
Figure 6.4: The frequency response at $\pm 5^o$ without the crosstalk canceller

Comparing Figure 6.2 and Figure 6.4, the *DA* at $\pm 30^o$ is larger than the *DA* at $\pm 5^o$ because there is small difference of transmission distances to the both ears at small incidence angle.

6.2 FIR Form

In this section, we will first simulate the matrix-inverse designs in both time and frequency domains. Next, the direct LSE design is also implemented in both time and frequency domains. We will test different lengths of the filters to show how it affects the performance.

6.2.1 Matrix-Inverse FIR Design

The first step to design the FIR filter is to calculate the denominator of G^{-1}, $D(z) = g_{rr}(z)g_{ll}(z) - g_{rl}(z)g_{lr}(z)$, and uses the solutions from Equation (3.1.9) and Equation (3.1.19) to design the crosstalk canceller. Besides, we must add delay for causality. In [8], we know that the extra delay can determine the performance of the crosstalk canceller. Here for comparison, we have taken the optimum delay for each FIR order in our simulations.

6.2.1.1 Design in Time Domain

The HRTFs we used are of 512-point. In the simulation, we will try different numbers of filter lengths to see how it affects the crosstalk cancellation.

Figure 6.5 and Figure 6.6 show the results of crosstalk cancellation over a pair of loudspeakers at $\pm 5^{o}$ with the crosstalk canceller of 50-tap and an extra 67 sample delay. The results of *DA* and *EI* are 13.78 dB and 19.08 dB^2, so the FOM are *CSF* = 11.93 dB, *EIF* = 7.39 dB^2, and *EN* = 0.1223.
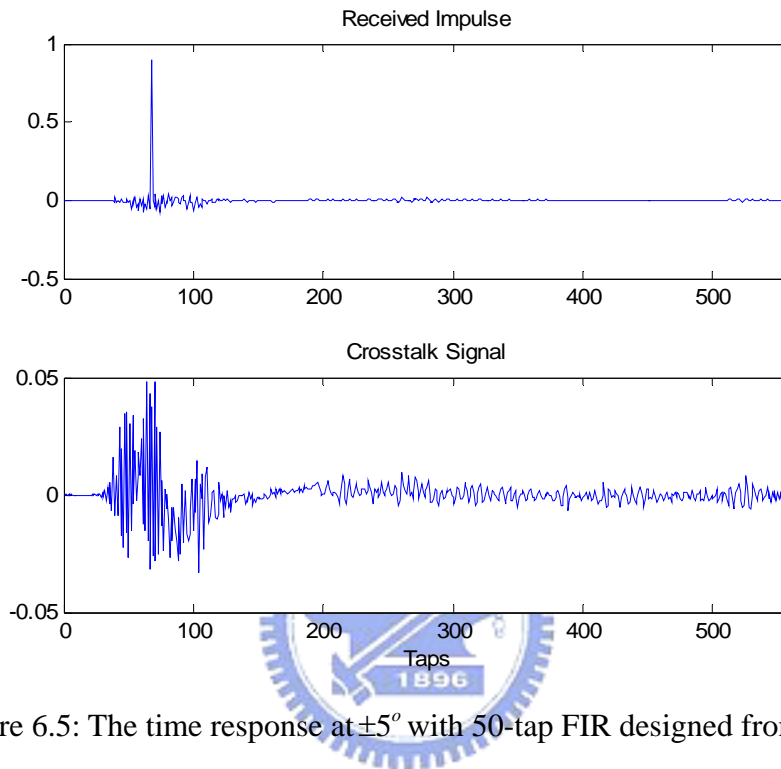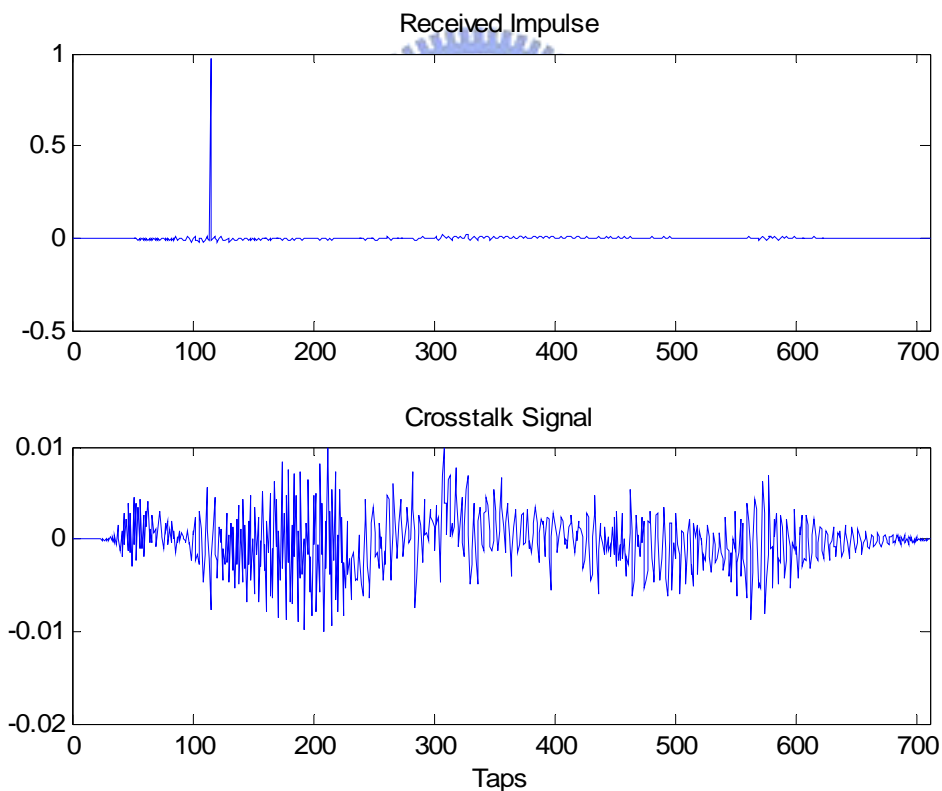

Figure 6.5: The time response at $\pm 5^{o}$ with 50-tap FIR designed from G^{-1}
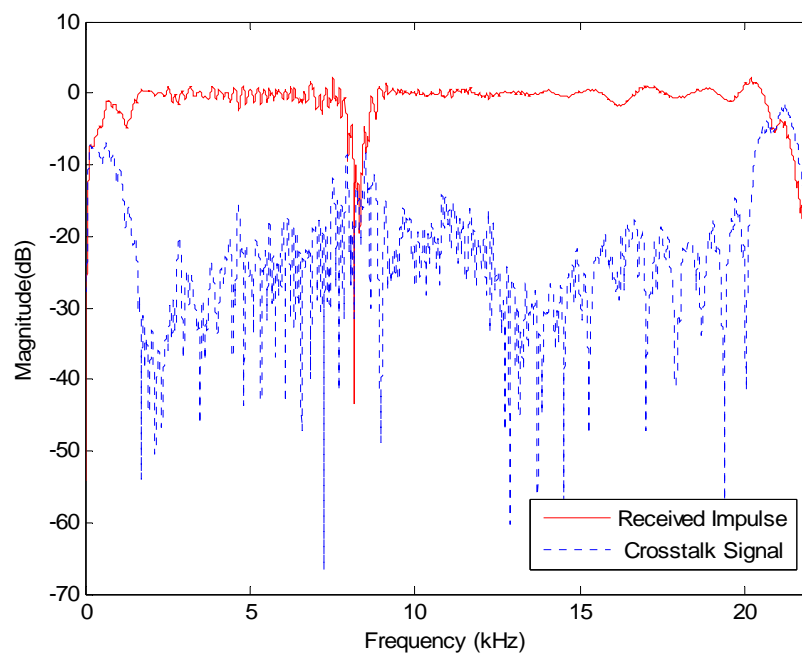
Figure 6.6: The frequency response at $\pm 5^{o}$ with 50-tap FIR designed from G^{-1}

Figure 6.7 and Figure 6.8 are the impulse responses of c_{11} and c_{21}.
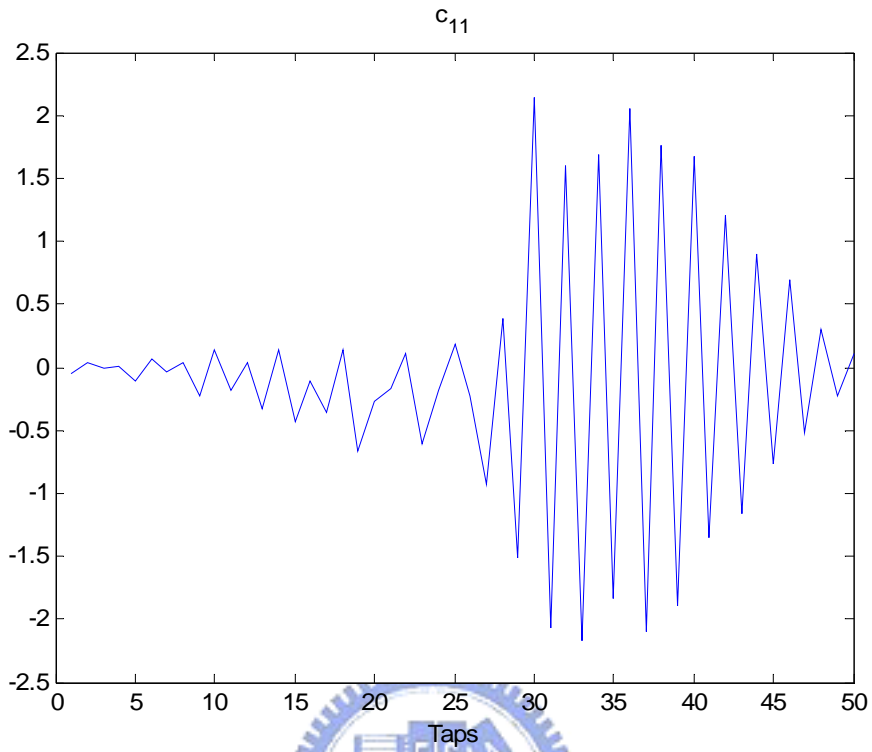
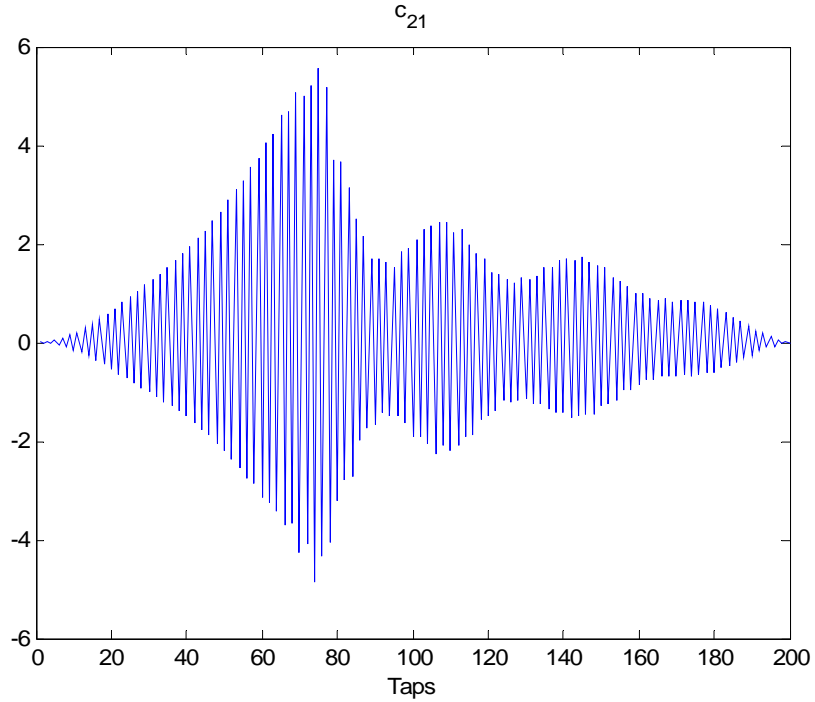Figure 6.7: Impulse response of c_{11} at $\pm 5^o$ designed from G^{-1}
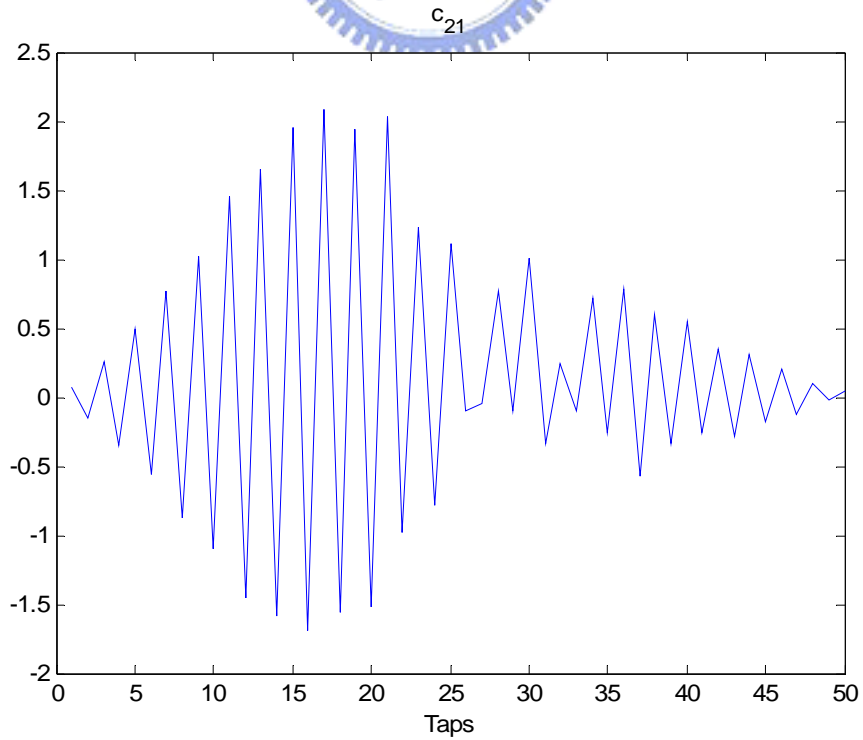
Figure 6.8: Impulse response of c_{21} at $\pm 5^o$ designed from G^{-1}

Because the denomination of the theoretical formulation in Equation (2.3.5) is often of non-minimum phase, the impulses have oscillations in the above two figures.

Next, Figure 6.9 and Figure 6.10 show the results that c_{11} and c_{21} are used with 200 taps and an extra 114 sample delay. In this simulation, *DA* and *EI* are 21.68 dB and 21.75 dB^2, in other words, the performance of crosstalk cancellation can reach *CSF* = 19.83 dB, *EIF* = 10.06 dB^2, and *EN* = 0.0287. We can find that the *CSF* and *EIF* become larger and the *EN* becomes smaller because the FIR filter length is increased.
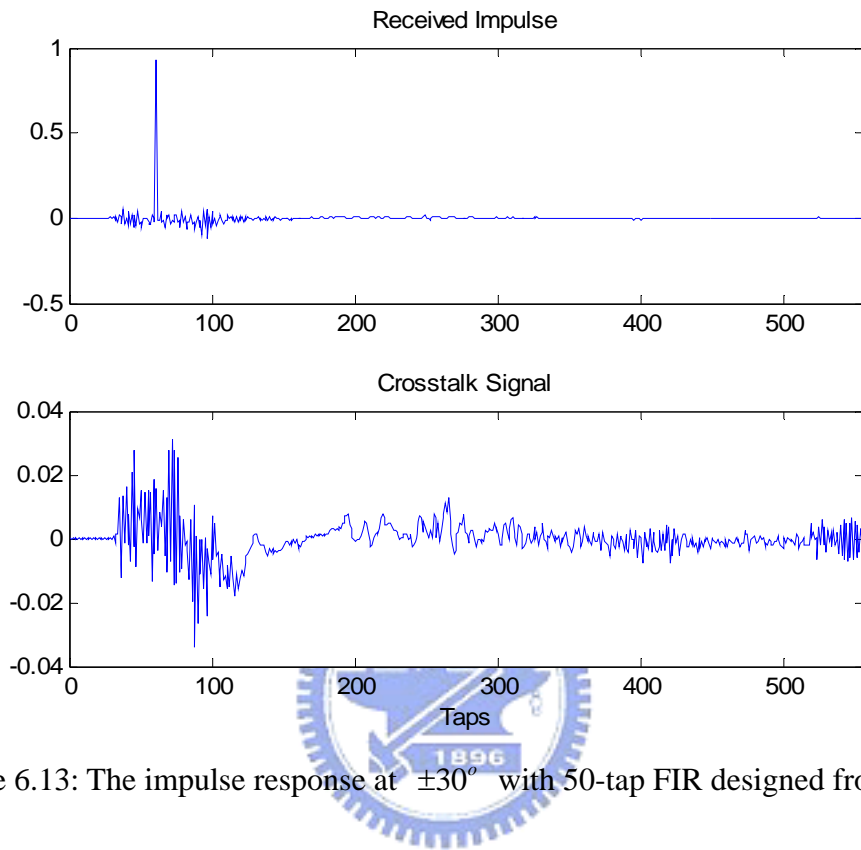

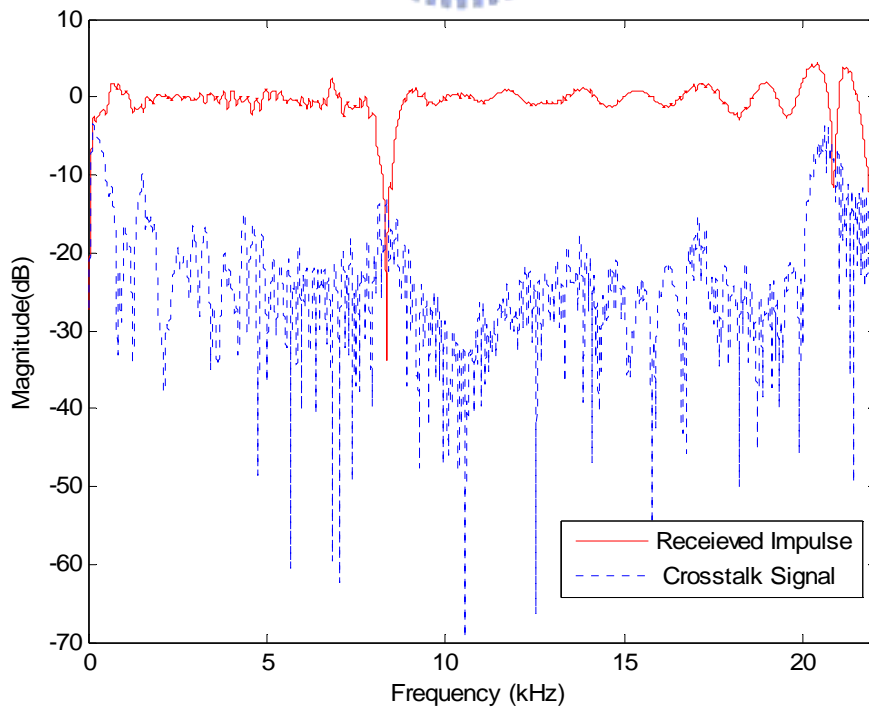Figure 6.9: The impulse response at $\pm 5^o$ with 200-tap FIR designed from G^{-1}
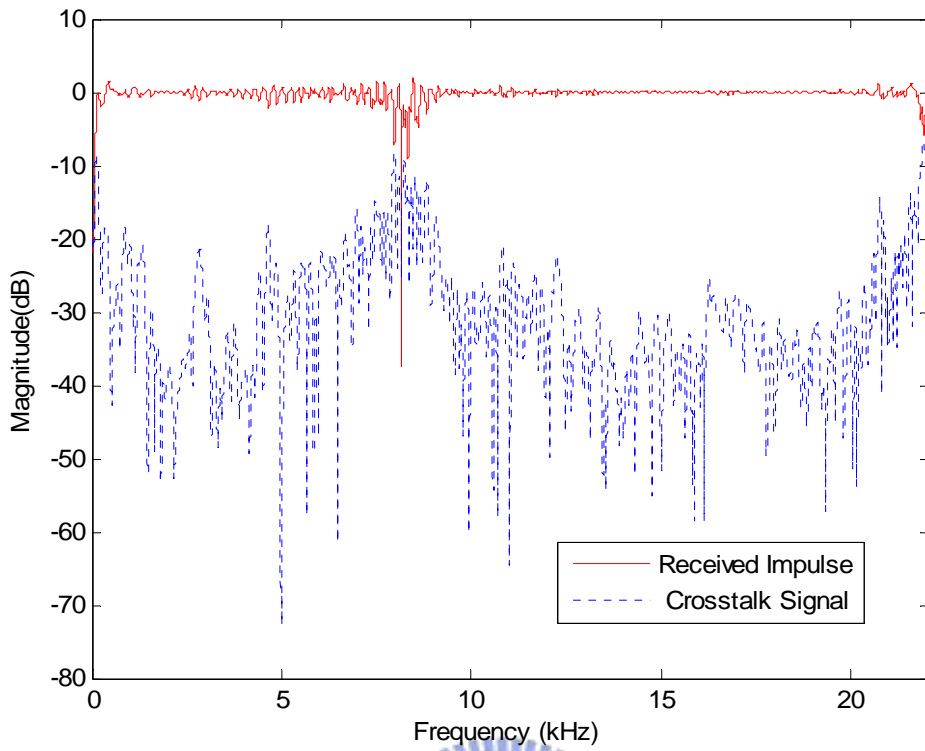
Figure 6.10: The frequency response at $\pm 5^o$ with 200-tap FIR designed from G^{-1}

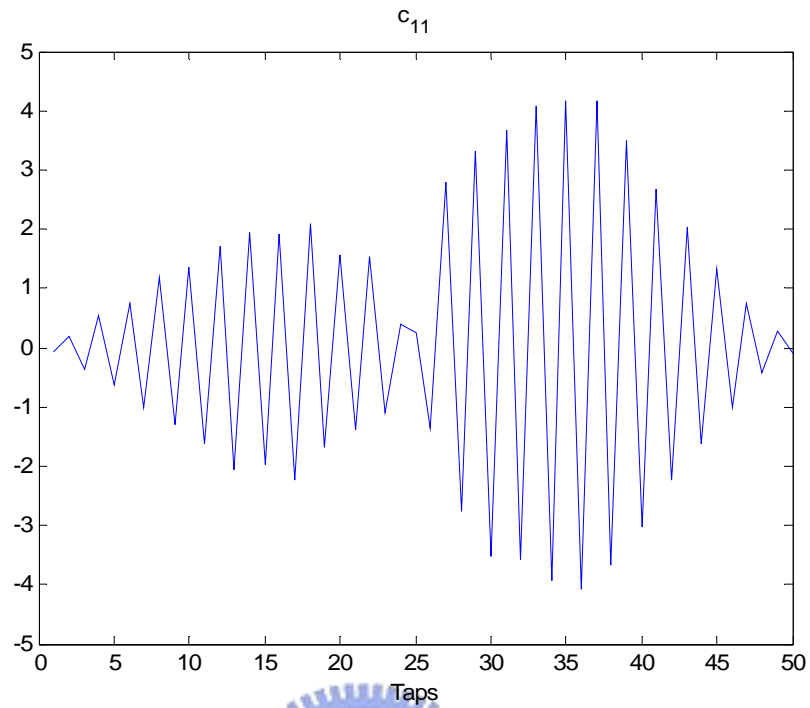Figure 6.11 and Figure 6.12 are the impulse responses of $\mathbf{c}_{11}$ and $\mathbf{c}_{21}$.
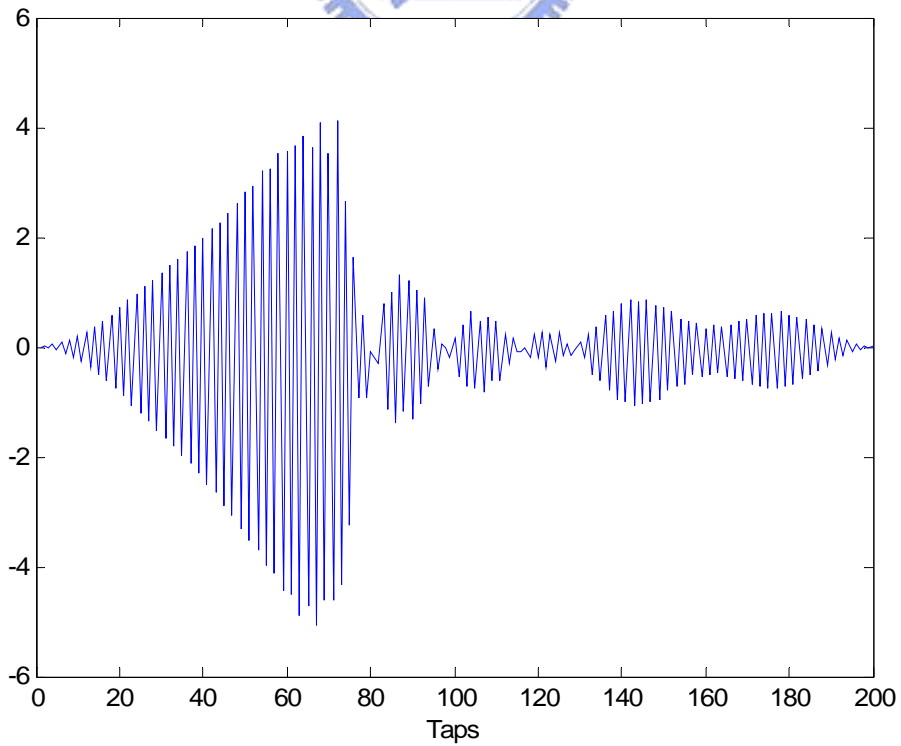
Figure 6.11: Impulse response of $\mathbf{c}_{11}$ at $\pm 5^o$ designed from G^{-1}
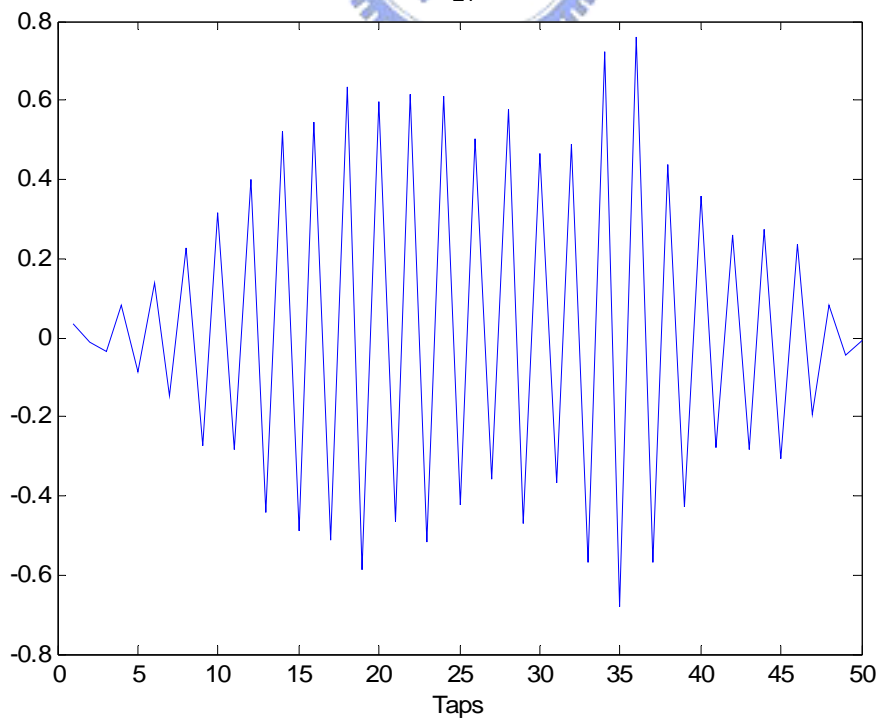
Figure 6.12: Impulse response of $\mathbf{c}_{21}$ at $\pm 5^o$ designed from G^{-1}

Next, let the filter length N equal 25, 50, 100 and 200 taps successively, and the results are listed in Table 6.1.

Filter Order	*EN*	*CSF* (dB)	*EIF* (dB^2)
25	0.1977	13.31	6.11
50	0.1223	11.93	7.39
100	0.0502	21.92	8.36
200	0.0287	19.83	10.06

Table 6.1: FOM at $\pm 5^o$ of FIR form designed from G^{-1}
in the time domain

From Table 6.1, as the order increases, the *EN* decreases, and *CSF* and *EIF* increase.

In other words, as the length of the filters increases, the performance also gets better.

Next, we will test the situation of the loudspeaker pair set at $\pm 30^o$. Figure 6.13

and Figure 6.14 show the results of crosstalk cancellation with 50 taps and an extra 60 sample delay. *DA* and *EI* are 16.99 dB and 20.71 dB^2, so the performance can reach *CSF* = 8.54 dB, *EIF* = 8.71 dB^2, and *EN* = 0.1344.
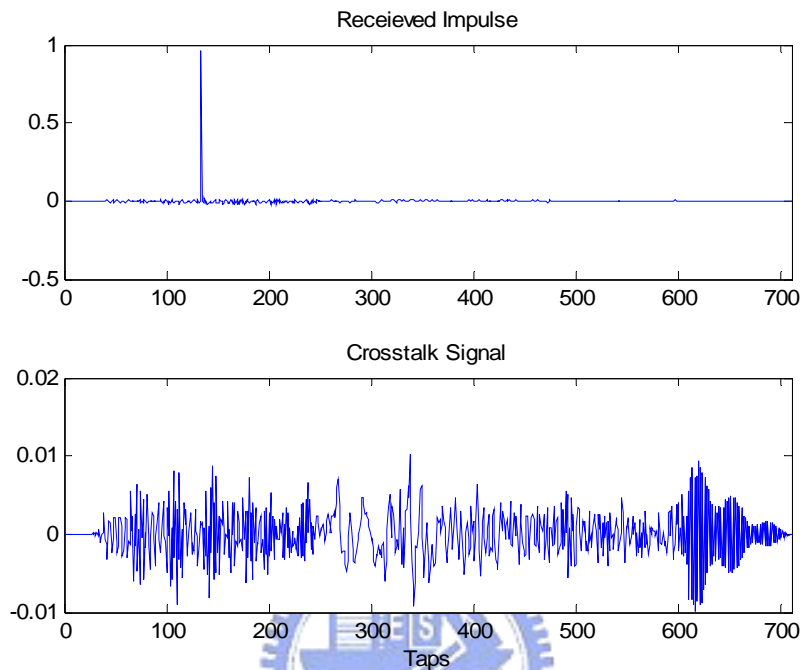

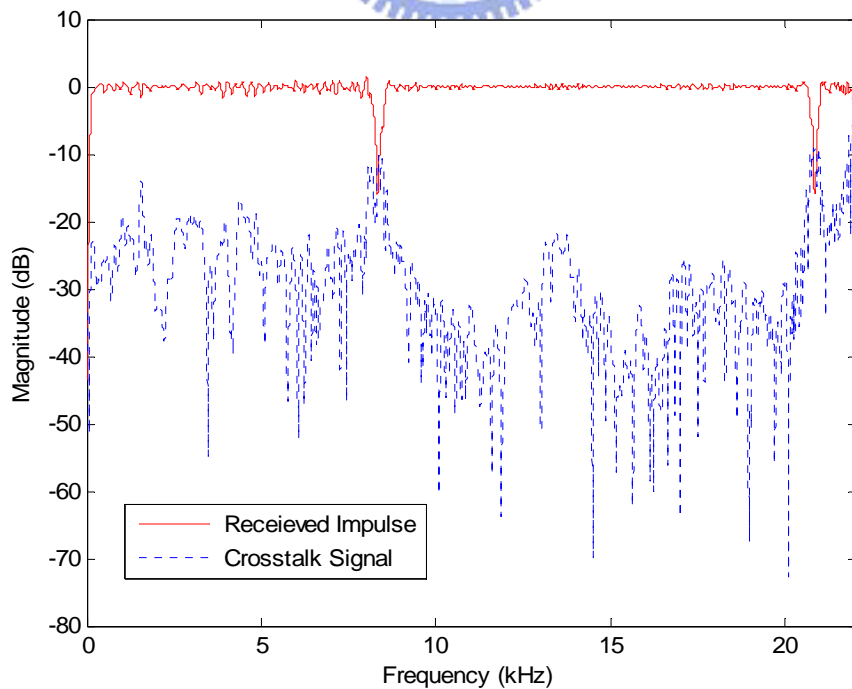Figure 6.13: The impulse response at $\pm30^o$ with 50-tap FIR designed from G^{-1}

Figure 6.14: The frequency response at $\pm30^o$ with 50-tap FIR designed from G^{-1}

Figure 6.15 and Figure 6.16 are the impulse responses of c_{11} and c_{21}.

c_{11}

Figure 6.15: Impulse response of c_{11} at $\pm 30^o$ designed from G^{-1}

c_{21}

Figure 6.16: Impulse response of c_{21} at $\pm 30^o$ designed from G^{-1}

Figure 6.17 and Figure 6.18 show the results of the crosstalk cancellation with 200 taps and an extra 133 sample delay. In this simulation, *DA* and *EI* are 21.43 dB and 21.61 dB^2, so the performance of crosstalk cancellation can reach *CSF* = 12.98 dB, *EIF* = 9.61 dB^2, and *EN* = 0.03.

Figure 6.17: The impulse response at $\pm 30^o$ with 200-tap FIR designed from G^{-1}

Figure 6.18: The frequency response at $\pm 30^o$ with 200-tap FIR designed from G^{-1}

The following two figures are the impulse responses of the filters $\mathbf{c}_{11}$ and $\mathbf{c}_{21}$.
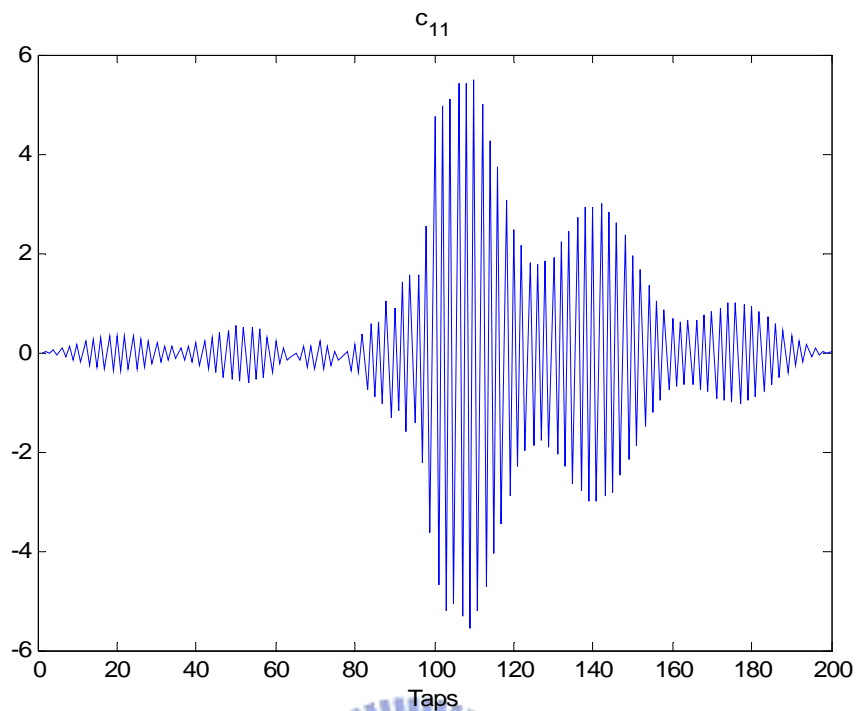
Figure 6.19: Impulse response of $\mathbf{c}_{11}$ at $\pm 30^o$ designed from G^{-1}
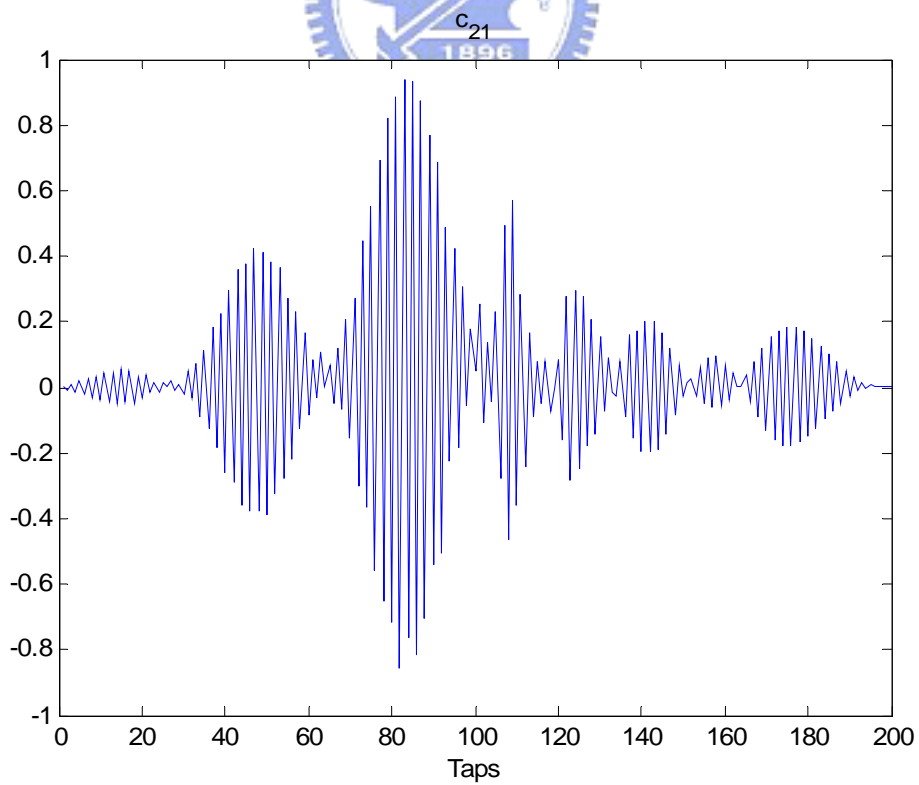

Figure 6.20: Impulse response of $\mathbf{c}_{21}$ at $\pm 30^o$ designed from G^{-1}

We can find that the *CSF* and *EIF* also become larger and *EN* becomes smaller.

Next, we use $N = 25, 50, 100, 200$ taps successively, and the results are listed in Table

6.2.

Filter Order	*EN*	*CSF* (dB)	*EIF* (dB^2)
25	0.2425	5.89	5.82
50	0.1344	8.54	8.71
100	0.0578	10.43	9.55
200	0.0322	12.98	9.61

Table 6.2: FOM at $\pm30^o$ of FIR form designed from G^{-1}
in the time domain

6.2.1.2 Design in Frequency Domain

We also use different lengths to see the effect to the performance of the crosstalk

cancellation. Table 6.3 lists the results of the loudspeaker pair at $\pm5^o$ and Table 6.4

lists the results of the loudspeaker pair at $\pm30^o$.

Filter Order	*EN*	*CSF* (dB)	*EIF* (dB^2)
25	0.1631	13.98	6.53
50	0.0843	17.21	8.12
100	0.0441	19.62	9.10
200	0.0274	21.73	9.87

Table 6.3: FOM at $\pm5^o$ of FIR form from G^{-1} in the frequency domain

Filter Order	EN	CSF (dB)	EIF (dB^2)
25	0.2059	7.09	5.88
50	0.089	10.71	8.42
100	0.0512	13.00	8.51
200	0.0307	13.08	9.66

Table 6.4: FOM at $\pm 30^o$ of FIR form designed from G^{-1} in the frequency domain

We will compare design in the time domain and design in the frequency domain. Figure 6.21 and Figure 6.22 are plotted the *EN* of designs in the time and frequency domain with different taps.
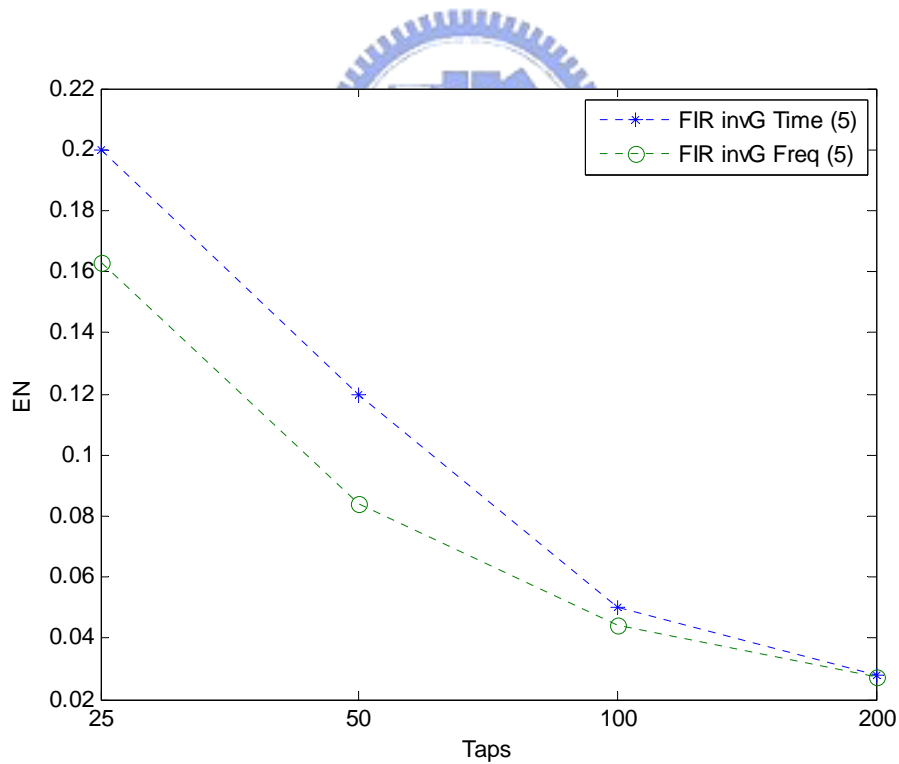

Figure 6.21: Comparison of design from G^{-1} at $\pm 5^o$ in time and frequency domains
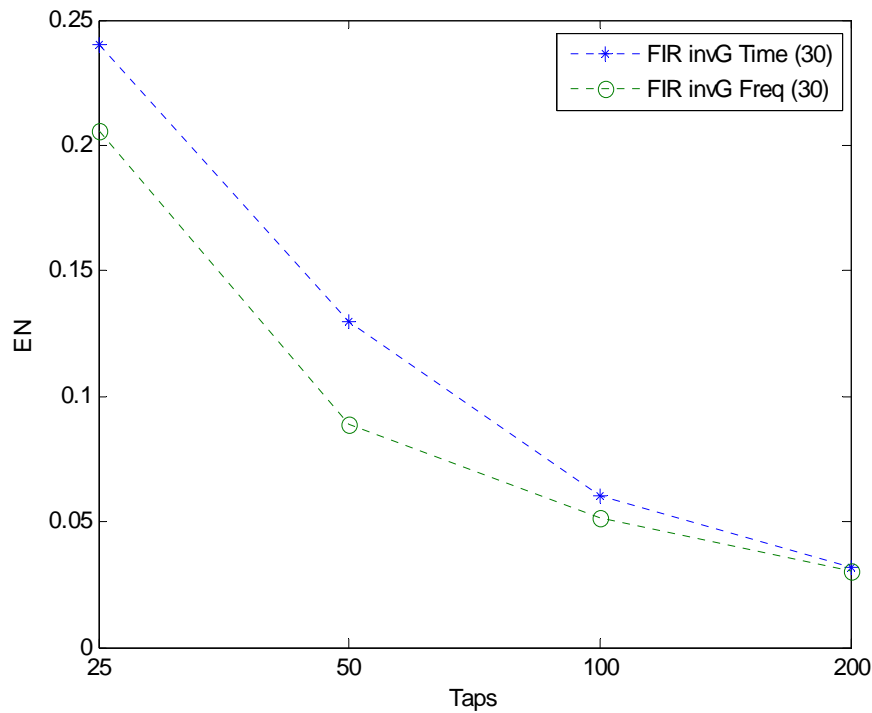
Figure 6.22: Comparison of design from G^{-1} at $\pm 30^o$ in time and frequency domains

From Figure 6.21 and Figure 6.22, the performance of design in the frequency domain is better than design in the time domain.

6.2.2 Direct LSE FIR Designed

6.2.2.1 Design in Time Domain

In this section, we will design the crosstalk canceller filters by using the direct LSE method. Similarly in this section, we will test different filter lengths of crosstalk cancellers to see the how it affects the crosstalk cancellation. Figure 6.23 is the results of the *EN* with different filter order at $\pm 5^o$. We plot the *EN* with respect to the tap number from 1 to 200.
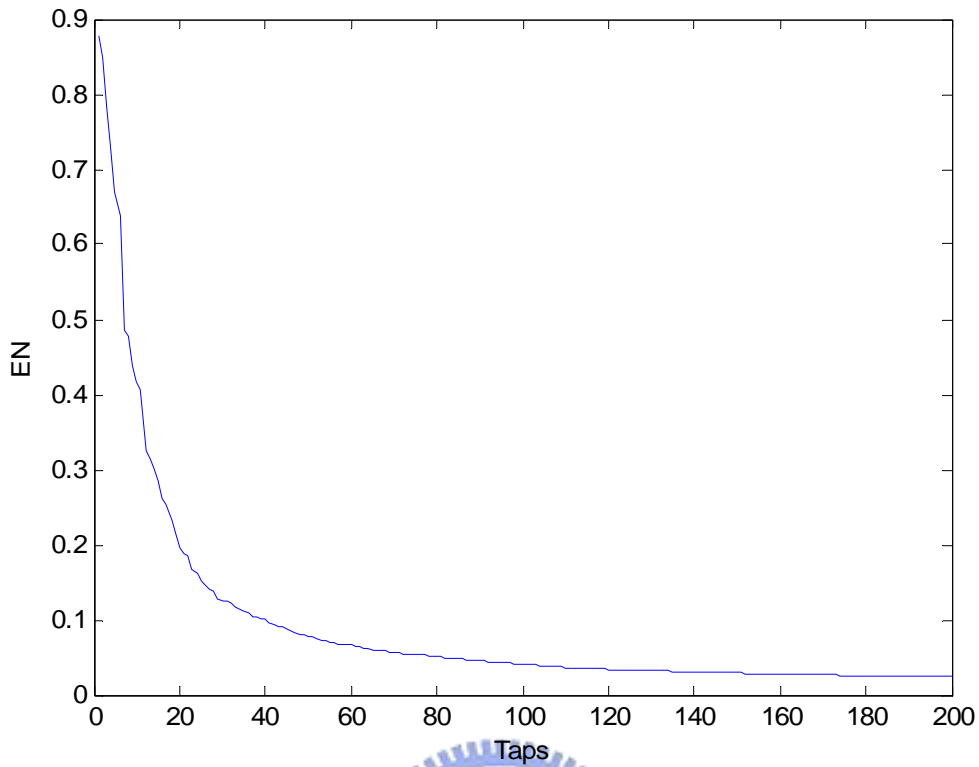
Figure 6.23: *EN* with different taps at $\pm 5^o$ using direct LSE in time domain

From the figure, it is obvious that the order of filters is increased and the performance of the crosstalk cancellation gets better. Similarly, we take four orders, 25, 50, 100 and 200 to see the FOM and the results are listed in Table 6.5.

Filter Order	*EN*	*CSF* (dB)	*EIF* (dB^2)
25	0.1506	15.69	6.49
50	0.0782	18.22	8.03
100	0.0408	20.01	9.20
200	0.0244	21.18	10.16

Table 6.5: FOM at $\pm 5^o$ of FIR form designed using direct LSE
in the time domain

Next, we simulate the situation the loudspeaker pair placed at $\pm 30^o$ and also

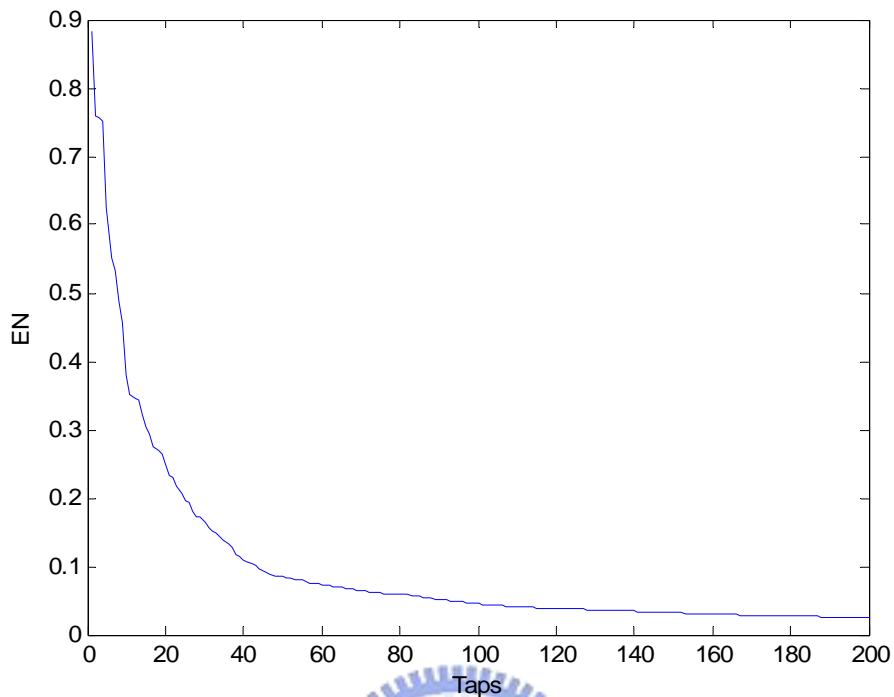plot the *EN* with respect to the tap number from 1 to 200 in Figure 6.24.

Figure 4.24: *EN* with different taps at $\pm 30^{o}$ using LSE in time domain

We also simulate four orders, 25, 50, 100 and 200, to see the FOM at $\pm 30^{o}$, and the results are listed in Table 6.6.

Filter Order	*EN*	*CSF* (dB)	*EIF* (dB^2)
25	0.1964	7.86	5.79
50	0.0843	11.43	8.59
100	0.0451	12.27	8.80
200	0.0253	11.99	10.11

Table 6.6: FOM at $\pm 30^{o}$ of FIR form designed using LSE
in the time domain

Next, we will compare the direct FIR LSE with the matrix-inverse design, and the *EN*

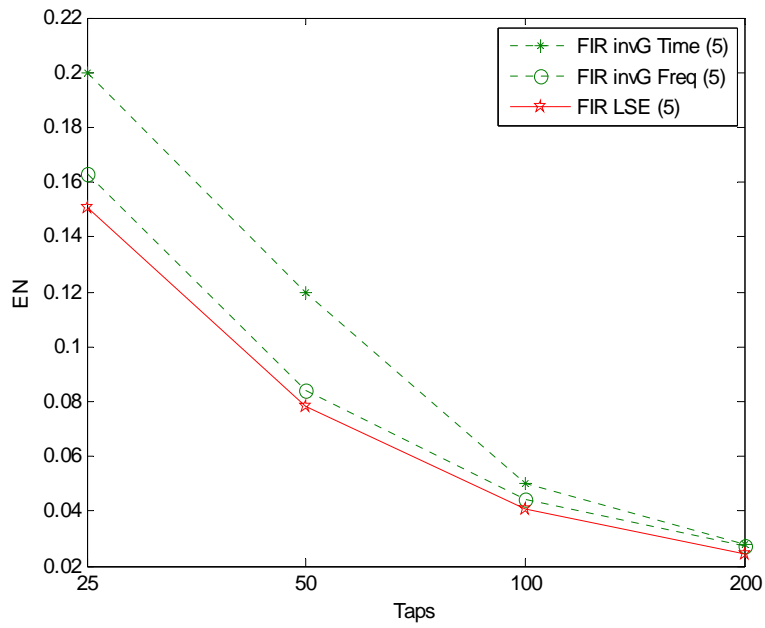of the two methods at $\pm 5^{o}$ and $\pm 30^{o}$ is plotted in Figure 6.25 and Figure 6.26.

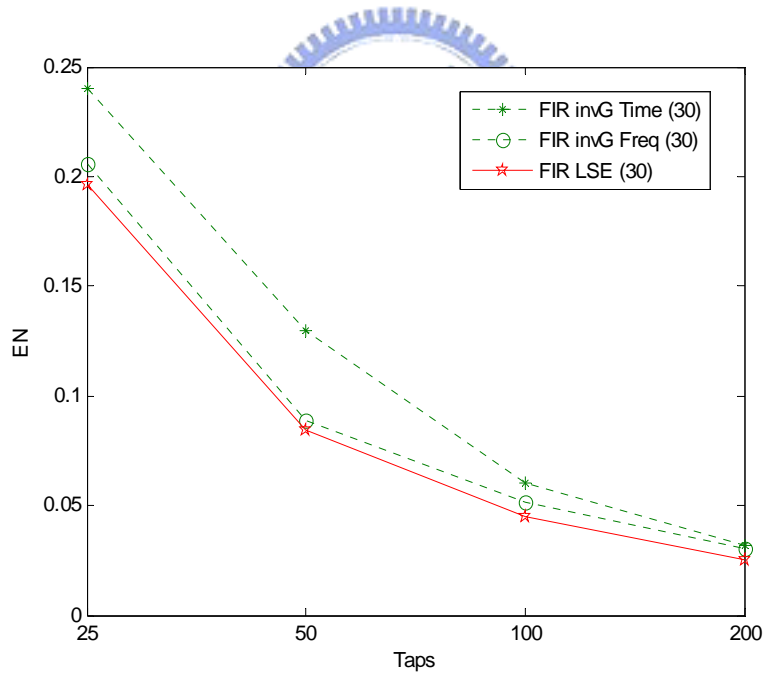Figure 6.25: Comparison between the direct FIR LSE and matrix-inverse at $\pm 5^o$

Figure 6.26: Comparison between the direct FIR LSE and matrix-inverse at $\pm 30^o$

From the above figures, the performance of the direct LSE FIR method is better

than design from matrix-inverse.

6.2.2.2 Design in Frequency Domain

In this section, we will simulate the crosstalk canceller designed in the frequency domain and the results of the loudspeaker pair placed at $\pm 5^{o}$ and $\pm 30^{o}$ are listed in Table 6.7 and Table 6.8.

Filter Order	EN	CSF (dB)	EIF (dB^2)
25	0.1506	15.69	6.49
50	0.0782	18.22	8.03
100	0.0408	20.01	9.20
200	0.0244	21.18	10.16

Table 6.7: FOM at $\pm 30^{o}$ of FIR form designed using LSE
in the frequency domain

Filter Order	EN	CSF (dB)	EIF (dB^2)
25	0.1964	7.86	5.79
50	0.0843	11.43	8.59
100	0.0451	12.27	8.80
200	0.0253	11.99	10.11

Table 6.8: FOM at $\pm 30^{o}$ of FIR form designed using LSE
in the frequency domain

Comparing Table 6.7 and Table 6.8 with Table 6.5 and Table 6.6, the results are the same. In other words, the performance designed in the frequency domain is the same as that designed in the time domain.

6.3 IIR Form

6.3.1 Filters Designed from Matrix-Inverse

6.3.1.1 Design in Time Domain

So far, we have found the crosstalk canceller in FIR forms. We want to design the filters in IIR form instead of FIR form. In Section (4.1.1), we will use IIR to approximate the theoretical solutions. We use different total taps (the sum of one denominator and one numerator) to simulate the performance and the results of the loudspeakers pair placed at $\pm 5^o$ and $\pm 30^o$ at are listed in Table 6.9 and Table 6.10.

Filter Order	*EN*	*CSF* (dB)	*EIF* (dB^2)
25	0.1947	14.63	5.59
50	0.1068	18.10	6.86
100	0.0462	20.79	8.84
200	0.0271	20.86	9.95

Table 6.9: FOM at $\pm 5^o$ of IIR form G^{-1} in the time domain

Filter Order	*EN*	*CSF* (dB)	*EIF* (dB^2)
25	0.2338	3.9	7.52
50	0.1089	9.82	8.05
100	0.0578	11.66	9.24
200	0.0318	12.28	9.62

Table 6.10: FOM at $\pm 30^o$ of IIR form G^{-1} in the time domain

6.3.1.2 Design in Frequency Domain

In this section, we will design the filters from matrix inverse G^{-1} in the frequency domain and we also use different total taps to see the performance. The results of the loudspeaker pair placed at $\pm 5^o$ and $\pm 30^o$ are listed in Table 6.11 and Table 6.12.

Filter Order	EN	CSF (dB)	EIF (dB²)
25	0.1631	13.99	6.54
50	0.0833	16.01	8.29
100	0.0349	20.05	9.78
200	0.0244	21.44	10.36

Table 6.11: FOM at $\pm 5^o$ of IIR form G^{-1} in the frequency domain

Filter Order	EN	CSF (dB)	EIF (dB²)
25	0.2059	7.09	5.90
50	0.089	10.634	8.45
100	0.0457	11.3141	9.34
200	0.0301	11.85	9.8

Table 6.12: FOM at $\pm 30^o$ of IIR form G^{-1} in the frequency domain

We will compare the IIR design between in the time and frequency domains and also compare IIR and FIR designs from matrix-inverse. The results of loudspeaker

pair placed at $\pm 5^o$ and $\pm 30^o$ are in Figure 6.27 and Figure 6.28.
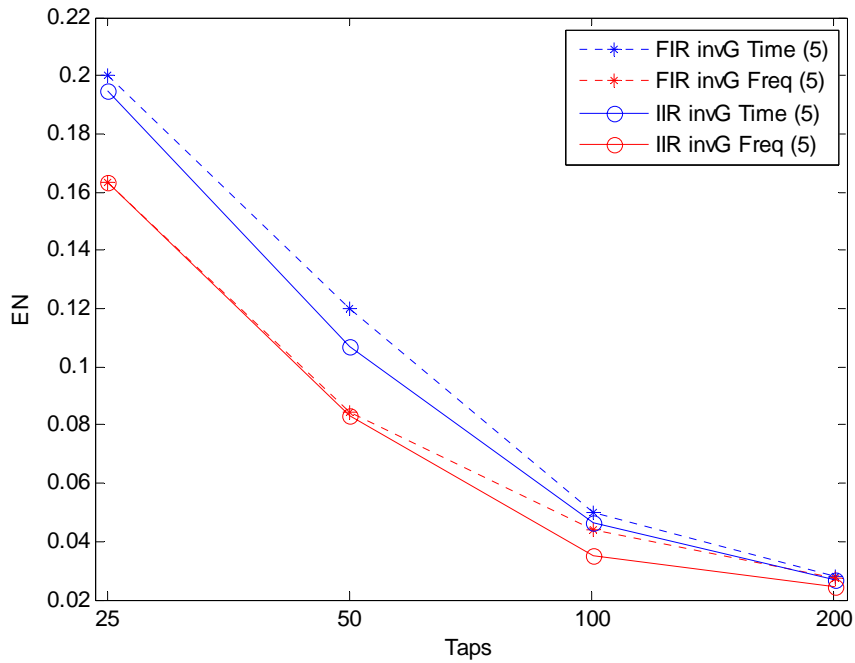
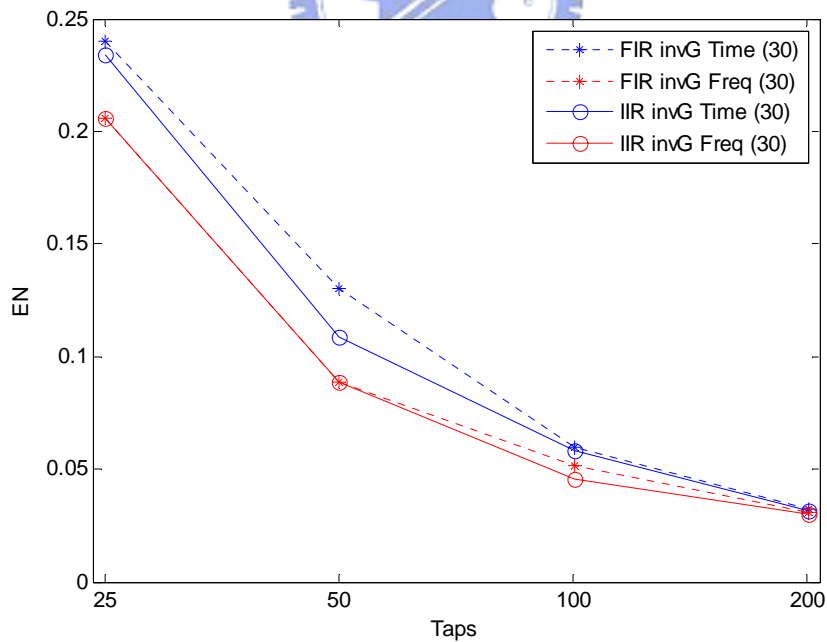Figure 6.27: Comparison with FIR and IIR designs from G^{-1} at $\pm 5^o$ in the time and frequency domains

Figure 6.28: Comparison with FIR and IIR designs from G^{-1} at $\pm 30^o$ in the time and frequency domains

In Figure 6.27 and Figure 6.28, blue and red lines represent designs in the time

and frequency domains, and solid and dotted lines represent designs in IIR and FIR.

We can find that the performance of IIR is better than FIR and design in the frequency

domain is better than design in the time domain.

6.3.2 Common-Pole IIR Design

6.3.2.1 Design in Time Domain

In this section, we will simulate the performance of the common-pole IIR model.

We simulate different total combined taps (the sum of two numerators and one

denominator) and the results of loudspeaker pair placed at $\pm 5^o$ and $\pm 30^o$ are listed

in Table 6.13 and Table 6.14.
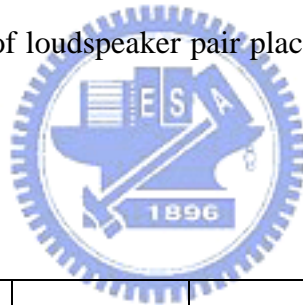
Filter Order	*EN*	*CSF* (dB)	*EIF* (dB^2)
50	0.1437	13.89	7.5
100	0.0614	15.43	9.09
200	0.0336	18.26	9.85
400	0.0222	21.47	10.39

Table 6.13: FOM at $\pm 5^o$ of common pole model

Filter Order	*EN*	*CSF* (dB)	*EIF* (dB^2)
50	0.1964	7.86	5.79
100	0.0756	6.79	10.19
200	0.0369	11.02	9.38
400	0.0224	11.706	10.54

Table 6.14: FOM at $\pm 30^o$ of common pole model

We compare the common-pole model with IIR design from matrix-inverse, and the EN of common-pole model and matrix-inverse at $\pm 5^o$ and $\pm 30^o$ are plotted in Figure 6.29 and Figure 6.30.
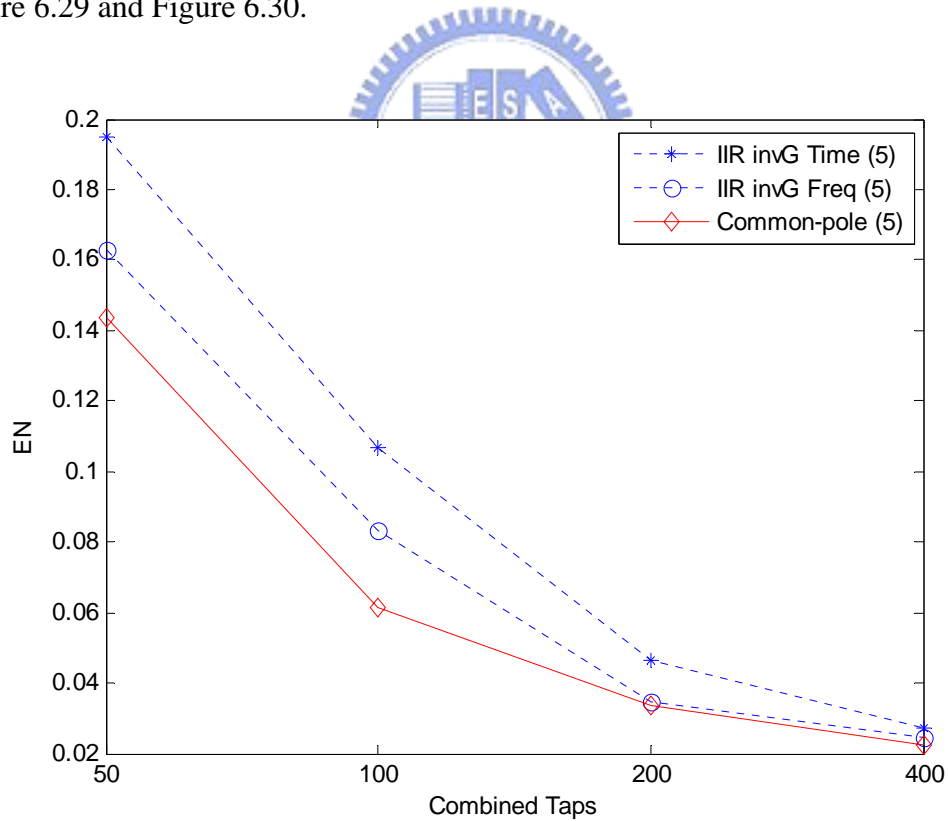

Figure 6.29: comparison between common-pole and IIR design form G^{-1} at $\pm 5^o$
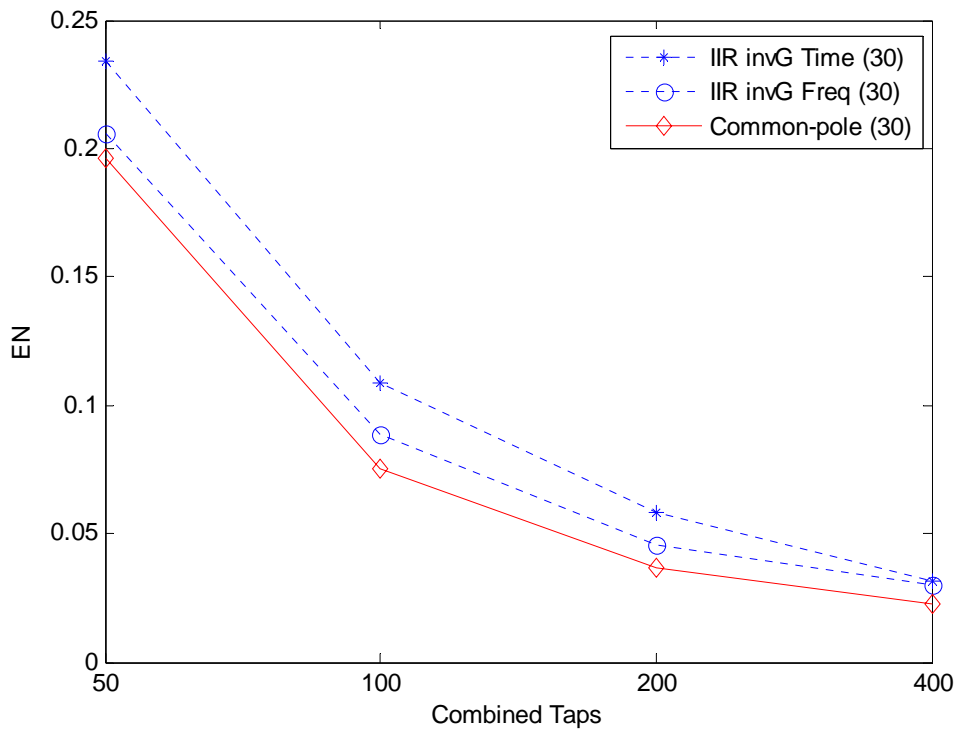
Figure 6.30: comparison between common-pole and IIR design form G^{-1} at $\pm 30^{o}$

In Figure 6.29 and Figure 6.30, the blue lines represent the *EN* from matrix-inverse design in the both domains and the red line represents the common pole model. We can find the performance of common is better than the design from matrix-inverse.

We compare the common pole model with LSE FIR form which is the best performance in FIR designs. The *EN* results which loudspeaker pair placed at $\pm 5^{o}$ and $\pm 30^{o}$ with different total combined taps of the common pole model and direct LSE FIR model are plotted in Figure 6.31 and Figure 6.32.
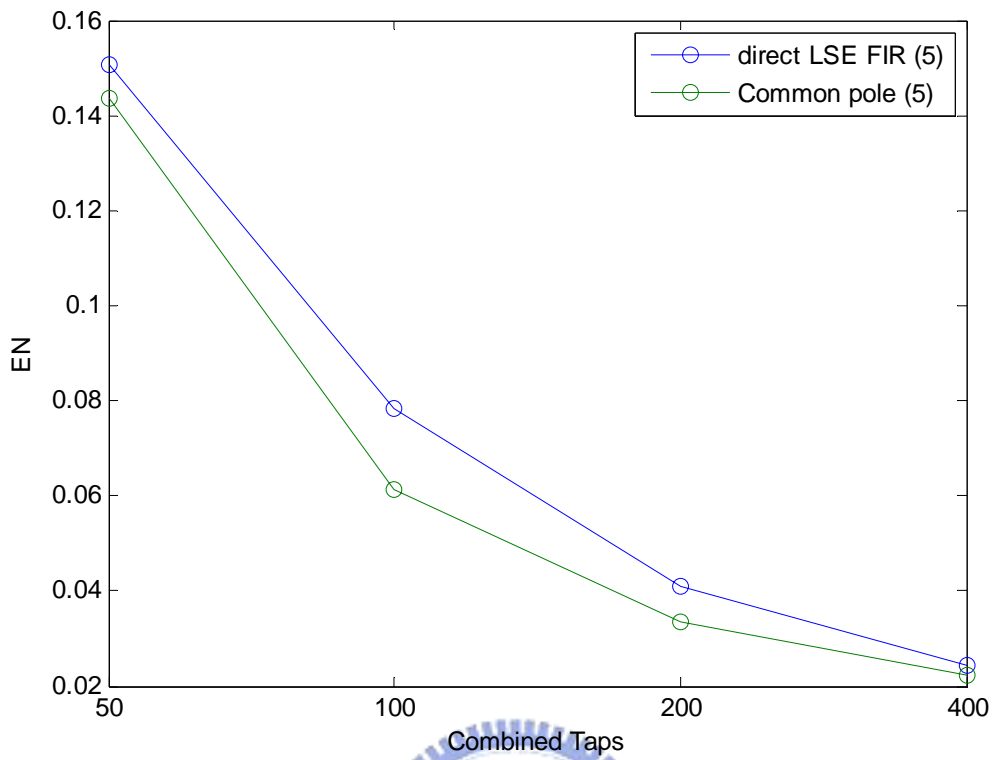
Figure 6.31: Comparison of *EN* with direct LSE FIR and Common-pole IIR at $\pm 5^o$
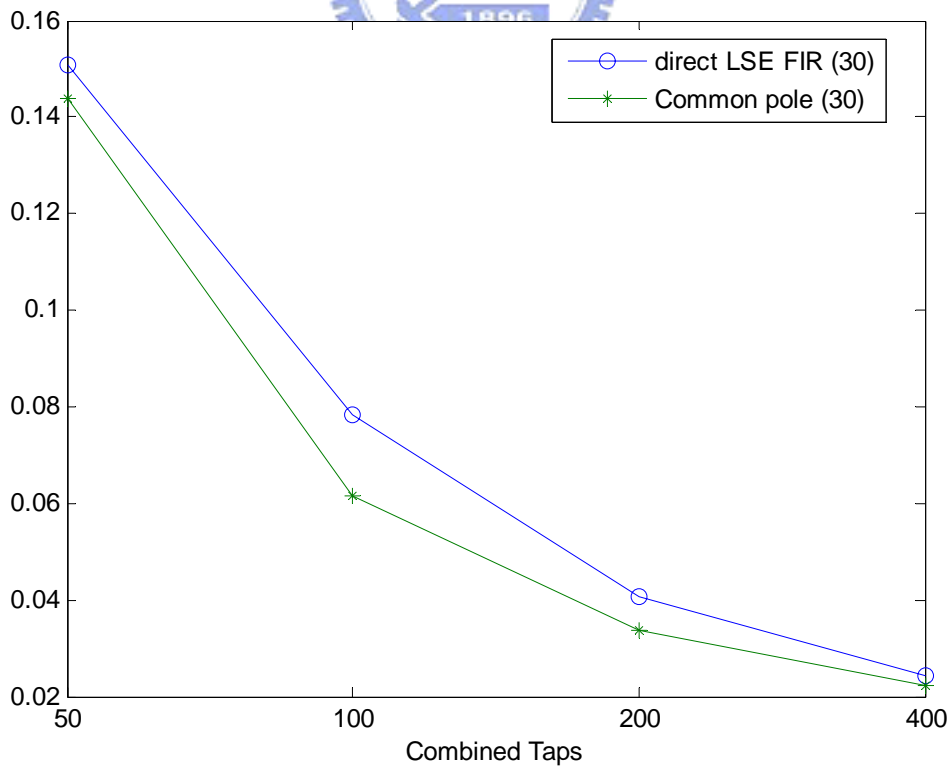
Figure 6.32: Comparison with direct LSE FIR and Common-pole IIR at $\pm 30^o$

From these above two figures, the performance of the common pole model is better than the direct LSE FIR model. However, we want to know that the improvement performance of the common pole model is from the IIR structure or the common part, so we compare the performance between common pole model and direct LSE FIR model with same total taps instead of combined taps. The *EN* results loudspeaker pair placed at $\pm5^o$ and $\pm30^o$ are plotted in Figure 6.33 and Figure 6.34.
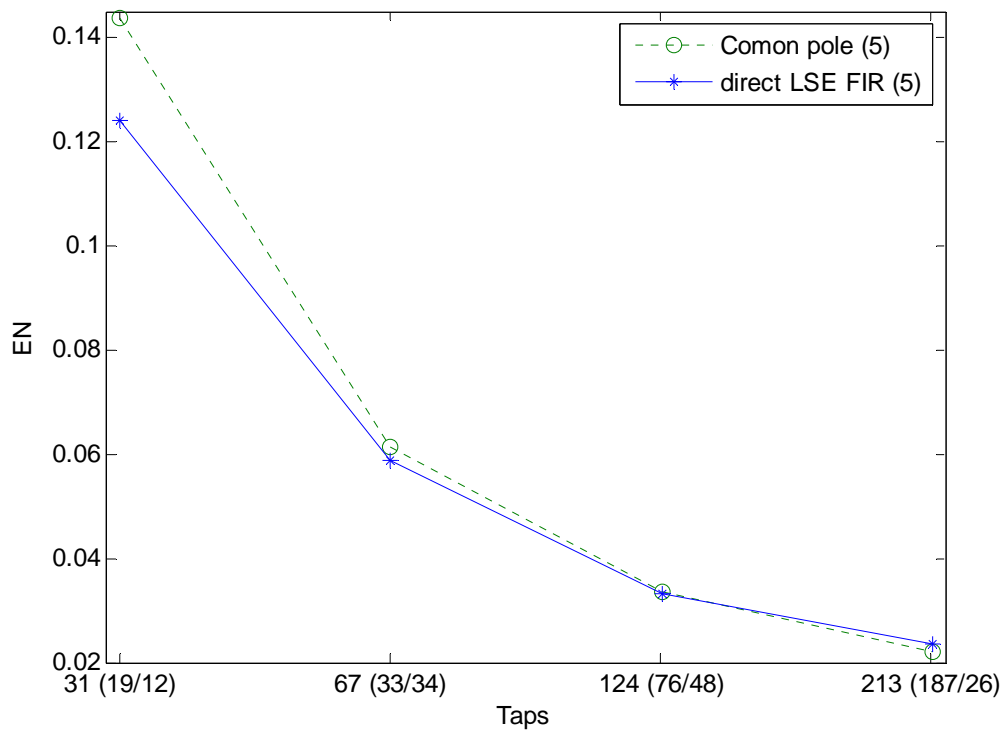
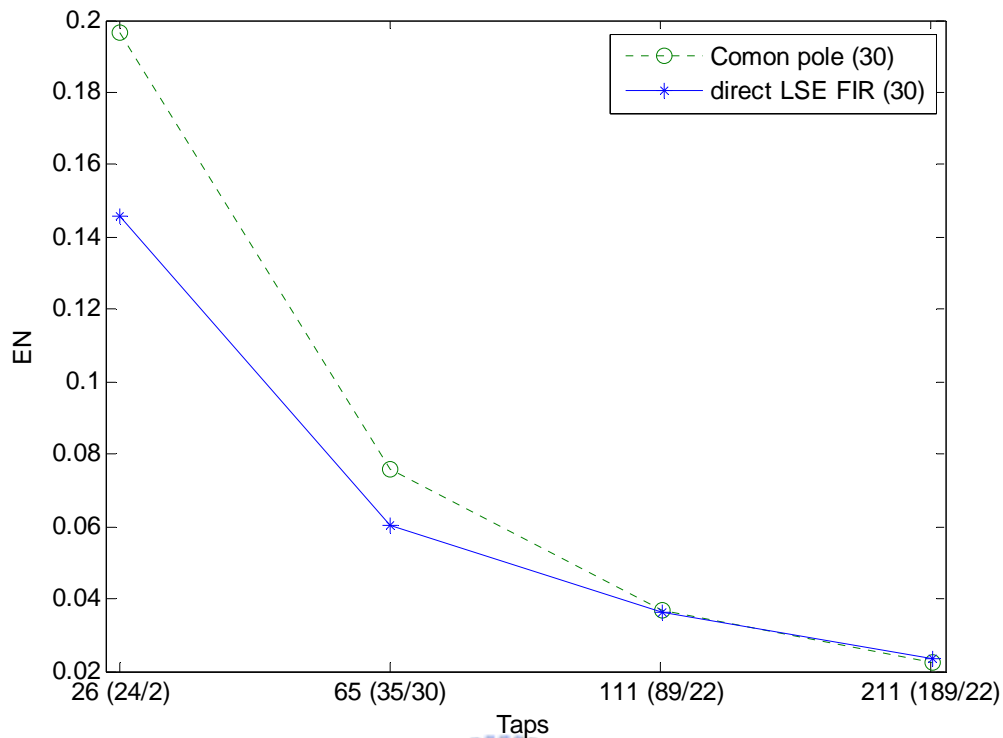Figure 6.33: Comparison between common pole and direct LSE FIR using same total taps at $\pm5^o$

Figure 6.34: Comparison between common pole and direct LSE FIR using same total taps at $\pm 30^{o}$

From Figure 6.33 and Figure 6.34, the *EN* of the direct LSE FIR is better than common pole model with small taps. In other words, we can know the performance of the common pole model is from the common part and not from the IIR.

6.3.2.1 Design in Frequency Domain

In this section, we will simulate the common pole model in the frequency domain, and find *EN* of difference total taps. The results loudspeaker pair placed at $\pm 5^{o}$ and $\pm 30^{o}$ are shown in Table 6.15 and Table 6.16.

Total tap \ Category	50	100	200	400
Time Domain	0.1437	0.0614	0.0336	0.0222
Frequency Domain	0.1437	0.0614	0.0336	0.0222

Table 6.15: Comparison Common-pole IIR at $\pm 5^o$ in time and frequency domains

Taps \ Category	50	100	200	400
Time Domain	0.1964	0.0756	0.0369	0.0224
Frequency Domain	0.1964	0.0756	0.0369	0.0224

Table 6.16: Comparison Common-pole IIR at $\pm 30^o$ in time and frequency domains

From the above tables, the performance of design in the time domain is the same as design in the frequency domain.

6.4 Robust Crosstalk Canceller

As we know, the head movement will eliminate the performance of the crosstalk cancellation. We first design a crosstalk canceller with the loudspeakers placed at $\pm 30^o$. We design the crosstalk canceller by using the direct LSE FIR method. Each filter length is 200 and the result is plotted in Figure 6.35. Figure 6.36 and Figure 6.37 show that the received signals in the frequency domain at both ears after the head rotating around $\pm 5^o$. We can find that the received signals at rotated head compare poorly with that the received signals at the fixed head. The flatness of these two figures in the high frequency band is bad, and the crosstalk suppression is, either. The FOM in Figure 4.37 are $EN = 0.3237$, $CSF = 4.486$ dB and $EIF = 9.249$ dB^2. Those in Figure 6.38 are $EN = 0.3325$, $CSF = 5.747$ dB and $EIF = 9.785$ dB^2.
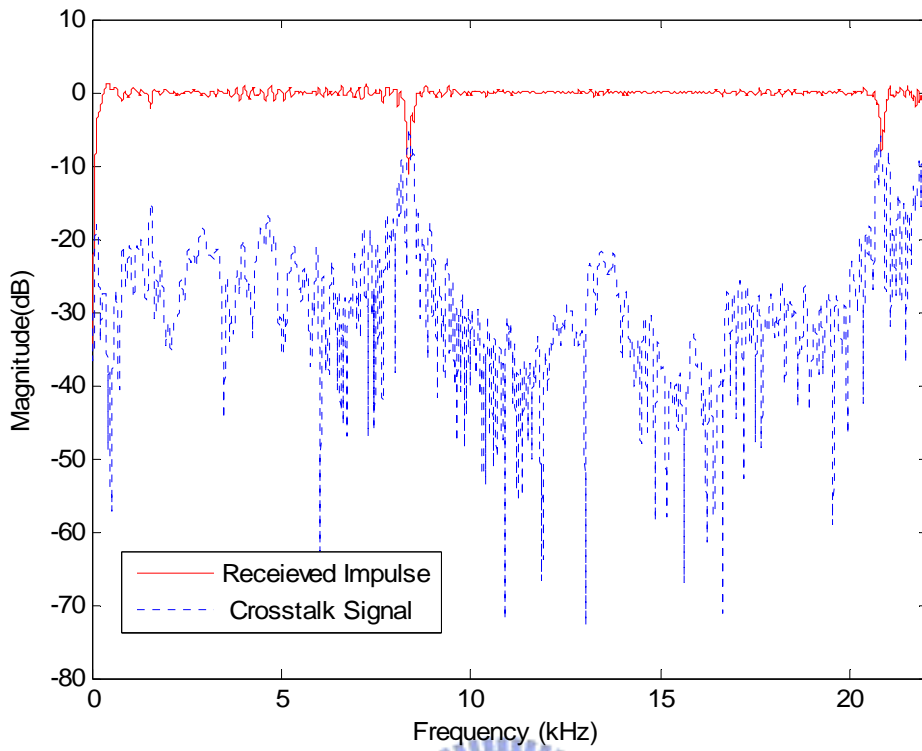
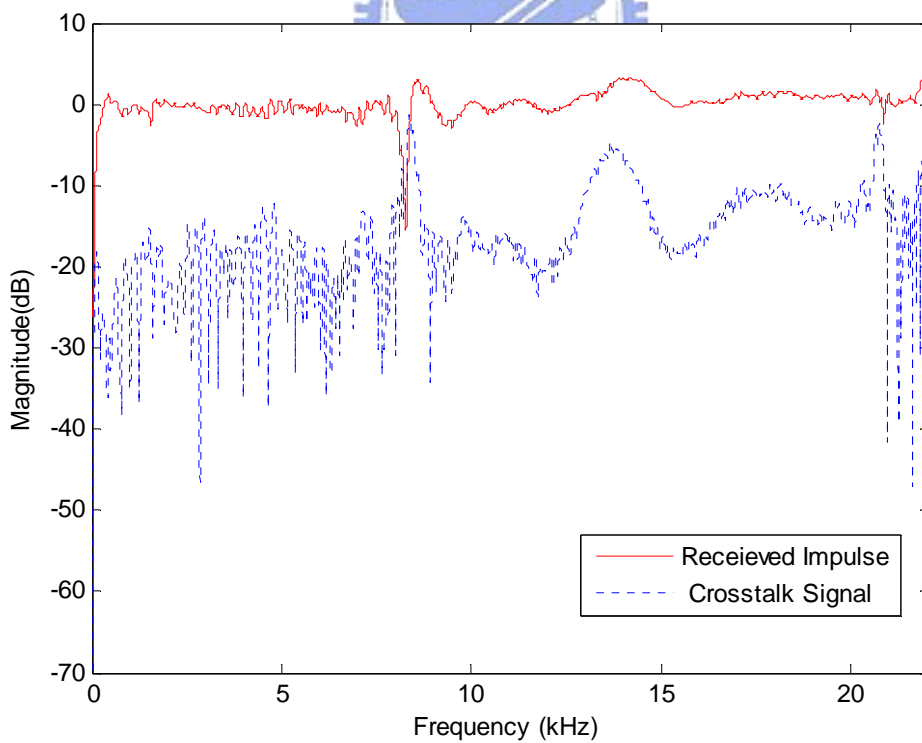Figure 6.35: The frequency response at $\pm30^o$ with 200-tap FIR designed using LSE

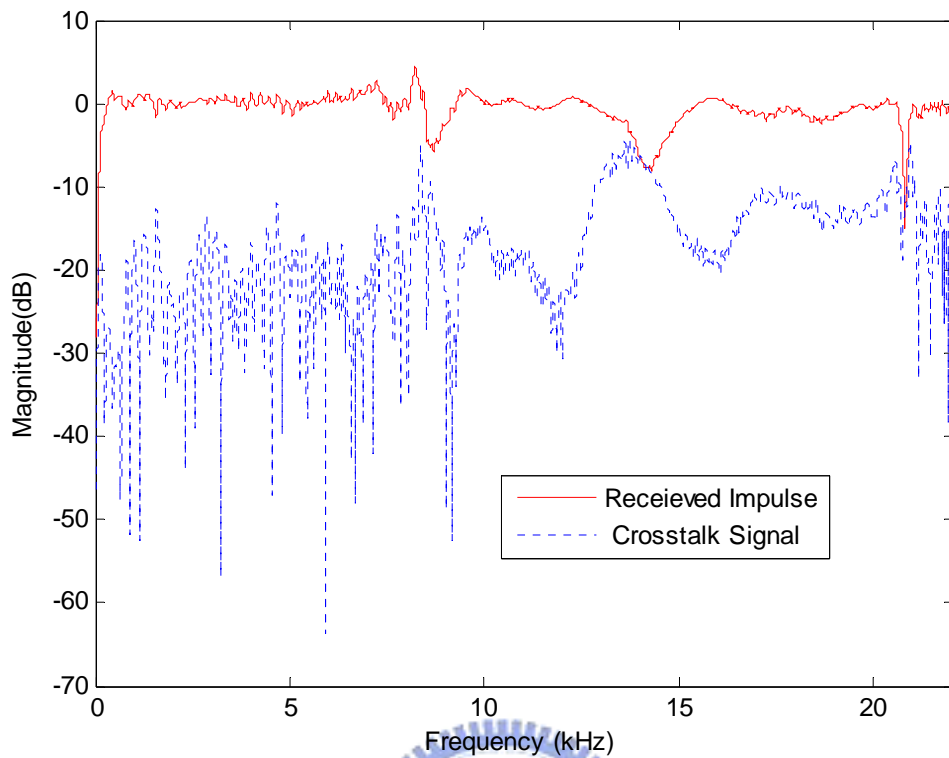Figure 6.36: The frequency response with head rotated $+5^o$ at $\pm30^o$

Figure 6.37: The frequency response with head rotated -5° at $\pm30^o$

Compared with the FOM of the fixed head, we can find the *EN* and *CSF* are worse very much. Therefore, we should take the robust crosstalk canceller to eliminate the effect of the head movement.

When we design the robust crosstalk canceller, the delay compensation has to be considered. Figure 6.38 and Figure 6.39 show the frequency response of the received signals using robust design without delay compensations. The received signals in Figure 6.38 are at the fixed head and these in Figure 6.39 are at rotated head. Obviously, we can find the equalization is very bad in these two figures.
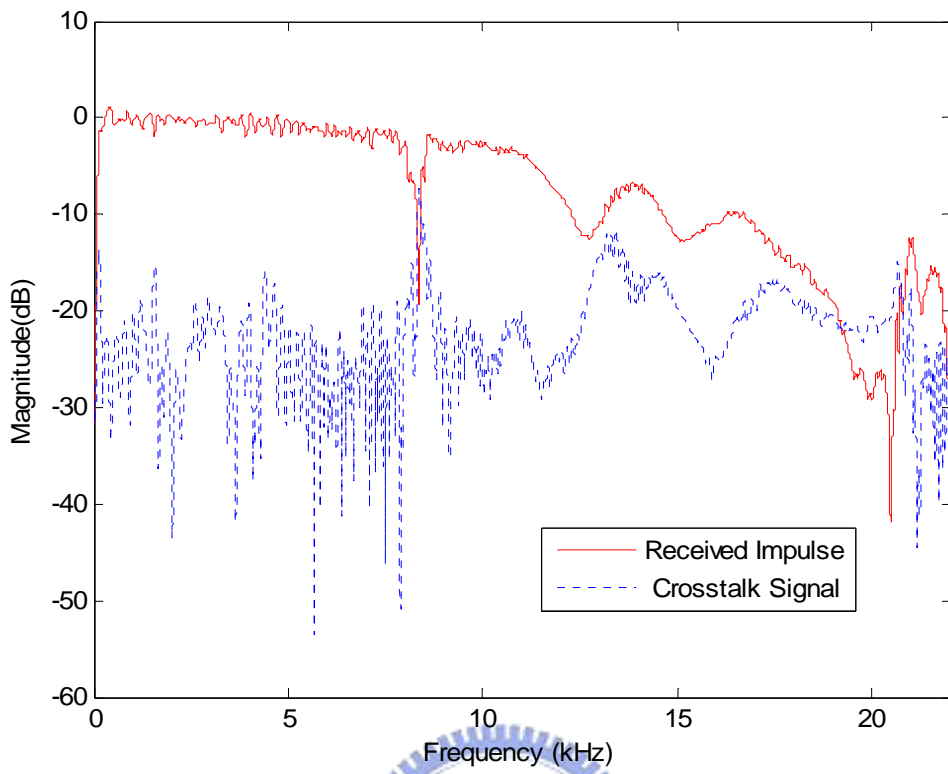
Figure 6.38: The frequency response at fixed head without compensation
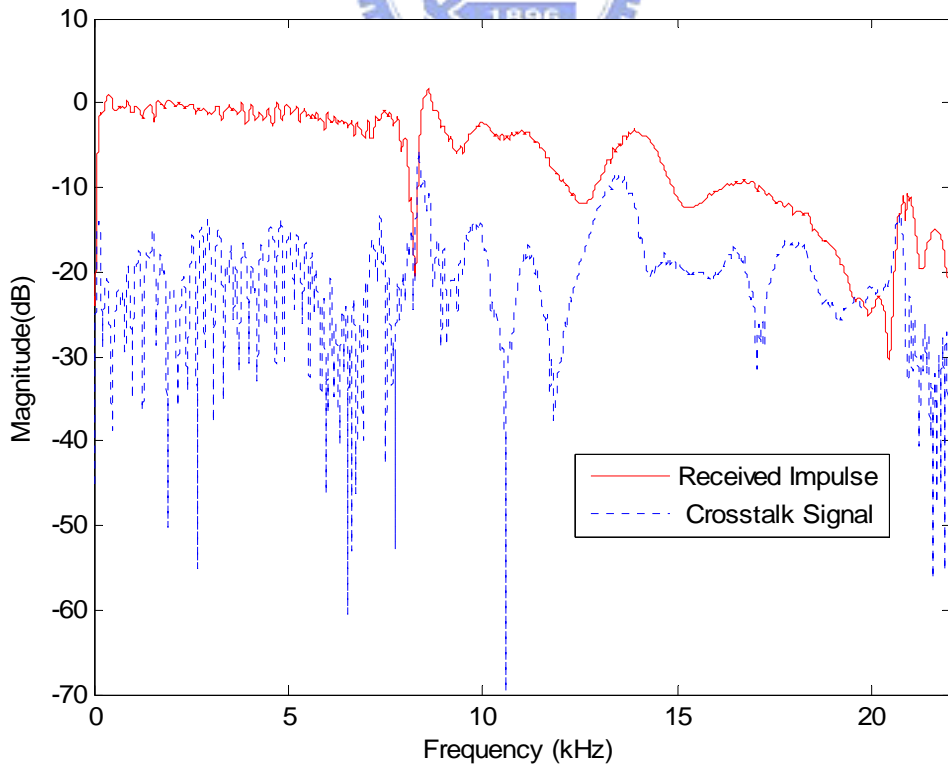

Figure 6.39: The frequency response at rotated head without compensation

Referring to from Equation (5.5), the delay compensation is calculated, and their value are ± 1. The desired compensated signal can be rewritten referring to Equation (5.6). Figure 6.40, Figure 6.41 and Figure 6.42 are the frequency responses with compensations. We can find the equalization is improved very much. The performance of the robust crosstalk canceller will be listed in Table 6.13 to compare with and without considering robustness.
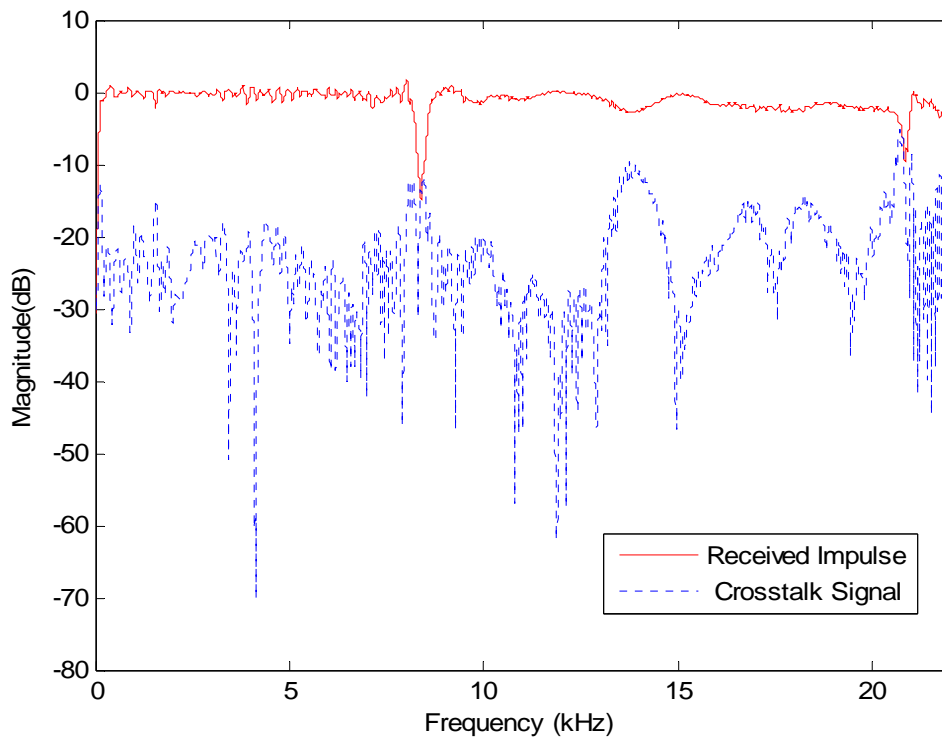

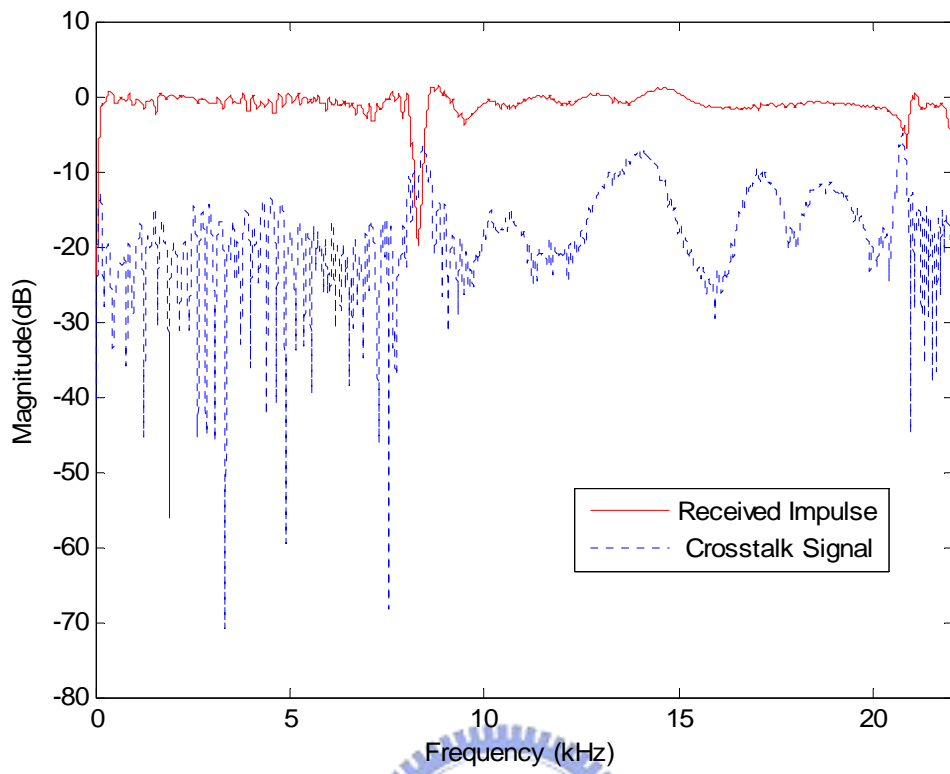Figure 6.40: The frequency response at fixed head with compensation

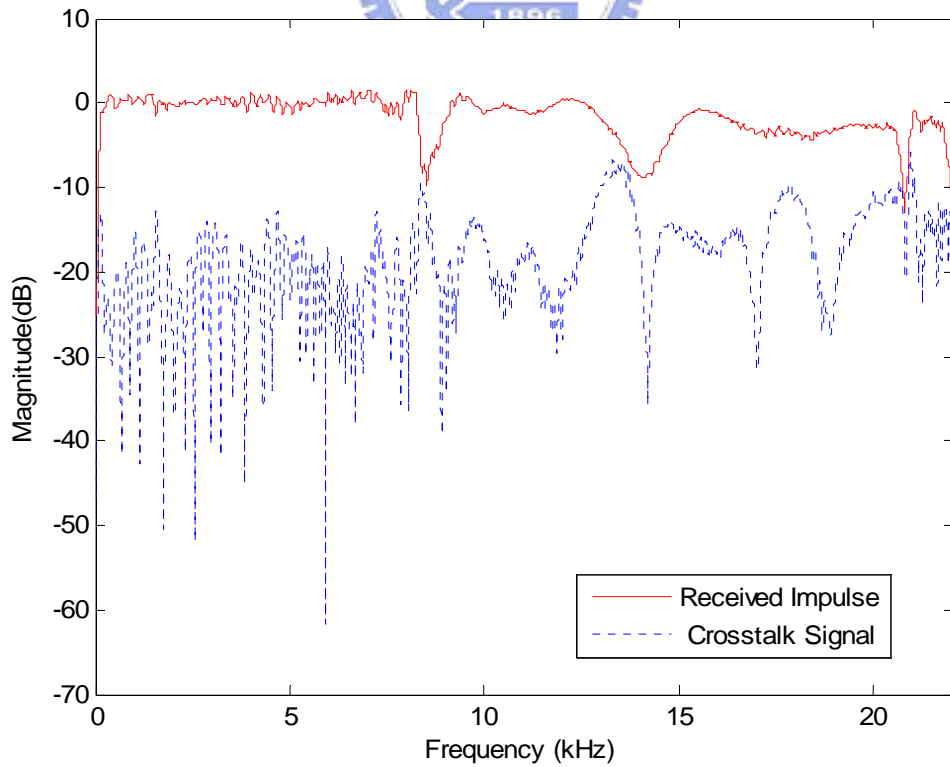Figure 6.41: The frequency response at rotated $+5^o$ head with compensation

Figure 6.42: The frequency response at rotated -5° head with compensation

In Table 6.17, 0°, $+5^o$ and -5° represent the places of the head. 0° is the fixed, $+5^o$ is rotated right around 5° and -5° is rotated left around 5°.

FOM / Category	Total *EN*			*CSF* (dB)			*EIF* (dB^2)		
	0°	$+5^o$	-5°	0°	$+5^o$	-5°	0°	$+5^o$	-5°
Non-robust	0.6814			11.99	4.49	5.75	10.11	9.25	9.79
	0.025	0.32	0.33						
Robust	0.5541			9.42	5.40	6.54	9.76	9.18	9.45
131	0.055	0.22	0.28						

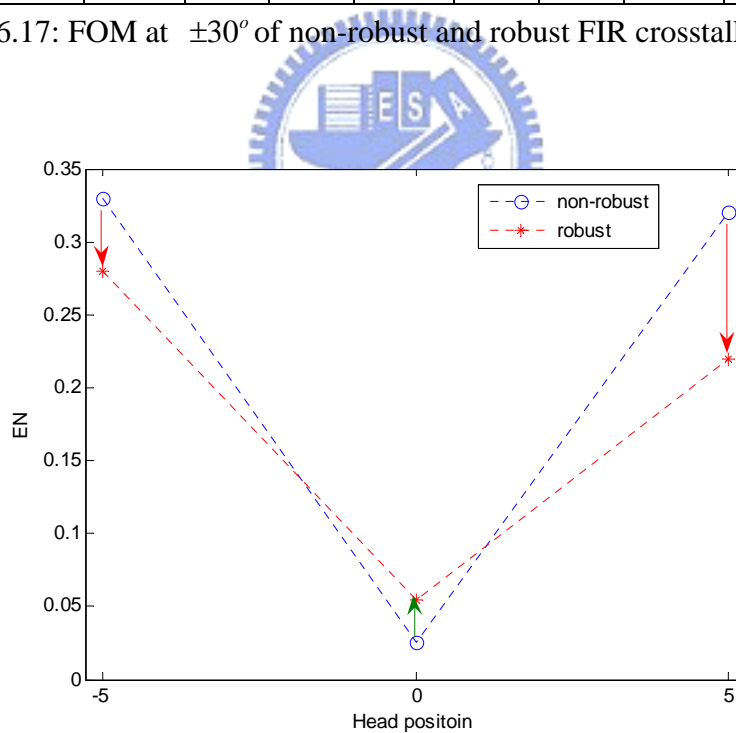Table 6.17: FOM at $\pm30^o$ of non-robust and robust FIR crosstalk cancellers

Figure 6.43: *EN* between robust design and non-robust design at $\pm30^o$

We list the each *EN*, *CSF*, and *EIF* at different places. Besides, we also list the total *EN* to compare the performance of the robust crosstalk cancellation. From Table 6.17 and Figure 6.43, we can find that the total *EN* becomes smaller with the robust

crosstalk canceller than the one with non-robust crosstalk canceller. Therefore, the performance gets better. The *EN* of the fixed head becomes bad, but those of the rotated heads become better. Because our method is to strike a balance in a region, we do not handle a specific place. Our goal is to minimize the total error in the region. The next table is to show that the performance gets better as the filter length increases.

Taps / Category	25	50	100	200
Non-robust	1.019	0.7768	0.6926	0.6814
Robust	0.9593	0.6863	0.5904	0.5541

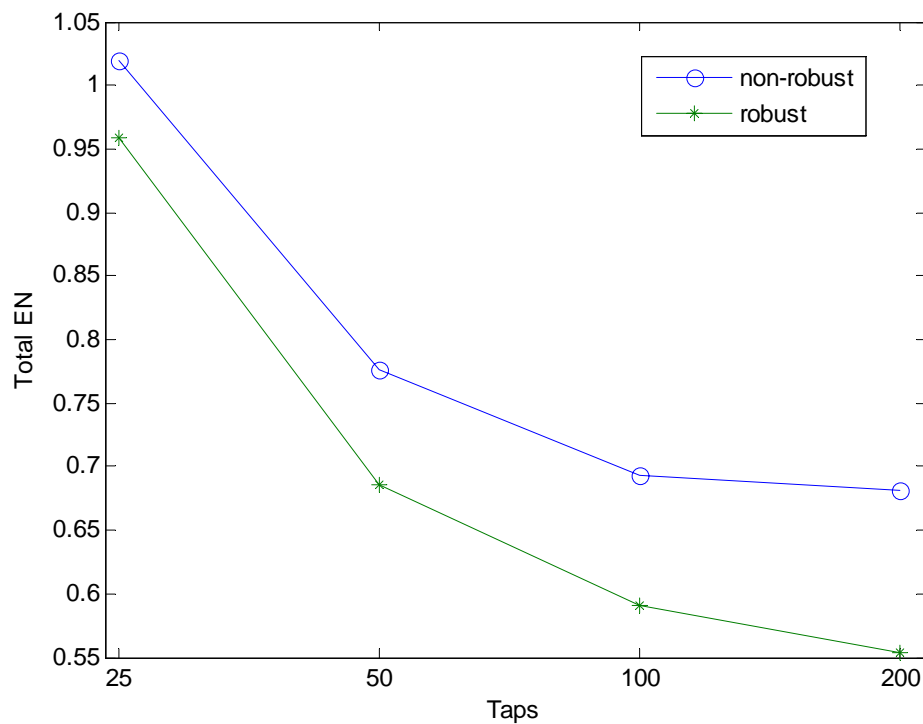Table 6.18: Total *EN* with different length at $\pm 30^o$

Figure 6.44: Total *EN* between robust design and non-robust design at $\pm 30^o$

From Table 6.18 and Figure 6.44, it is obvious that the performance with robust crosstalk cancellers is better than that with non-robust crosstalk cancellers.

Next, we will try the small angle at $\pm 5^o$. Figure 4.45 is the result at the fixed

head, and Figure 4.46 and Figure 4.47 show the received signals at rotated-head

places with non-robust crosstalk canceller designed by LSE 200-tap FIR. Similarly,

we must calculate the delay compensation first. From the delay-compensated equation,

we can find the compensated delay is ±1. Then, we can design the robust crosstalk

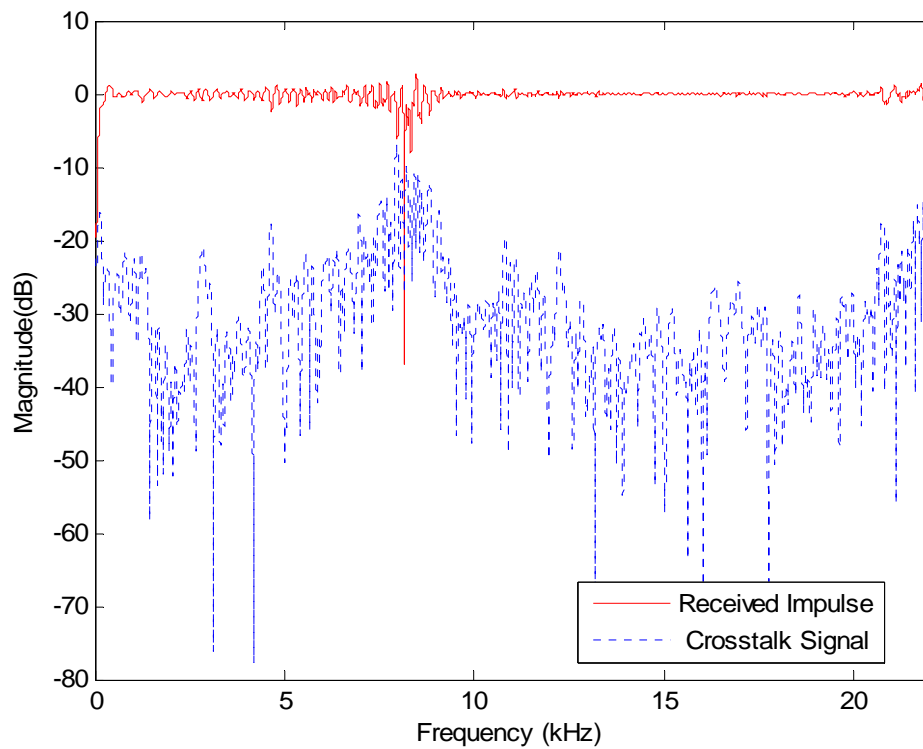canceller. Here, we use the crosstalk canceller with 200 taps and extra 121 sample

delay.

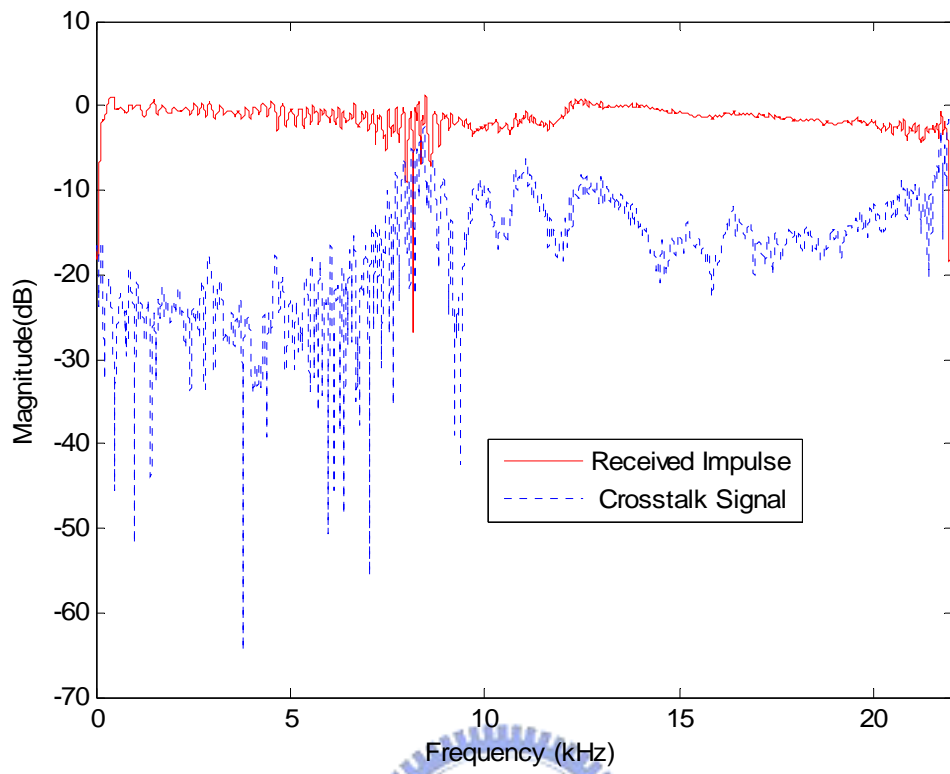Figure 6.45: The frequency response at $\pm 5^{o}$ with 200-tap FIR designed using LSE

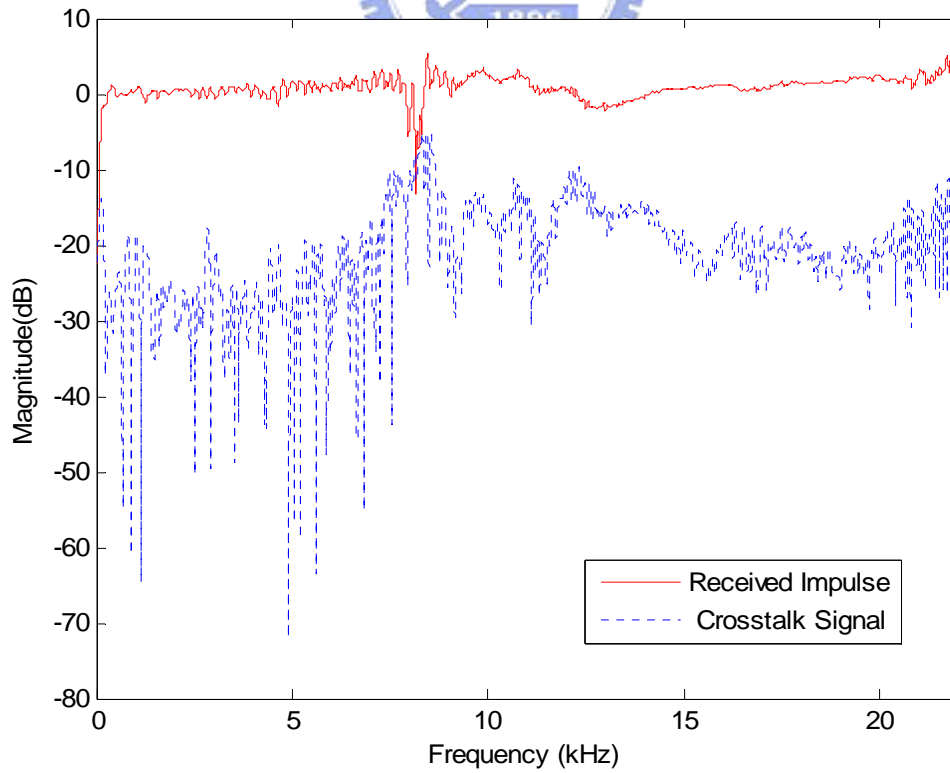Figure 6.46: The frequency response with head rotated $+5^o$

Figure 6.47: The frequency response with head rotated -5°

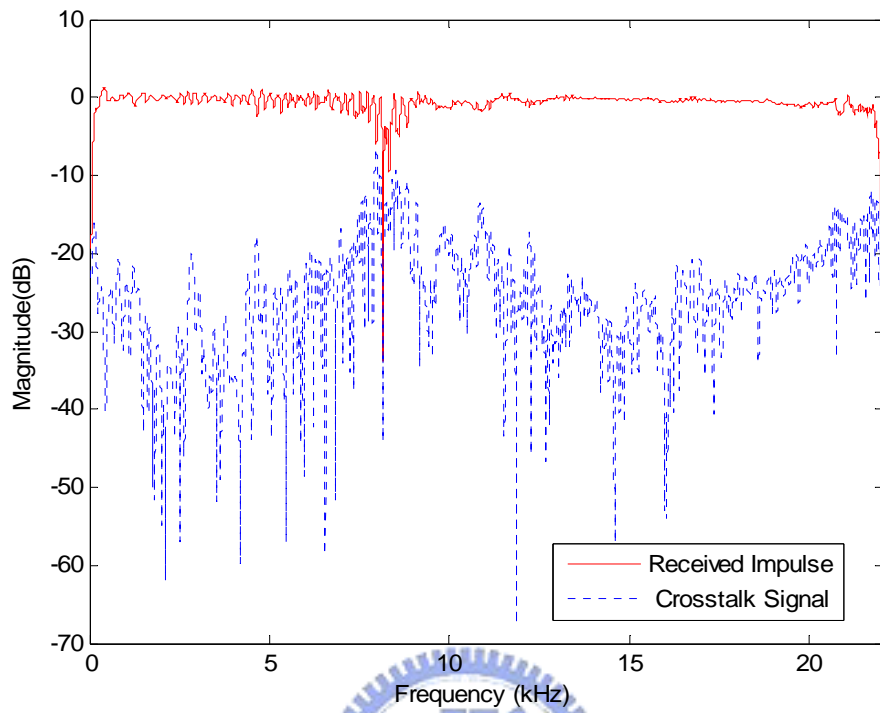Figure 6.48, Figure 6.49 and Figure 6.50 are the received signals at three locations where the head moved.

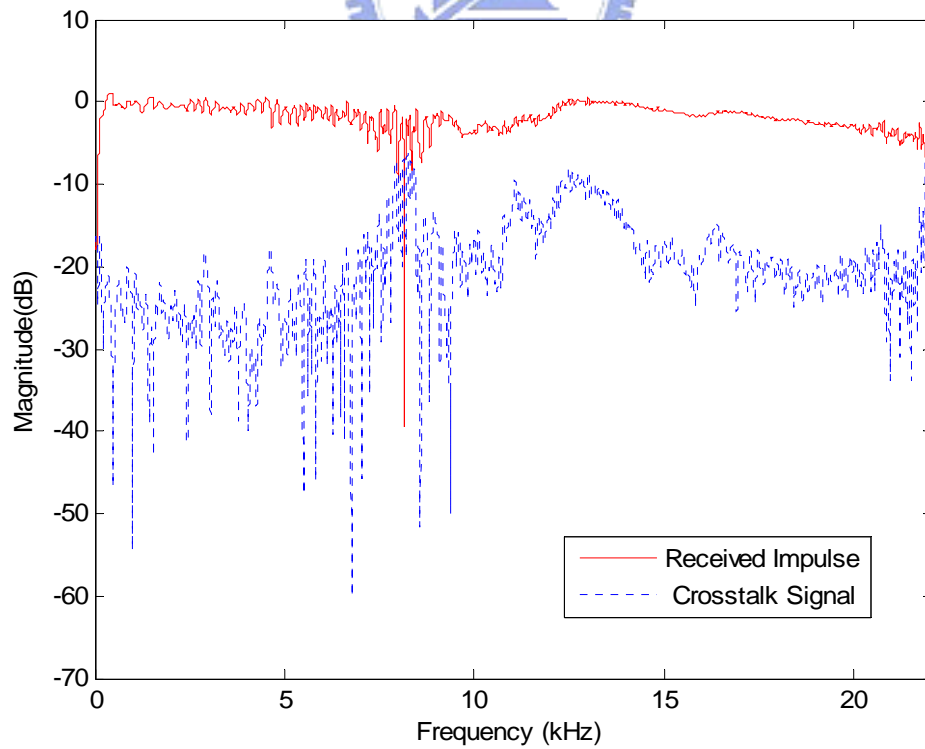Figure 6.48: The frequency response at fixed head with compensation

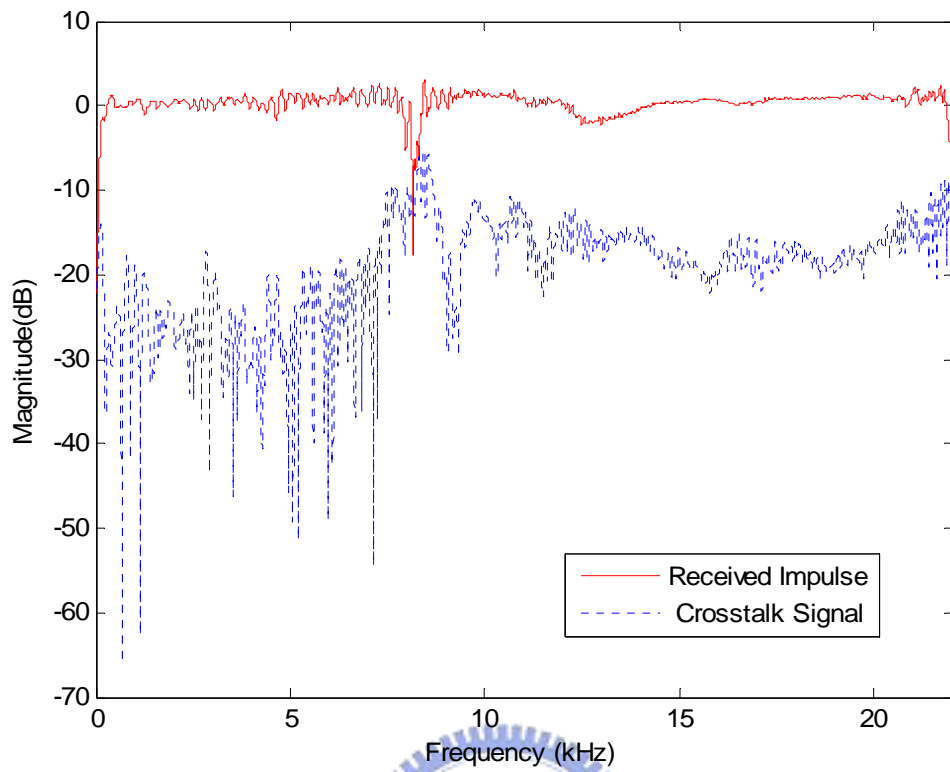Figure 6.49: The frequency response at rotated $+5^o$ head with compensation

Figure 6.50: The frequency response at rotated -5° head with compensation

The FOM of the three locations is listed in detail in Table 6.19.

FOM	Total *EN*			*CSF* (dB)			*EIF* (dB^2)		
Category	0°	$+5^o$	-5°	0°	$+5^o$	-5°	0°	$+5^o$	-5°
Non-robust	0.3002			21.18	9.85	16.57	10.16	10.05	9.40
	0.0244	0.1395	0.1364						
Robust	0.2579			18.54	12.86	15.29	9.97	9.79	9.50
121	0.0348	0.1301	0.093						

Table 6.19: FOM at $\pm5^o$ of non-robust and robust FIR crosstalk cancellers
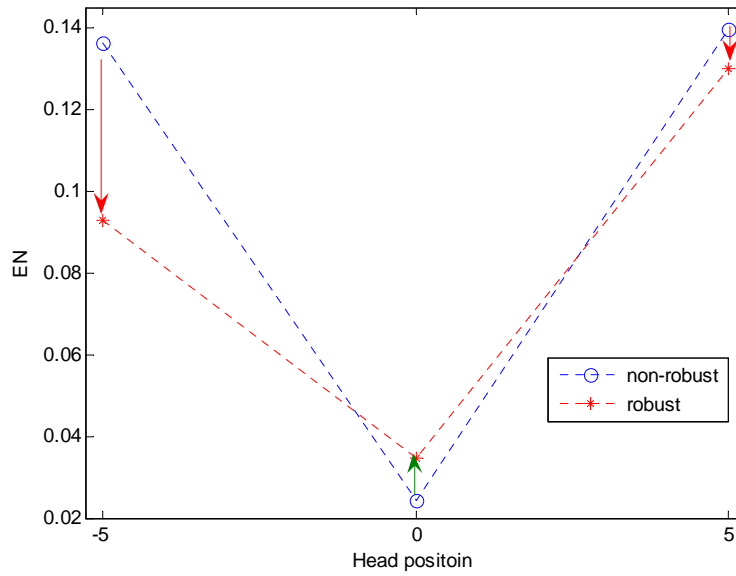
Figure 6.51: *EN* between robust design and non-robust design at $\pm 5^{o}$

From Table 6.19 and Figure 6.51, the total *EN* with robust crosstalk canceller is

smaller than the one with non-robust crosstalk canceller. Similarly, the origin becomes

worse but other places are improved. Table 6.20 is listed the results compared total

EN with different taps. From Figure 6.52 and Table 6.20, we can know that the robust

crosstalk canceller can reduce the total *EN*.

Taps Category	25	50	100	200
Non-robust	0.6172	0.4140	0.3281	0.3002
Robust	0.5991	0.3908	0.2959	0.2579

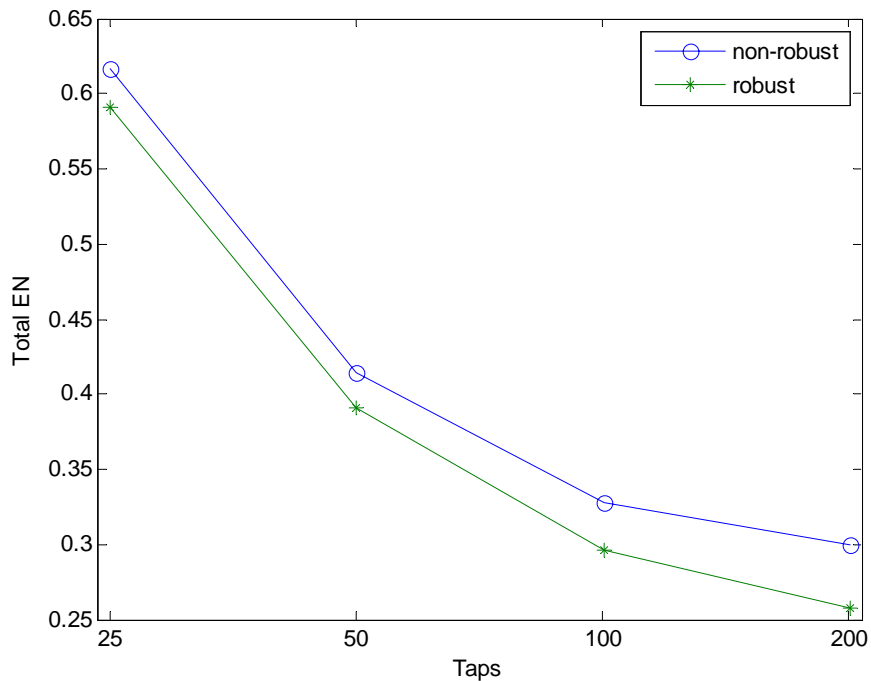Table 6.20: Total *EN* with different length at $\pm 5^{o}$

Figure 6.52: Total *EN* between robust design and non-robust design at $\pm 5^o$

Next, we will test the common-pole model in robust method. We test different

combined tap number of the robust common pole model and compare with design in

robust direct LSE FIR model. The results of total *EN* with the loudspeaker pair at

$\pm 5^o$ are listed in Table 6.21.

Taps / Category	50	100	200	400
Direct LSE FIR	0.5991	0.3908	0.2959	0.2579
Common pole IIR	0.5991	0.383	0.2879	0.2574

Table 6.21: Comparison between robust direct LSE FIR and
robust common pole IIR at $\pm 5^o$

We also test the loudspeaker pair at $\pm 30^o$ and the results are the same as the

robust direct LSE FIR model.

Chapter 7

Conclusions

In the thesis, we investigate two methods, matrix-inverse and LSE to design crosstalk cancellers and each method can be implemented in FIR or IIR forms. Matrix- inverse design in the frequency domain is better than in the time domain. In FIR form, we can find that the performance of LSE FIR is better than that designed from matrix-inverse. In IIR form, the common-pole model is also better than that designed from matrix-inverse. Compare LSE FIR with common-pole model, the common pole model is better because the LSE FIR model can be seen as the special case of common-pole model.

Then we propose the region-equalized method to eliminate the effect of the head movements. In the simulations, we can find that the performance at the origin place is worse than that designed with single one place, but the performance of other places are improved. It is proved that the region-equalized method can reduce the total difference between the received signals and the desired signals.

Bibliography

[1] Durand R. Begault, "3-D Sound for Virtual Reality and Multimedia AP Professional," 1994.

[2] W. G.. Gardner, "3-D Audio Using Loudspeakers," Wave Art, Inc, 1998.

[3] B. Gardner, K. Martin, "HRTF measurements of a KEMAR dummy-head microphone", MIT. Media Lab Perceptual Computing Technical Report No. 280, 1994.

[4] J. L. Bauck and D. H. Cooper, "Generalized transaural stereo and applications", J. Audio Eng. Soc., vol. 44, no. 9, pp. 683-685, Sept. 1996.

[5] Lopez-Poveda, E. A., and Meddis, "A physical model of sound diffraction and reflections in the human concha", J. Acoust. Soc., 3248–3259, 1996.

[6] J. Blauert, "Spatial Hearing: The Psychophysics of Human Sound Localization", M.I.T. Press, Cambridge, MA, 1983.

[7] M.R. Schroeder and B.S. Atal "Computer simulation of sound transmission in rooms," IEEE Conv. Record, 7:150-155, 1963.

[8] P. Y. Wang, "3D Sound Reproduction via Two Loudspeakers," NCTU Master These, 1997.

[9] D. B. Ward and G. W. Elko, "Optimum loudspeaker spacing for robust crosstalk cancellation", in Proc. IEEE ICASSP' 98, Seattle, WA, vol. 6, pp. 3541-3544.

[10] D. B. Ward and G. W. Elko, "Effect of loudspeaker position on the robustness of acoustic crosstalk cancellation," IEEE Signal Processing Lett,, vol. 6, pp. 106-108, May 1999.

[11] D. B. Ward and G. W. Elko, "A new robust system for 3-D audio using loudspeakers," In Proceedings of ICASSP 2000, Istanbul, Turkey, June 2000.

[12] J. Yang, W. S. Gan, and S. E. Tang, "Development of virtual sound imaging system using triple elevated speakers," IEEE Trans. on Consumer Electronics, vol. 50, no. 3, Aug. 2004.

[13] A. Mouchtaris, P. Reveliotis, and C. Kyriakakais, "Inverse filter design for immersive audio rendering over loudspeaker," IEEE Trans. on Multimedia, vol. 2, no. 2, June 2000.

[14] A.V. Oppenheim and R. W. Schafer, "Discrete Time Signal Processing", Prentice Hall, 1999.

[15] M. C. Chen, and S. F. Hsieh, "Common acoustical-poles/zeros modeling for 3D sound processing," IEEE ICASSP'2000、pp. 785-788.

[16] Y. F. Chow, "Acoustical Spatial Equalization," NCTU Master These, 1994.