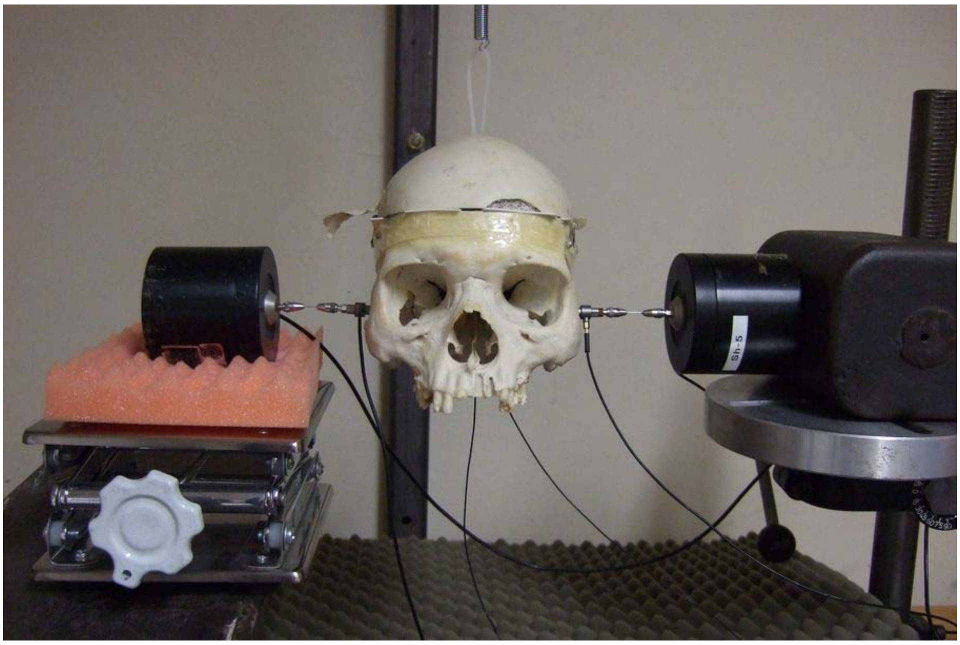
CHALMERS



Application of cross-talk cancellation to the improvement of binaural directional properties for individuals using bone anchored hearing aids (BAHA)

Master's Thesis in the Master's programme in Sound and Vibration

CHIAO-LING LIAO

Department of Civil and Environmental Engineering
Division of Applied Acoustics
Chalmers Room Acoustics Group
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2010

Master's Thesis 2010:10

MASTER'S THESIS 2010:10

Application of cross-talk cancellation to the improvement of binaural directional properties for individuals using bone anchored hearing aids (BAHA)

Chiao-Ling Liao

Department of Civil and Environmental Engineering
Division of Applied Acoustics
Chalmers Room Acoustics Group
CHALMERS UNIVERSITY OF TECHNOLOGY
Göteborg, Sweden 2010

Application of cross-talk cancellation to the improvement of binaural directional properties for individuals using bone anchored hearing aids (BAHA)

Master's Thesis 2010:10

Department of Civil and Environmental Engineering
Division of Applied Acoustics
Chalmers Room Acoustics Group
Chalmers University of Technology
SE-41296 Göteborg
Sweden

Tel. +46-(0)31 772 1000

Reproservice / Department of Civil and Environmental Engineering
Göteborg, Sweden 2010

Application of cross-talk cancellation to the improvement of binaural directional properties for individuals using bone anchored hearing aids (BAHA)
Master's Thesis in the Master's programme in Sound and Vibration
Chiao-Ling, Liao
Department of Civil and Environmental Engineering
Division of Applied Acoustics
Chalmers Room Acoustics Group
Chalmers University of Technology

Abstract

In Bone Anchored Hearing Aids (BAHA), waves stimulated by one BAHA travel to not only the ipsilateral cochlea, but also travels to the contralateral cochlea. This is referred to as crosstalk in the skull in this thesis. It decreases the direction sensing ability for individuals who wear one or two BAHAs. This is especially true at frequencies below 1 kHz, at the so-called anti-resonance at which the contralateral cochlea perceives higher energy than the ipsilateral cochlea. At the higher frequency, the natural attenuation of vibration signal in human head can improve the binaural directional properties.

In this thesis, a crosstalk cancellation algorithm is applied to make an artificial attenuation to the intercranial wave signal. It works from 250 Hz to 1 kHz and attenuates the crosstalk signal 2 to 10 dB at different frequencies.

Keywords: Crosstalk, Crosstalk Cancellation, Bone Anchored Hearing Aid, Intercranial Transfer Function, ILD

Contents

Acknowledgements

To accomplish my master thesis, I would like to express my gratitude to Professor Mendel Kleiner. He guided me through the thesis and always provided me very good and inspirational answers. I would also express my gratitude to Mr. Kristian Åsnes in Cochlear for providing this interesting topic and giving me great advice. Moreover, I would like to thank Börje Wijk who always smiled and helped me with the equipment and measuring set-ups.

I would like also thank all the people that helped me and cheered me up during the period of doing this thesis. Especially the one who found a skull for me that I could use during the experiments, and the one who helped me a lot with DSP board.

At the end, I would like to thank my family; they support me to come to Sweden and make my dream come true.

1. Introduction

Bone anchored hearing aids have been investigated for around twenty years. It helps many people with impaired hearing to hear the world. The application started from unilateral fitting of BAHA; few years later, people tried to use bilateral fitting at two sides of head. It is approved [1, 2, 3] to have more benefits compare to unilateral fitting situation. However, the crosstalk signals that transfer to the contralateral cochlea makes the binaural processing of the patients who are fitted bilaterally with BAHAs less than for people with normal hearing.

The crosstalk problem is a quite common issue in the electronics and audio area. The application of crosstalk cancellation to the BAHA has never been tried before. It is interesting to know that how the crosstalk cancellation will improve the binaural processing and what will be the application limitations.

In this thesis, four crosstalk cancellation algorithms are investigated to find which one of them that is most suitable to cancel the crosstalk signals in the skull. The investigation is carried out by doing the measurement on a dry skull.

The topic of this thesis was suggested by Mr. Kristian Åsnes in Cochlear.

2. The hearing and Bone Anchored Hearing Aids (BAHA)

2.1. Hearing

2.1.1. Air conducted hearing

When mentioning "hearing", what most people think of is the auditory system shown in Figure 2.1. Modern audiology theory was established by Békésy in the middle of 20th century. The sound in the space perceived by a person is processed through a series of stages: first, the auditory signal is captured by the ear pinna and transferred to the middle the ear through ear canal. In the end of the canal a ear drum is attached to the ossicles. When the sound wave stimulates the ear drum, the sound wave therefore converts into vibration to the oval window of the cochlea through structure the borne path. The basilar membrane (BM) of the cochlea works analogue of an electrical filter, moves differently according to incoming vibration signals, and the mechanical movement excites the hearing cells and, at the end, to the brain makes the hearing sense function works. To perceive the sound though this path is defined as the air conduction path.

In later investigations, researchers had found there is another path that makes people hears, the bone conduction path, which is also the main property used in bone anchored hearing aids. This will be introduced in next section.

2.1.2. Bone conducted hearing

The hearing mechanism of a human, according to Tonndorf, can be made as a block diagram shown in Figure 2.2. There are three paths that the sound transfers to the inner ear. The air conduction path is introduced in previous section, the rest of two paths are processed via the temporal bone, which is a part of the human head bones and is indicated in Figure 2.3. One can experience the bone conduction by listen to his or her own voice when talking; the sound is different from one's own recorded voice played by the loudspeaker. That is because the sound one hears from the speaker is only through the air conduction path, but when talking, the sound one hears is the combination of the energy transferred via bone conduction and air conduction paths.

According to Tonndorf, there three bone conduction paths that bring the sound to the auditory system.

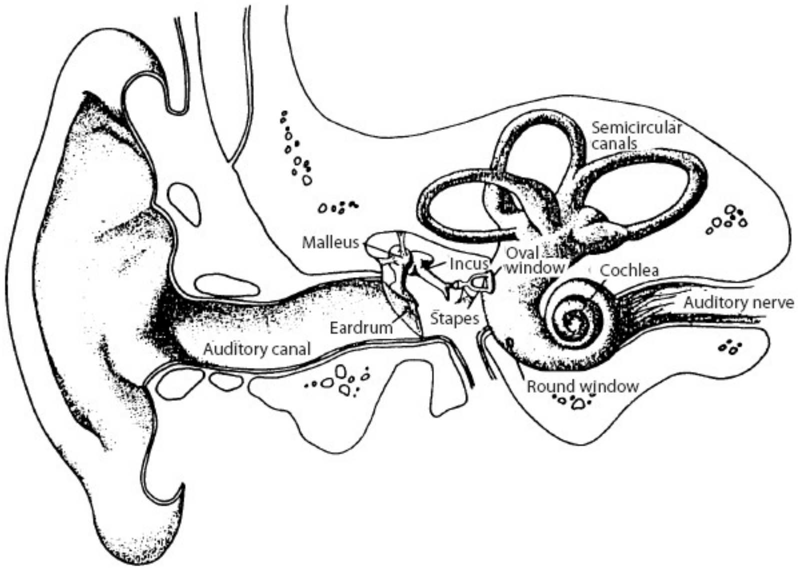


Figure 2.1.: The auditory system[4]

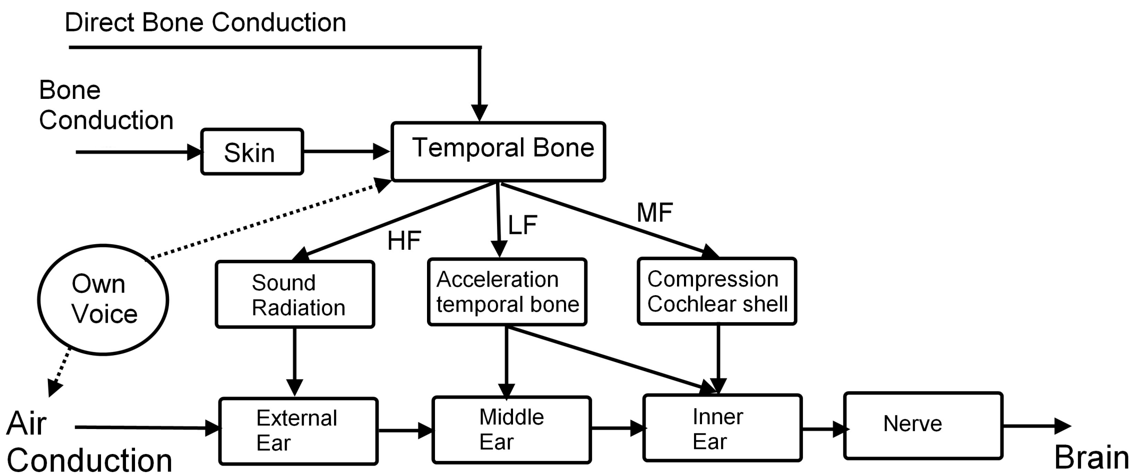


Figure 2.2.: Air and bone conduction schematic. After [5]

1. High frequency: the sound energy radiates into the external ear canal.

2. Low frequency: the inertial response of the middle ear ossicles and the inner ear fluids.

3. Middle frequency: the compression response of the middle ear ossicles and the inner ear fluids.

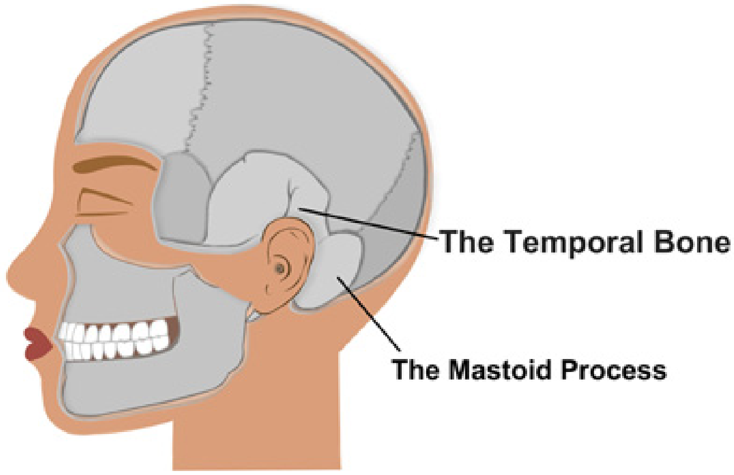


Figure 2.3.: The temporal bone and the mastoid process[6]

2.1.3. Binaural hearing

To be able to tell the direction and how far of the sound source is a very important skill for human beings and animals so that one can detect where is the dangerous thing to prevent or where is the target to get. To detect where, how far and from which direction the sound comes is called "localization". To complete the task, one needs two important cues: interaural level difference (ILD) and interaural time difference (ITD). They are also called binaural cues.

In Figure 2.4, the sound source locates at S, the distance from the source to the right ear is SR, and the distance to the left ear is SL. Since SR is shorter than SL, the traveling time will be different. The longer distance the more time it will take, therefore it makes the ITD. ITD is an important cue especially for the frequency below 1.5 kHz [7]. When the signal travels to the left ear through path SL, the face or the head can be an obstacle that attenuates the signal at high frequency, because the wave length is short compares to the size of the head. The signal to the left ear is therefore attenuated and less then the signal that reaches the right ear; this causes the ILD. The phenomenon is also called "head shadow effect". ILD is more useful at frequencies above 1.5 kHz. The combination of ILD and ITD can cover the whole hearing frequency range and provide the cues of localization.

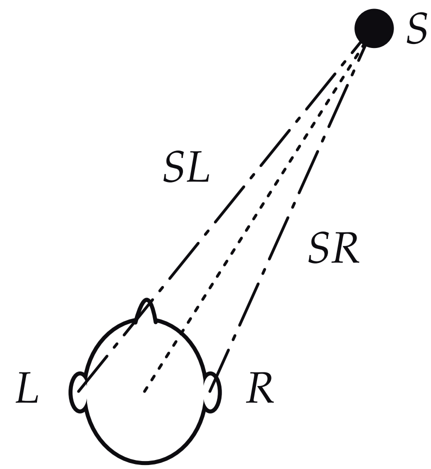


Figure 2.4.: The distance from the sound source to two ears

2.2. Hearing aids

2.2.1. Types of hearing aid

For the three transmitting paths of the hearing mechanism, there are corresponding hearing aids to assist people who have impaired auditory system, but the hypothesis is that the patient's cochleae have be confirmed to be working, otherwise one needs cochlear implant.

The most common type of hearing aid is through the air conduction path, such as in the ear cannel (ITC), behind the ear (BTE) and etc, is designed and manufactured in many companies. For the bone conduction paths, the early design of bone condudcted hearing aid is shown in left side of Figure 2.5. A metal spring hooks the actuator on the mastoid processing of the patient. The force against the head made by the metal spring is kept around 3 to 4 Newton in order to ensure the vibration signal can be transmitted to the head. Since the continued force on the head makes patient uncomfortable, and also because of other reasons [4], this type of bone conduction hearing aid is not well accepted.

Because of the progress of the material technology, one can place a titanium implant into the head bone to transmit the signal without causing irritation or inflammation, which is defined as the direct bone conduction. This method solves the problems caused by the original bone conducted hearing aid, at least the static force on the head can be removed [4]. There are two types of the direct bone conduction hearing aids. The first one is so called "transcutaneous bone conduction hearing aid" which is shown in the right side of Figure 2.5, the other one is "percutaneous bone conduction hearing aid" . The transcutaneous bone conduction hearing aid is composed of external part and the internal part with the skin and the soft tissue in the between. The external part contains a coil for the induction of the signal magnetic flux, the inner part is a permanent magnet. The force

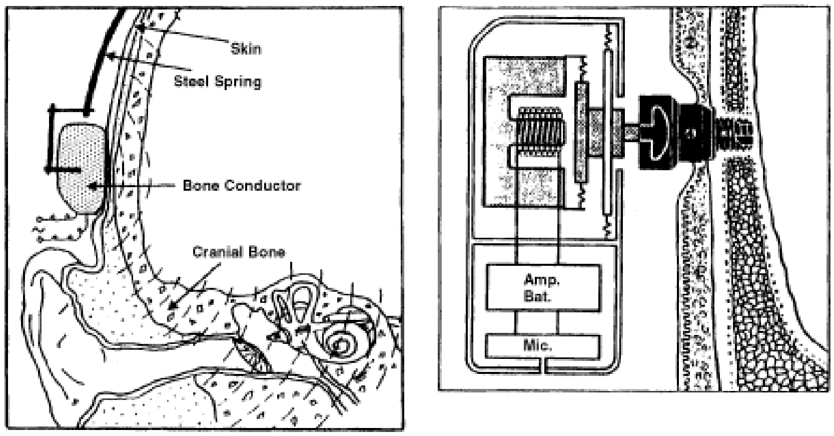


Figure 2.5.: Left:Original bone conduction hearing aid Right:transcutaneous bone conduction hearing aid [5]

between these two parts depends on the thickness of the skin and the damping of the soft issue; these cause problems of the application of the device [4]. A better transducer is therefore expected to solve these drawbacks.

2.2.2. Bone Anchored Hearing Aid

The bone anchored hearing aid (BAHA) is the percutaneous bone conduction hearing aid. It is developed since 1980′ by Anders Tjellström and Bo Håkansson in Göteborg, Sweden. A BAHA is composed of three main parts. The first part is a titanium fixture, similar to the dental implant body, implanted into the mastoid of the head bone where approximately 55mm behind the ear pinna. Second part is the sound processor unit, in which are microphone, digital signal processor (DSP), battery and a force transducer. The third part is an abutment between the other two parts. The implant process takes approximately three months to make the titanium fixture integrated with the bone.

The acoustical signal that is caught by the microphone in the sound processor unit is processed by the DSP and converted into vibration signal by the actuator (vibrator). The wave travels through the head bone to both sides of the cochleae.

2.2.3. Vibration transmission in the skull

How does the wave travel in the skull? How is the vibration mode of the skull under stimulating? For bone anchored hearing aid application, people concern more about the problems such like: Where is best position of the vibration actuator? How will the wave transmit to the cochlea? What is the intracranial transfer function? Those questions have

Figure 2.6.: Bone Anchored Hearing Aid (BAHA)[8]

been investigated for decades by doing the experiments using dry skulls, cadavers and the live human subjects [9, 10, 11].

Stenfelt has made the research of the vibration characteristics of the skull with a dry skull[10] and six cadavers [9]. From his research, it is assured that the closer the stimulation to the cochlea, the ipsilateral (near side) cochlea will get the larger signal and the larger attenuation to the contralateral (far side) cochlea. Therefore the most proper position for BC stimulating is at the temples.

Under stimulation, the skull moves as a rigid body at low frequencies so that both cochleae get the same level signal. It normally happens under 500 Hz. Beyond that until around 1 kHz, the “antiresonance phenomena” were found both in dry skull and cadavers experiments. The “antiresonance” is defined that under stimulating, the signals reach contralateral cochlea are larger than those reach ipsilaterl one.

At higher frequencies, the attenuation will be between -10 to 20 dB, it changes with different frequency.

The wave motion of the cranial bone under stimulation was found complex. According to Stenfelt, it is combination of longitudinal and transverse wav in the head when stimulated. Because of the complex structure of the cranial bone the head, the wave transmission is therefore quite complex, for example transmitted with different phase and group velocity at different frequencies.

2.2.4. Unilateral fitting of BAHA and bilateral fitting of BAHAs

More than15, 000 patients have been fitted the BAHAs worldwide without serious infection problems in clinic. Most of them fit BAHA unilaterally [1]. As to at which side of head to fit depends on the patients’ preference if both of the cochleae are working well, otherwise it is fit at the impaired cochlea side. The benefit of fitting bilateral BAHAs has been carried out by the Bosman [3] and Claudia [1]. By doing the hearing tests, the results indicates that bilateral fitting BAHAs can improve the spacial hearing ability, improved sound localization, and improve speech recognition. From the engineering aspect of view, Stenfelt [2] has also proved the benefit of bilateral fitting BAHAs, better hearing thresholds from the front and better overall hearing ability from the surround, but he also indicates "due to the cross-hearing of bone conducted sound, the binaural processing for the patient fitted bilaterally with BAHAs is less than for normal binaural air conduction hearing", the cross-hearing, in the report is defined as "crosstalk". The crosstalk signal transfers to the contralateral cochlea because of the antiresonance or low attenuation at low frequency range.

3. Crosstalk and Crosstalk Cancellation

3.1. Crosstalk

Crosstalk occurs when reproducing binaural recorded signals with two loudspeakers. One only expects to receive the signal from the right speaker at the right ear, and same for the left one. Unfortunately there are always the signal from the opposite speakers through the transfer path C_{LR} and C_{RL}; see Figure 3.1.Where C_{LR} represent the signal from the right side speaker to the left side of ear. Those signals are crosstalk signals.

In the application of two BAHAs fitted on the head, when two BAHAs are working, the signal reaches the cochlea is the combination of the waves from ipsilateral (near side) and contralateral (far side) BAHA, as shown in Figure 3.2. The signal from BAHA to ipsilateral cochlea is defined as direct signal, and the one to the contralateral cochlea is the crosstalk signal. The crosstalk signal is alternatively called transcranial signal in the thesis as introduced in chapter 2.

3.2. Crosstalk cancellation Algorithms

To cancel the crosstalk signals, the crosstalk cancellation is required. This technology was first introduced by Bauer in 1961, and put into practice by Schroeder and Atal in 1963 [12]. There are four methods that have been tried in this project: the ideal crosstalk cancellation, the adaptive crosstalk cancellation method, the recursive crosstalk cancellation and the fast deconvoultion method. Each algorithm has different drawbacks, advantages and limitations when applied to reproduce the stereo signal with two speakers. One has never applied any algorithm to cranial sound reproduction before, so that which of the algorithms is proper still an open problem. In this section, the algorithms will be introduced, and the analysis of its application in the skull can be found in chapter six.

3.2.1. Ideal crosstalk cancellation

In order to explain how the basic crosstalk cancellation work, the transmission of the signal in Figure 3.1 can be expressed by a mathematical matrix form,

$$\begin{bmatrix} Z_R \\ Z_L \end{bmatrix} = \begin{bmatrix} C_{RR} & C_{RL} \\ C_{LR} & C_{LL} \end{bmatrix} \begin{bmatrix} X_R \\ X_L \end{bmatrix} \tag{3.1}$$

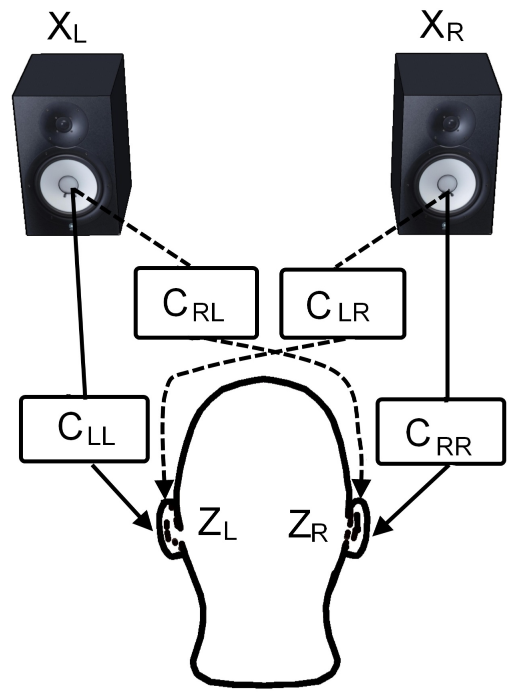


Figure 3.1.: Acoustic transfer funtions between two loudspeakers and th ears of a listener

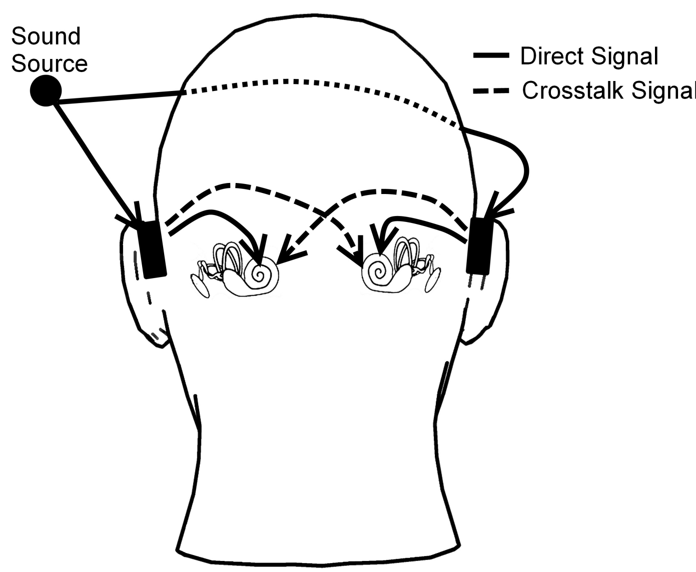


Figure 3.2.: Intercranial direct and crosstalk signal

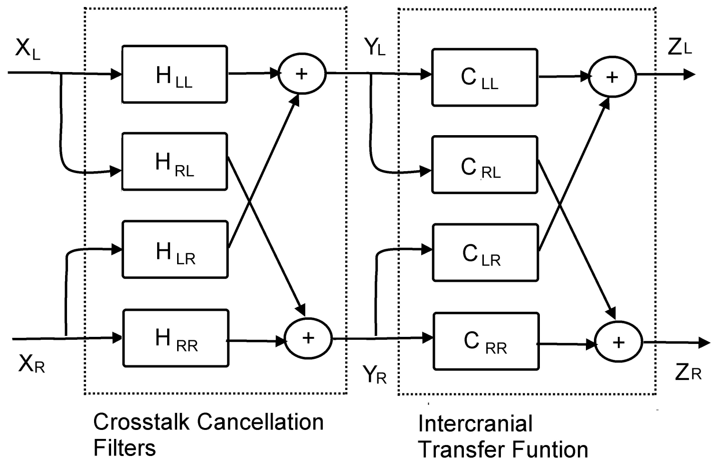


Figure 3.3.: Block diagram of the crosstalk cancellation algorithm

Where C matrix is the transfer function, C_{LL},$C_{RR}$are the direct paths, C_{RL},C_{LR} are the crosstalk paths, Z matrix is the signal at the ear and X matrix represents the signal emitted from the speakers. To cancel the crosstalk signals, one matrix H should be found. By multiplying H with the transfer function matrix C, one can get a desired identity matrix, shown in Equation (3.2). This matrix H is called the crosstalk cancellation matrix.

$$\begin{bmatrix} C_{RR} & C_{RL} \\ C_{LR} & C_{LL} \end{bmatrix} \begin{bmatrix} H_{RR} & H_{RL} \\ H_{LR} & H_{LL} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{3.2}$$

Equation (3.1) can be rewritten as Equation (3.3), and the system can be plot in block diagram as shown in Figure 3.3.

$$\begin{bmatrix} Z_R \\ Z_L \end{bmatrix} = \begin{bmatrix} C_{RR} & C_{RL} \\ C_{LR} & C_{LL} \end{bmatrix} \begin{bmatrix} H_{RR} & H_{RL} \\ H_{LR} & H_{LL} \end{bmatrix} \begin{bmatrix} X_R \\ X_L \end{bmatrix} \tag{3.3}$$

C matrix multiply H matrix becomes an identity matrix; the H matrix is, in another word, the inversion of the C matrix, therefore matrix H can be calculated using the formula below:

$$H = \frac{1}{C_{RR}C_{LL} - C_{RL}C_{LR}} \begin{bmatrix} C_{LL} & -C_{RL} \\ -C_{LR} & C_{RR} \end{bmatrix} \tag{3.4}$$

In the application of reproducing recording signals with two loudspeakers, at low frequencies, the product of multiplication of the direct paths, $C_{LL} \cdot C_{RR}$, can be identical to

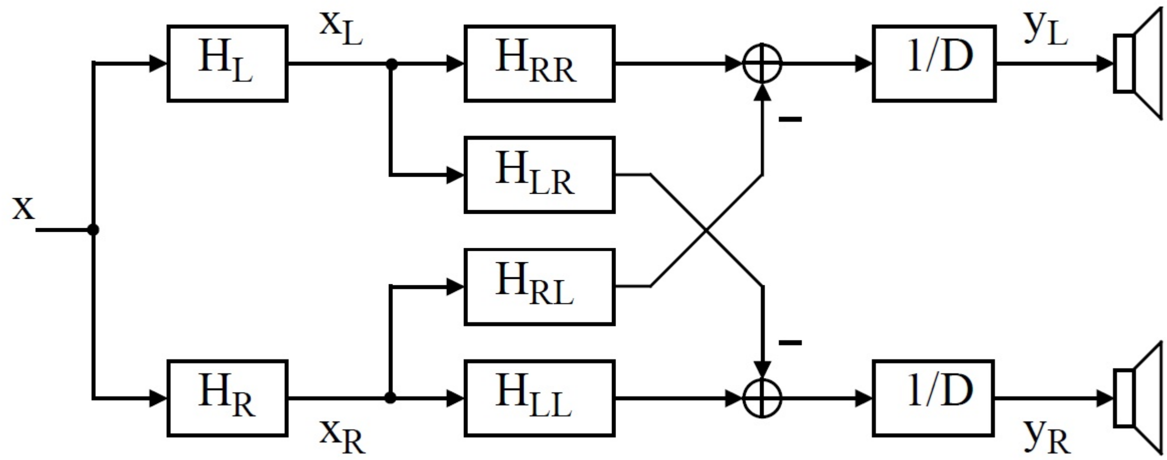


Figure 3.4.: Signal source binaural synthesizer cascaded with crosstalk cancellation filter. [res.: William G. Gardner 1998]

the product of multiplication of the crosstalk paths, $C_{RL} \cdot C_{LR}$, it makes the denomination of the Equation (3.4) very small, therefore the value of H matrix is boosted; this is called "ill-condition". One should prevent ill condition in real application, because of the dynamic limitation of the transducers or the amplifier.

Except the ill condition problem, non-minimum phase properties of the transfer function make the inversion of the transfer impulse response a problem in designing of the crosstalk cancellation filter. Several researches have published some crosstalk cancellation methods to prevent the problems above, but each of them have different limitations when it comes to the phase.

3.2.2. Recursive crosstalk cancellation algorithm and the ME-LMS method

Recursive crosstalk cancellation

Ideal crosstalk introduced above is just a concept, one should develop the algorithm that is applicable. The first usage of crosstalk cancellation carried out by Schroeder and Atal (1963) [12], the schematic figure is shown in Figure 3.4. The crosstalk cancellation filter is implemented using four feedforward filters and two inverse determinant filter, where $D = H_{LL}H_{RR} - H_{LR}H_{RL}$.

The straightforward way to apply the Equation (3.4) is using four feedforward filters, as introduced above. Two recursive filter topologies which can also implement the inverse matrix are shown in Figure 3.5. The left one has been used by Iwahara and Mori [12] to implement crosstalk cancellers.

The system equations for the topology in Figure 3.5a are:

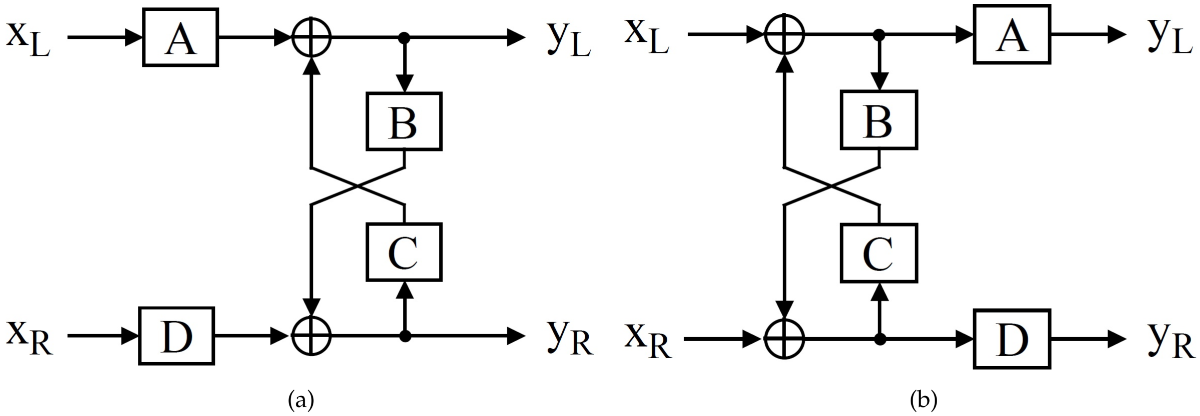


Figure 3.5.: Recursive toplogies for implementing the 2x2 inverse matrix

$$\begin{aligned} y_L &= \mathrm{A}x_L + \mathrm{C}y_R \\ y_R &= Dx_R + By_L \end{aligned} \tag{3.5}$$

The coefficients in Equation (3.5) are:

$$\begin{aligned} \mathrm{A} &= \frac{1}{C_{LL}} \\ B &= -\left(\frac{C_{LR}}{C_{RR}}\right) \\ C &= -\left(\frac{C_{RL}}{C_{LL}}\right) \\ D &= \frac{1}{C_{RR}} \end{aligned} \tag{3.6}$$

The mathematical expression of the system shown in Figure 3.5b are in Equation (3.7)

$$\begin{aligned} y_L &= \mathrm{A}\left(x_L + \frac{C}{D}y_R\right) \\ y_R &= D\left(x_R + \frac{B}{A}y_L\right) \end{aligned} \tag{3.7}$$

The coefficients in Equation (3.7) are given as below:

$$\mathrm{A} = \frac{1}{C_{LL}}$$

$$
\begin{aligned}
B &= -\left(\frac{C_{LR}}{C_{LL}}\right) \\
C &= -\left(\frac{C_{RL}}{C_{RR}}\right) \\
D &= \frac{1}{C_{RR}}
\end{aligned} \tag{3.8}
$$

The important feature for these two algorithms is that the predicted signal is fed back to the opposite channel to cancel the crosstalk signals. Therefore the value of B and C and proper delay in Equation (3.6) and (3.8) are very critical in implementation.

Multi-error least mean square method

The multi-error least mean square (ME-LMS) method has investigated by Nelson et al [13]. It is an adaptive filter that can manage the non-minimum problem with combination the usage of receiving and feeding back the signal to update the filter. In the 3D reproduction with two loudspeakers, two microphones are placed in the ear canals of the dummy head. The whole system can be plot in a block diagram in Figure 3.6. The block was circled by dash line, it is same as the one in Figure 3.3, but the Z_L and Z_R in Figure 3.3 are optimal results that are identical to the input signal to the speaker. In real application, the real signals one obtained at the microphones, $\hat{Z}_L$ and $\hat{Z}_R$, are usually deviated from the desired signals that one would get. The desired signals Z_R and Z_L are should be the same as x_R and x_L, but with some reasonable time delay Δ. The time delay is introduced because of the propagation time to the receivers on the dummy head. The deviation between the desired signals and the real measured signals are defined as error signals e_L and e_R in the block diagram. One uses the error signal as a feedback to update and calculate a better filter h in ME-LMS method. The detail can be found in appendix A.

3.2.3. Fast deconvolution algorithm

Deconvolution is a quite common terminology that refers to a calculation method that one can use to obtain the input data from the output data, more specifically; it is an "inversion" in signal processing area. This method is a very useful tool in many areas. In audio applications, it is called "equalization" for one channel deconvolution or used for the multi-channel sound reproduction.

Fast deconvolution method designed in frequency domain

Fast deconvolution is a method based on Fast Fourier Transform (FFT) in combination with the least square sense [14]. Two regularization parameters, gain factor and shape function are crucial for designing the crosstalk cancellation filter coefficients [15]. An example that demonstrates how the fast deconvolution method works is shown in Figure

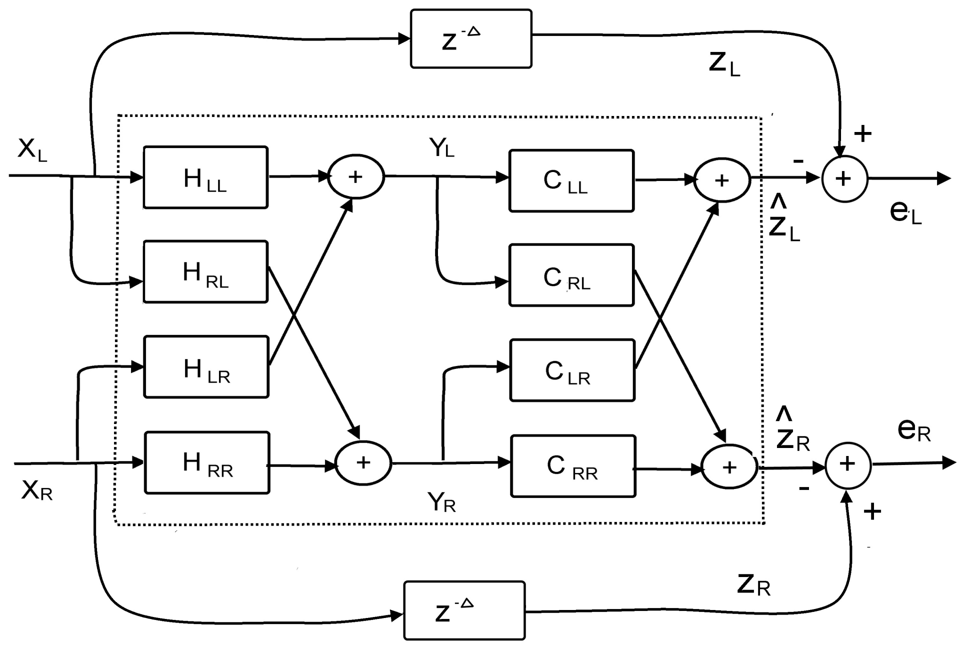


Figure 3.6.: System scheme of ME-LMS

3.7 and Figure 3.8 shows the block diagram of the model. Since the system is assumed to be working in discrete time, the conventional z-transform notation is used, where the $v(z)$ is the loudspeaker input signal, $w(z)$ is output signal from the microphone, $C(z)$ is the electro-acoustical transfer function, $e(z)$ is the reproduction error vector, $u(z)$ is the recorded signal, $H(z)$ represents the crosstalk cancellation filter coefficient, $d(z)$ is the desired signals and the $A(z)$ is the target function. The signal $u(z)$ is delayed m samples. When there are two loudspeakers and two microphones in the sound field, the transfer function matrix C can be written in 2x2 matrix as mentioned before.

The relationships among the variables are denoted as bellow:

$$v\left(z\right)=H_{m,A}\left(z\right)u\left(z\right) \tag{3.9}$$

$$w\left(z\right)=C\left(z\right)v\left(z\right) \tag{3.10}$$

$$d\left(z\right)=z^{-m}A\left(z\right)u\left(z\right) \tag{3.11}$$

$$e\left(z\right)=d\left(z\right)-w\left(z\right) \tag{3.12}$$

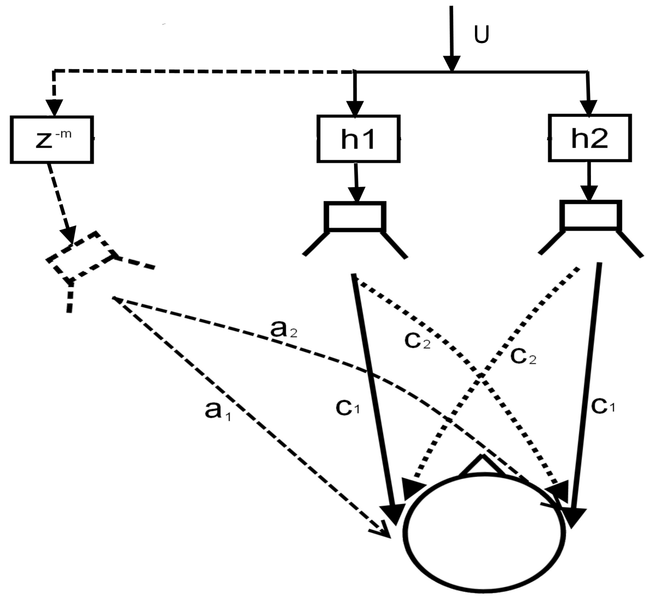


Figure 3.7.: An example that demostrates the fast deconvolution method

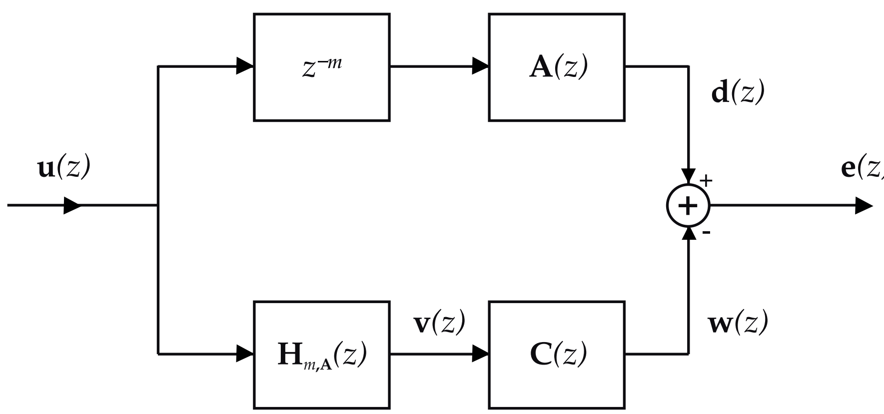


Figure 3.8.: Block diagram of the fast deconvolution method example

The cost function J has to be introduced first to explain how the least square sense is applied in the fast deconvolution. If only the gain factor β is considered, the cost function can be defined as the Equation (3.13) . In the equation, the delay m is assumed to be zero and the z is replaced by $e^{j\omega\Delta}$, where ω is the angular frequency and Δ is the sampling interval.

$$J\left(e^{j\omega\Delta}\right) = e^{H}\left(e^{j\omega\Delta}\right) e\left(e^{j\omega\Delta}\right) + \beta v^{H}\left(e^{j\omega\Delta}\right) v\left(e^{j\omega\Delta}\right) \tag{3.13}$$

The first part $e^{H}e$ is the "performance" term, and $\beta v^{H} v$ is the "effort penalty" term. β

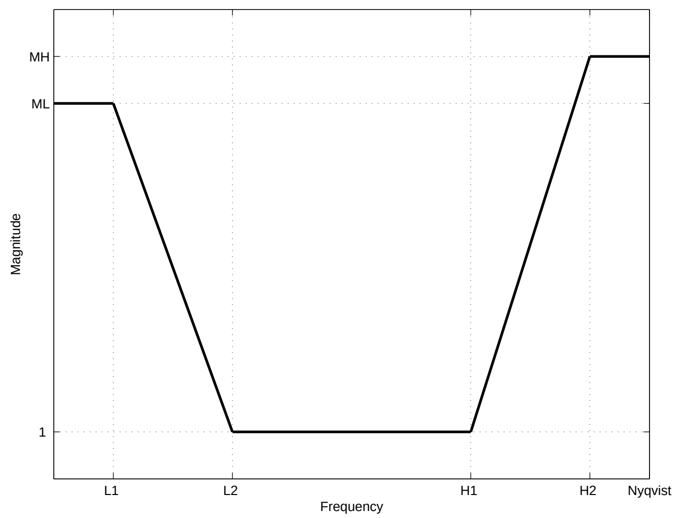


Figure 3.9.: A suggested shape function for audio-related problems

is a positive real number, a regularization parameter that determines how much weight to assign to the effort term. By varying β from zero to infinity, the solution changes from minimizing only the performance error to minimizing only the effort penalty. From Equation (3.9) to Equation (3.13), in order to minimize J in the least square sense, the $H_{o,A}$ must be given as:

$$H_{o,A}\left(e^{j\omega\Delta}\right) = \left[C^H\left(e^{j\omega\Delta}\right)C\left(e^{j\omega\Delta}\right) + \beta I\right]^{-1} C^H\left(e^{j\omega\Delta}\right)A\left(e^{j\omega\Delta}\right) \tag{3.14}$$

When the desire signal $d(z)$ is identical to the input signal $u(z)$, the matrix $A(z)$ can be replaced by an identity matrix. Equation (3.14) can then be rewritten as:

$$H_{o,I}\left(e^{j\omega\Delta}\right) = \left[C^T\left(z^{-1}\right)C\left(z\right) + \beta I\right]^{-1} C^T\left(z^{-1}\right) \tag{3.15}$$

If the shape function $B(z)$ is considered, the Equation (3.15) will be rewritten as Equation (3.16). $B(z)$ is a z-transform of a digital filter that amplifies the frequencies that should not be boosted by the crosstalk cancellation network.

$B(z)$ is a frequency dependent regularisation, it suppresses the signal value boosted due to ill-condition. In conventional crosstalk cancellation application, the shape function is suggested as shown in Figure 3.9[16]. Where B_L is the low frequency asymptotic, B_H is a high frequency asymptotic value. In the middle frequency range, $|B|$ is one. B_L and B_H are usually much greater than one. The frequencies f_{L1}, f_{L2}, f_{H1} and f_{H2} defines the two transition bands.

$$H\left(z\right) = \left[C^T\left(z^{-1}\right)C\left(z\right) + \beta B^T\left(z^{-1}\right)B\left(z\right)\right]^{-1} C^T\left(z^{-1}\right)z^{-m} \tag{3.16}$$

Both the shape function and the gain factor are found by trial and error; one can not calculate the optimal values of them. Moreover, there are definitely some advantages of using this method, but there is a big restrain that the filter derived from this method must be a long filter, at least 4 times length of the measured transfer function [14]. An improved approach can be made from the time domain aspect, and will be introduced next.

Fast deconvolution method designed in time domain

For a single channel deconvoultion, assume that the length of the crosstalk cancellation filter h is N_h, and that the length of the measured transfer impulse response c is N_c. A matrix C is defined as in Equation (3.17) below, so that $c * h = C \centerdot h$, where $*$ is the convolution operator. Therefore C is a matrix of which the number of column is N_h, and the number of rows is $N_h + N_c - 1$.

$$C = \begin{bmatrix} c(0) & & \\ \vdots & \ddots & \\ c(N_c - 1) & \ddots & c(0) \\ & \ddots & \vdots \\ & & c(N_c - 1) \end{bmatrix} \tag{3.17}$$

The performance term and effort penalty term are therefore changed into $E = e^T e$ and $V = \mathbf{v}_b^T \mathbf{v}_b$. Vector $\mathbf{v}_b$ is the input voltage to the loudspeaker which is the signal after original signal, v, is filtered by the shape function b. In this mathematical expression, $\mathbf{v}_b$ equals to b convolutes v, $\mathbf{v}_b = b * v$. Same as C matrix, there is another matrix that B can be found, $b * \mathbf{v} = B \centerdot \mathbf{v}$, B is a matrix of which the number of column is N_h and the number of rows is $N_h + N_b - 1$

$$B = \begin{bmatrix} b(0) & & \\ \vdots & \ddots & \\ b(N_b - 1) & \ddots & \\ & \ddots & b(0) \\ & & \vdots \\ & & b(N_b - 1) \end{bmatrix} \tag{3.18}$$

Substitute these variables into the cost function J, the equation for calculating filter h is therefore derived as below.

$$h = \left[C^T C + B^T B \right]^{-1} \centerdot C^T a_m \tag{3.19}$$

Where $B^T B$ and $C^T C$ are both N_h-by N_h matrix, a_m is the target transfer function, $N_h + N_c - 1$-by N_h matrix. If the filter length of target transfer impulse response with m samples delay is shorter than $N_h + N_c - 1$, zero padding can be used to extend the length.

The formula of multi-channel deconvolution can be easily extended from the single channel formula. Assume the multi-channel system has S speakers and R microphones. The filter h matrix is therefore defined as

$$h = \begin{bmatrix} h_1 \\ \vdots \\ h_S \end{bmatrix} \tag{3.20}$$

For a, d, w, and e, they are all column vectors made up of R, rather than S, FIR filters. The a_m becomes :

$$a_m = \begin{bmatrix} a_{m,1} \\ \vdots \\ a_{m,R} \end{bmatrix} \tag{3.21}$$

The effort V for multi-channels application is then defined as:

$$\mathbf{v}_b = \begin{bmatrix} \mathbf{v}_{b,1} \\ \vdots \\ \mathbf{v}_{b,S} \end{bmatrix} \tag{3.22}$$

And total energy V from s speakers are calculated as:

$$\mathbf{v}_b^T \mathbf{v}_b = \mathbf{v}^T \left(\oplus \sum B^T B \right) \mathbf{v} \tag{3.23}$$

Where

$$\oplus \sum B^T B = \begin{bmatrix} B^T B & 0 & \cdots & 0 \\ 0 & B^T B & \vdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B^T B \end{bmatrix} \tag{3.24}$$

The optimal crosstalk cancellation filter h is derived as in Equation (3.25)

$$h = \left[C^T C + \beta \oplus \sum B^T B \right]^{-1} \cdot C^T a_m \tag{3.25}$$

4. Measurement

In order to investigate the crosstalk cancellation effect of the BAHAs in the skull, two major data are compulsory: the head related transfer function (HRTF) of the BAHA and the intercranial transfer function. These will be introduced next.

4.1. The Intercranial transfer function measurement setup

A dry human skull is used to investigate the intercranial transfer function in this project. Figure 4.1 shows the setup scheme of measuring the transfer function of the skull; the equipments are list in Table 4.1. The measurements were carried out in the vibro-acoustic laboratory in Applied Acoustics Department in Chalmers. The skull was originally cut into parts, the upper part was like a lid so that one can open it and place the accelerometer inside; a thick porous adhesive is in between the two parts to create a continuing connection and a hook holds the two parts tightly during the measurement; Figure 4.2 shows the details. The skull was hung in the spring to be simulated as a free body. Two force transducers were placed at each side of the skull. The measurement setup is shown in Figure 4.3.

If one can't measure the vibration exactly inside the cochlea, the alternative way is to measure the signal at the position very close to the cochlea. In Stenfelt's research, the accelerometers were placed at the accurate eminence (top portion of the petrous part of the temporal bone) [10]. Same for this project, two B&K 4374 accelerometers are mounted to the accurate eminences of the skull as shown in the Figure 4.4. Two B&K type 8203 force transducers are used to generate the vibration signal to the skull; they are mounted to the two sides of mastoid processing part of the skull. The frequency range which is interested in this project is quite wide, from 200 Hz up to 10 kHz, therefore a rigid mounting is

Equipment	Model
Force Transducer	Brüel & Kjærtype 8203
Acclerometer	Brüel & Kjærtype 4374
Acquisitions	MLSSA (Maximum length sequence system analyzer)
Charge Amplifier	Brüel & Kjær 2635
Amplifier	NAD 3020

Table 4.1.: Intercranial transfer function measurement equipment list

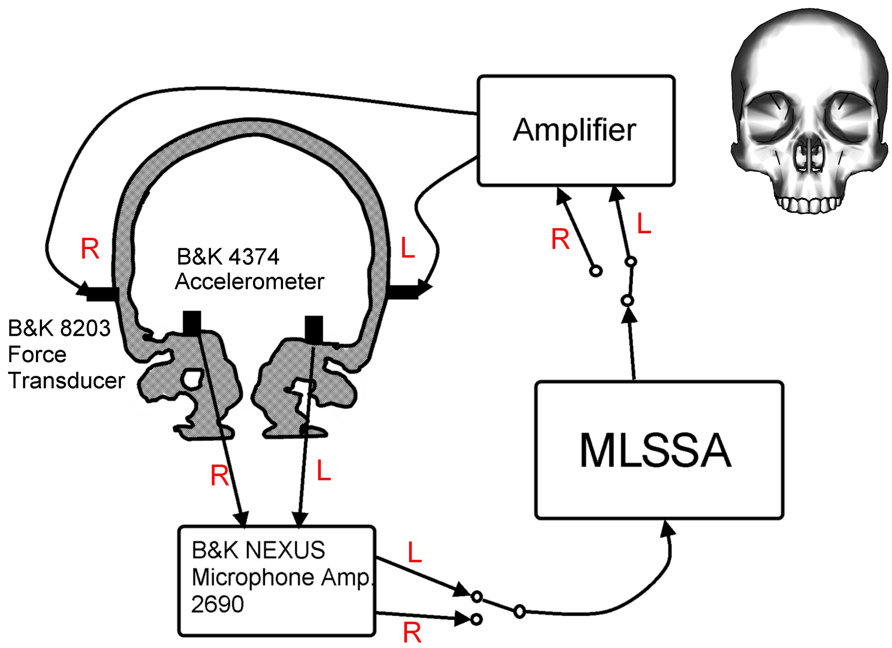
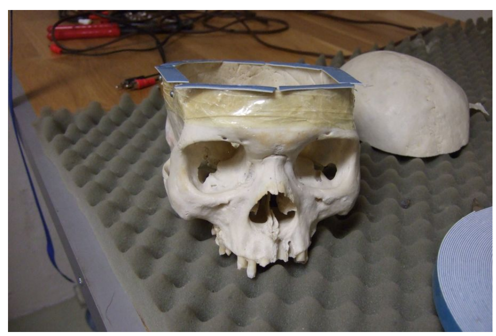


Figure 4.1.: Intercranial transfer function measurement setup scheme (with MLSSA)

(a)



(b)

Figure 4.2.: Dry skull

required. Especially at higher frequencies, the rigid mounting will not cause the distortion or low coherence problem. The optimal mounting method is using X-60 glue, but the drawback is that the glue can't be removed after the measurement [17]. Since this skull is not allowed to be treated such destructive mounting method, the alternative was to use a so called "bee's wax", a yellow wax from B&K. The corresponding drawback is that one should be careful to take care of the high frequency distortion and low coherence results.

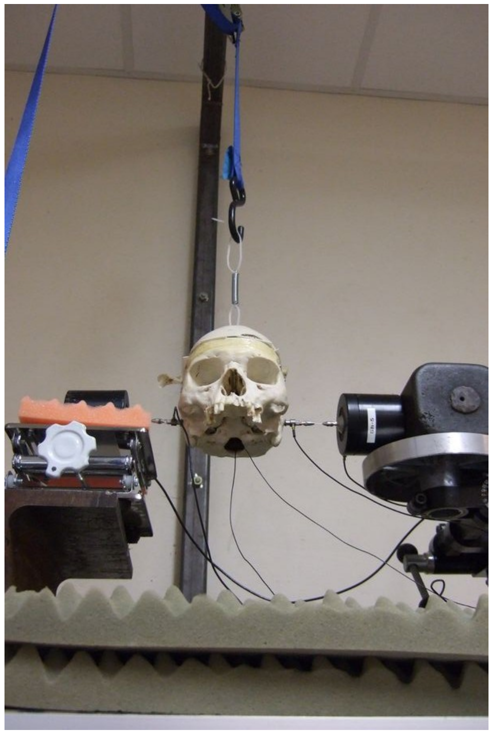
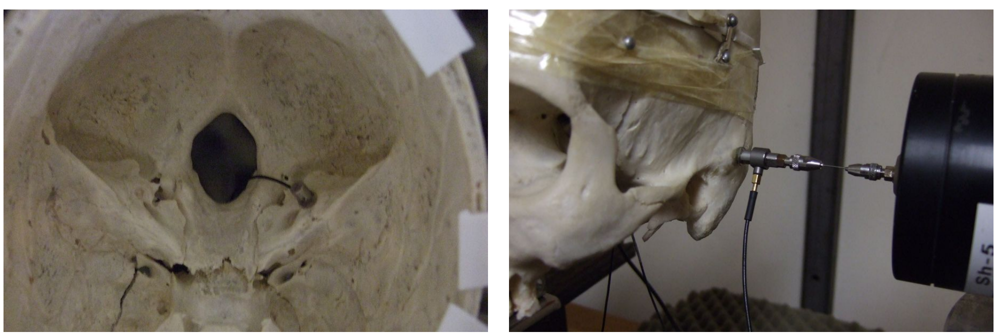


Figure 4.3.: Intercranial transfer function measurement setup



(a) accelerometer at the accurate eminence of the skull

(b) force transducer on the mastoid processing of the skull

Figure 4.4.: the accelerometer and the force transducer

In order to know if the force transducers and accelerometers are well mounted, the TXI acquisition station combination with the Trigger Happy analyzing software was used. The results will be discussed in chapter 6.

Since fast deconvolution in time domain is used for designing the filter, one needs the intercranial transfer impulse responses of the skull to implement the digital signal processing. To use the MLSSA as the acquisition and analyzing system is the proper solution.

The sampling rate of MLSSA was set to 32 kHz. Since the highest frequency we are interested in is 10 kHz, the ratio of sampling rate to 10000 is higher than the Nyquist factor to make sure there will be no aliasing problem. The setting are: MLS continues stimulus signal, Chebyshev filter, 10 times of measurement average for each impulse response/ transfer function.

The intercranial transfer functions/impulse responses were measured separately. For example, $FLAL$ is an intercranial transfer function measured when the skull is stimulated on left side by the force transducer and the vibration level is measured at the left cochle. With the same stimulus, while the vibration level is measured at the right cochlea, the intercranial transfer function $FLAR$ is obtained. Same definition for $FRAR$ and $FRAL$.

4.2. The Intercranial transfer function with crosstalk cancellation effect measurement

The intercranial transfer functions were obtained from previous measurement. The next step is to measure the crosstalk cancellation effect in this application. An extra device is needed in this case, the digital processor, to do the digital signal processing.

The DSP board is Texas Instrument TMS320C6713 which is featured with 32 bits floating point digital signal processor and a stereo input/output audio CODEC, AIC23. The algorithm is written in C program language, with the Code Compose Studio (CCS), one can do the programming in Windows. There is a USB cable connection between the DSP board and the PC, the algorithm can be built to the DSP board.

The main strategy of this structure is the usage of the ping-pong buffer and the circular buffer [18]. The algorithm block plot is in Figure 3.3, the crosstalk cancellation filters are h_{rr}, h_{rr}, h_{rl} and h_{lr}. They are designed by the method introduced in chapter 3, 256 taps of FIR filters.

After the DSP board was confirmed working and running well with the code, the measurement was carried out by the setup shown in Figure 4.5. When the DSP is working, the whole system becomes non-linear; MLSSA can't handle non linear system analyzing, therefore the recording method is introduced here. The equipment are the same as in Table 4.1, only the ananlyzer is changed. A personal computer with an external sound card, Creative E-MU type 0404 is used as the acquisition and the signal source. The test signal was white noise. The DSP board can be set to do the signal processing or bypass the signal directly from the CODEC analog input to CODEC analog output by writing the code and then simply switch the switch on the board. Therefore, one can measure the intercranial transfer function with and without crosstalk cancellation effect with this setup.

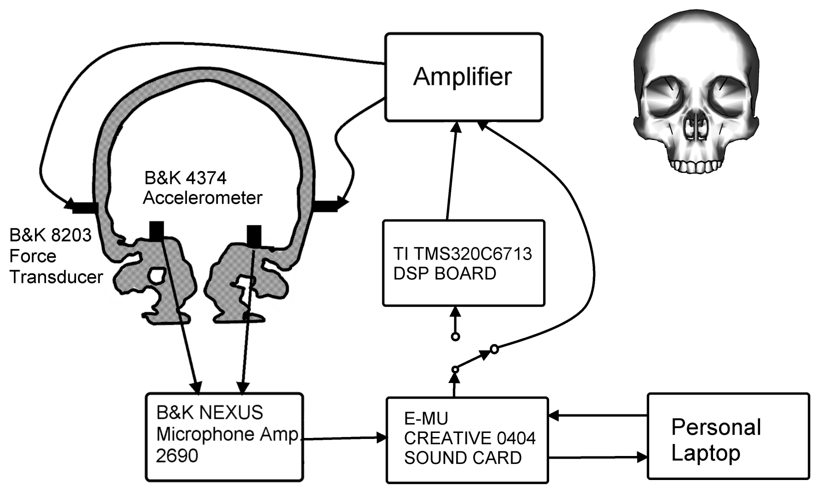


Figure 4.5.: Setup scheme of the intercranial transfer function measurement by recording the signals

4.3. Head Related Transfer Function (HRTF) measurement

The pinna of the ear, the head and the ear canal combines into a very complex filter thru which people perceive sound. To know the properties, one can measure head related transfer function (HRTF) to see the characteristics. HRTF is defined as the relative spectrum level difference of the sound source in the space to sound signal level perceived by the microphone which is placed in the ear cannel. When the patient wears the BAHAs, one is interested in the HRTF of the BAHA. In this section, to method used to measure the HRTF of the BAHA will be introduced.

This measurement was carried out in the anechoic chamber in Applied Acoustics Department of Chalmers. The experiment equipment list is in Table 4.2. A human torso and head simulator, KEMAR, was placed in the middle of the anechoic chamber on the turning table; the BAHA was placed 5 cm behind the ear pinna that is introduced in chapter 2. The sound source was played through a loudspeaker which was placed 3 meters from KEMAR [19], and as the same height as the ear of the simulator; the scheme is shown in Figure 4.6. Figure 4.7 shows the setup of the measurement that how the equipment are connected. The analyzer is MLSSA, same as what was used for the intercranial transfer function/ impulse response measurement. The test signal sent by MLSSA is continuous MLS (maximum least sequences) signal; the sampling rate was set 75 kHz. The measured data was saved in time and can then be called head related impulse response (HRIR). that can be transferred to HRTF via doing FFT (Fast Fourier Transform) later in the laptop with

Equipment	Model
Analyzer	MLSSA (Maximum length sequence system analyzer)
Microphone	Brüel & Kjær type 4134
Microphone Amplifier	Brüel & Kjær type 2804
Calibrator	Brüel & Kjær type 4231
Loudspeaker	YAMAHA NS-10M (R)
Amplifier	NAD 3020
Turn table	Outline ET1/ST1
Torso and head simulator	Knowles Electronics KEMAR
BAHA	Cofidential

Table 4.2.: HRTF measurement equipment list

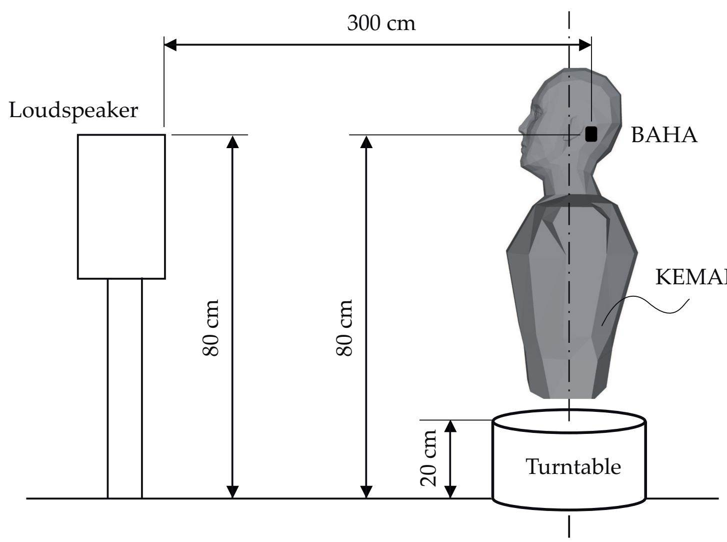


Figure 4.6.: HRTF measurement setup (details of the distances)

MATLAB to achieve higher frequency resolution then MLSSA can manage. The minimum frequency increment of MLSSA is 157 Hz due to the limit hardware ability, it is too large for doing the analyzing in the low frequency range.

The turning table was used for measuring the directivity characteristic of the BAHA. The angle interval of the measurement is 5°. To measure 360°, one will get 71 HRIRs. A MACRO file on MLSSA was used so that the measurement of these 71 steps was run automatically.

The measurement results was equalized later by a reference data measured by using B&K 4134 microphone. After calibrating the microphone with B&K 4231 calibrator, the microphone was placed in the same height and the distance as the KEMAR to the speaker as shown in the Figure 4.8; the test signal and test setting of the MLSSA was the same as

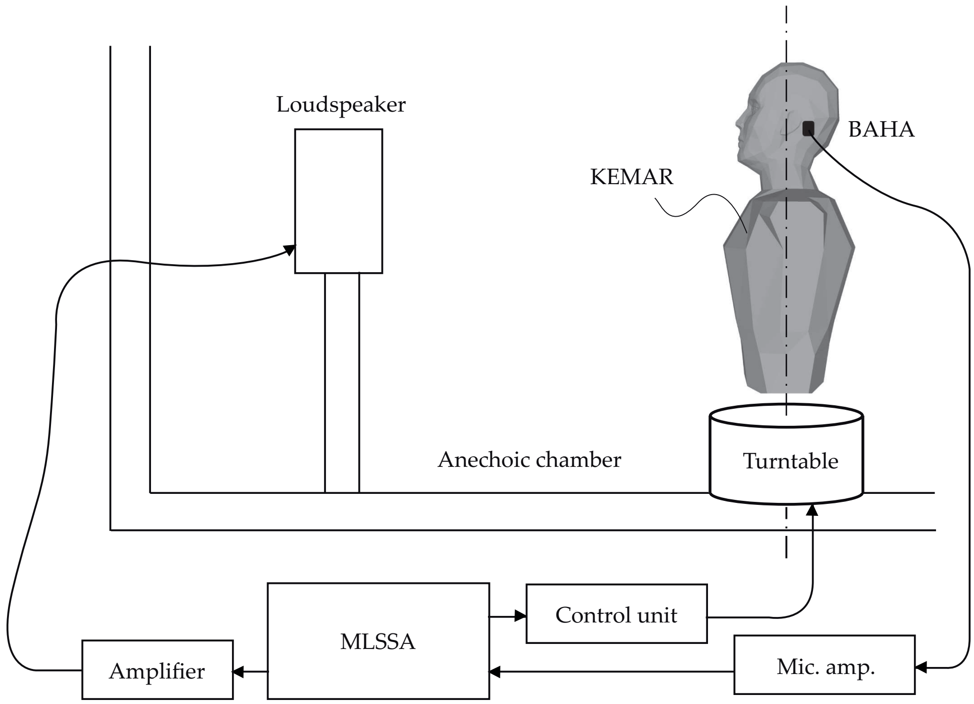


Figure 4.7.: HRTF measurement setup

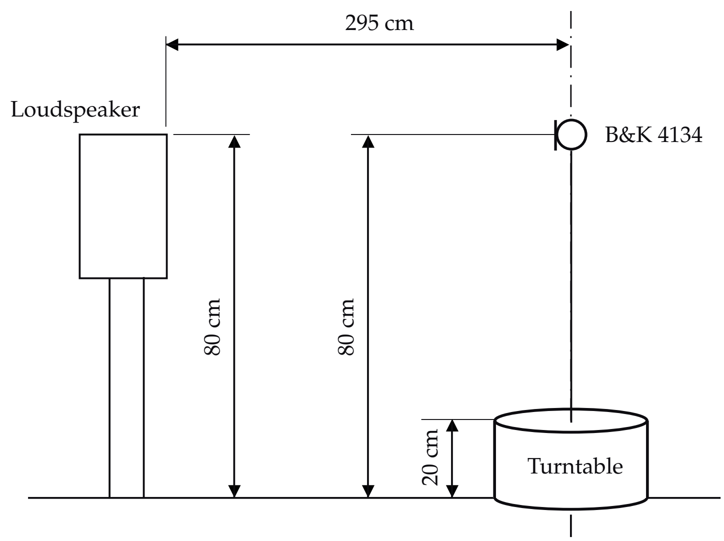


Figure 4.8.: HRTF measurement setup for reference data

for measuring BAHAs. The purpose of doing free field equalization is to get rid of the factors from the loudspeakers, the signal or noise, to have compensation of the response of the whole measurement system.

5. Data Processing

5.1. From the recorded time domain data to Frequency domain data processing

In the Figure 6.5, the sample rate of the sound card to acquire the signal was 48 kHz. The signal was started to be recorded since the test signal was played. The recording time was approximately 5 minutes.

To transfer the acquired data into frequency data, there are several steps of processing. The data processing of steps are listed below [20]:

1. Decide the block size, 8196 or 16372, the smaller size the more block numbers, but the recorded signals are very long, therefore the number of the blocks is not a problem concerned about.

2. Do the windowing to each block. The Hanning window was used in this project.

3. Do the FFT of the filtered time signals.

4. Do the average of all the frequency response of each block.

5. Transfer to 1/3 octave band.

5.2. The signals at two cochleae

The signals at cochleae in this project were calculated instead of measured. The signal at the right side of the cochlea, $Cochlea_R$, is calculated from Equation (5.1) as below [2]:

$$Cochlea_R = HRTF_{BAHA-R} \cdot TF_{IPSI} + HRTF_{BAHA-L} \cdot TF_{CONTRA} \tag{5.1}$$

Where TF_{IPSI} represents the intercrainal transfer function of the ipsilateral side path, in another word is the vibration wave transfers from the force transducer to the near side cochlea. From the measurement of the intercranial transfer function, there should be four ones, $FLAL$, $FLAR$, $FRAL$, $FRAR$. In order to simplify the calculation model, only the $FLAL$ and $FLAR$ are chosen for calculation. $FLAL$ represents TF_{IPSI}, and $FLAR$ represents TF_{CONTRA}. The signals reach the left side of cochlea can be calculated from Equation (5.2):

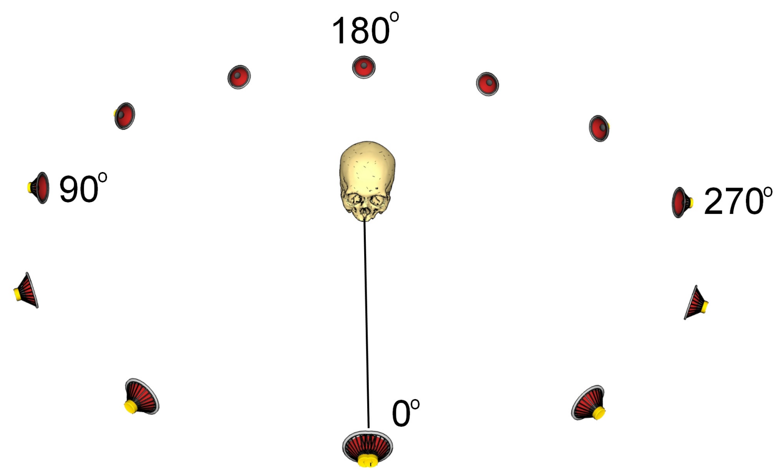


Figure 5.1.: Sound source from different azimuth angle

$$Cochlea_L = HRTF_{BAHA-L} \cdot TF_{IPSI} + HRTF_{BAHA-R} \cdot TF_{CONTRA} \tag{5.2}$$

In Equation (5.1) and (5.2), $HRTF_{BAHA-R}$ and $HRTF_{BAHA-L}$ are related such that: $HRTF_{BAHA-R}(\alpha) = HRTF_{BAHA-L}(-\alpha)$; α is the relative angle of the sound source to the middle of KEMAR, as Figure 5.1 shows. Minus angle means the anticlockwise direction and the positive angle means the clockwise direction.

5.3. ILD calculation

The interaural level difference is for normal hearing people was introduced in chapter two. In this project one is more interested to know the inter-cochlea level difference. The level difference can be also calculated from the Equation (5.3) [19] :

$$ILD = 20\log_{10}\left(\left|\frac{Cochlea_R}{Cochlea_L}\right|\right) [dB] \tag{5.3}$$

6. Results and Discussion

6.1. Intercranial Transfer function

The intercranial impulse response and intercranial transfer function are shown in Figure 6.1 and 6.2. In time domain, one can see the crosstalk signals, c_{RL} and c_{LR}, are delayed about 20 µs compared to the direct signal. The skull bone is also a quite damped material, therefore both signal die out after 8 ms. The intercranial frequency response in Figure 6.2 provides information which are more important and interesting. From a big scale point of view, one can see this skull is not symmetric, the direct and crosstalk transfer paths at two sides are not identical, but the trends are similar. Take the left side as an example, the level of the direct transfer function C_{RR} and transcranial transfer function C_{RL}, are almost same level below 500 Hz. This phenomenon fits what is described in chapter two, the skull moves as a rigid body. After 500 Hz, the level of crosstalk path is larger than the direct path until 1 kHz, the antiresonance is also shown here. Above the frequency, the level of transcranial path is attenuated, the higher frequencies, the more attenuation.

The usage of the TXI and trigger happy provides the information if the setup of the force transducers and accelerometers are decent. The transfer functions got by this measurement are compared with MLSSA ones in Figure 6.3. One can see the trends are the same. Figure 6.4 shows the coherence of this measurement. In the coherence plot, one can see below at least 6 kHz, the measurement setup is reliable. In case of the overtone situation that the coherence can't help to indicate, the THD (total harmonic distortion) is measured by the MLSSA, which is shown in Table 6.1. The measuring frequencies are prime numbers, it can simplify the analyzing situation due to the dimension or resonance. From the results, one can reassure the measurement setting is reliable.

The third method measuring of intercranial transfer function was used mainly when the crosstalk cancellation algorithm is applied through the DSP board. The intercranial transfer function without crosstalk cancellation can be checked first if this measurement

Frequency (Hz)	223	499	1193	1999	2999	4999
THD (%)	-20.4	-15	-32	-45.1	-37.2	-27.7
Odd (%)	-23.7	-16.9	-35.2	-49.7	-40.1	-35.7
Even (%)	-23.7	-19.5	-36.7	-46.9	-40.3	-28.5

Table 6.1.: Total harmonic distortion of the measurment setting

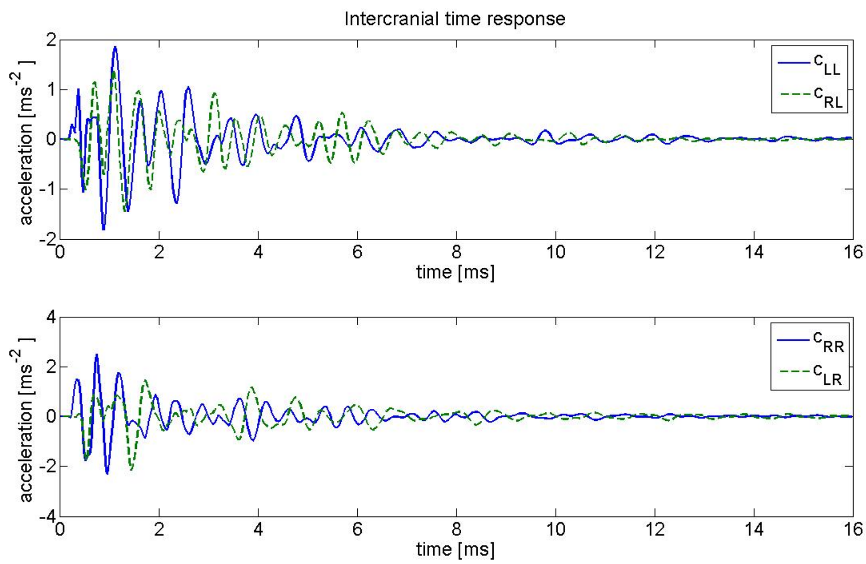


Figure 6.1.: Intercranial impulse response measured by MLSSA

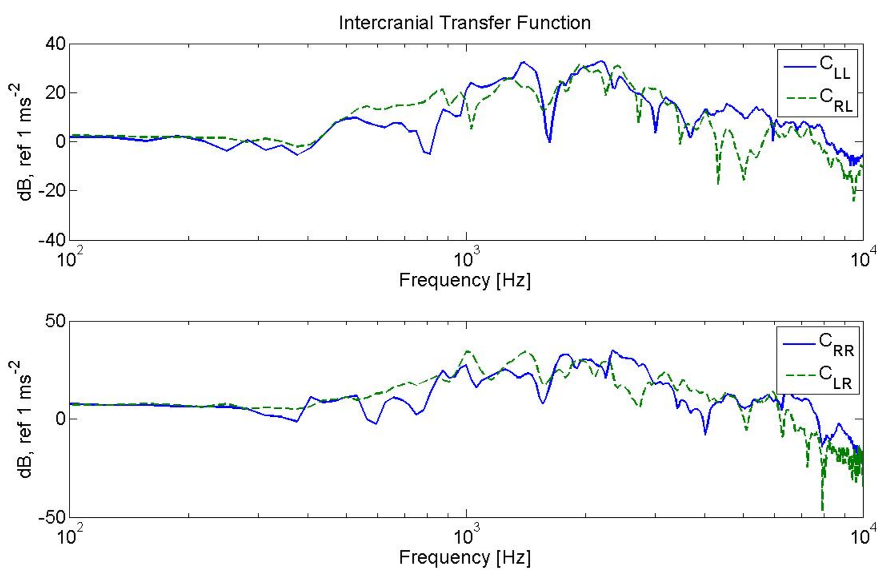


Figure 6.2.: Intercranial frequency response measured by MLSSA

method is feasible. The measured results are shown in Figure 6.5. Compared with the results from previous one in Figure 6.3, the trends of the curves are same, but the levels are different since it is very tricky to calibrate the signal that the laptop acquired in the third measurement method. One can therefore at least know these three methods are compatible, if one only compares the relative level differences of transfer functions of the

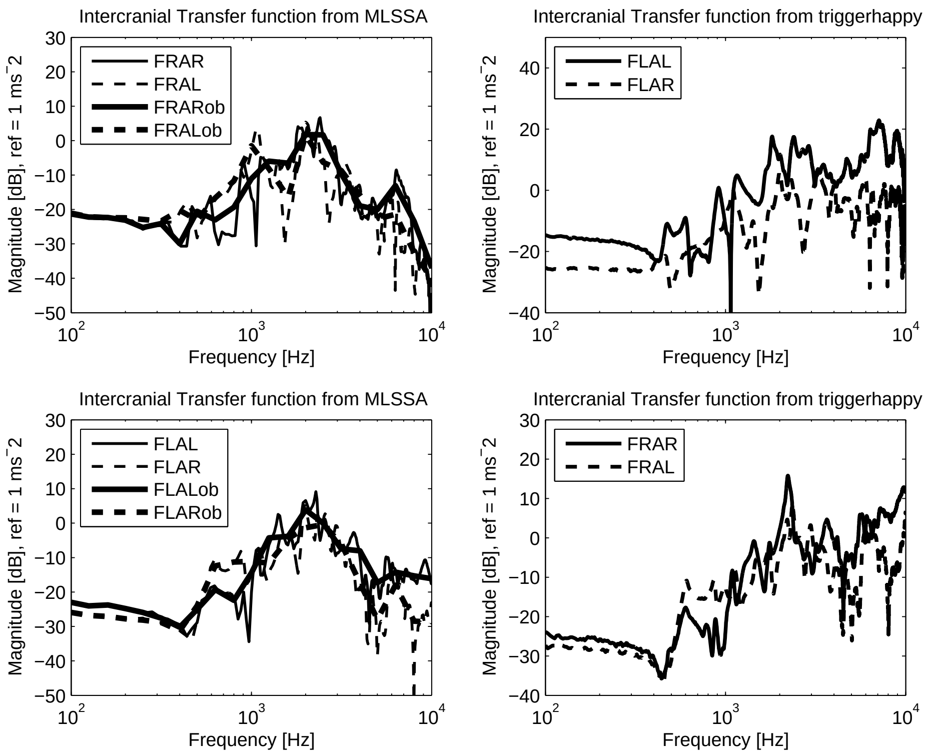


Figure 6.3.: Comparison of the Intercranial transfer function from TXI acquision and MLSSA

skull.

6.2. Crosstalk cancellation filter design

6.2.1. Crosstalk cancellation algorithms comparison

In order to study and to understand the algorithms, one example is used to investigate the crosstalk cancellation effects. This example is the HRTF (head related transfer function) of a KEMAR measured by the MIT media lab in the anechoic chamber, the source is incident from 30 degree of azimuth angle and 0 degree of elevation angle. The results are plotted in Figure 6.6; the x axis represents the frequency range and the y axis shows the attenuation level of the crosstalk signal refers to the direct signal. If there is no attenuation, the y axis value will be 0 dB. The solid line in Figure 6.6 is the natural attenuation caused by the head shadowing, without any crosstalk cancellation.

The others curves show different attenuation effect, in average, the attenuation is large enough to achieve the binaural listening, but the corresponding application limitations have to be considered before applying to the head skull. These algorithms are carried out

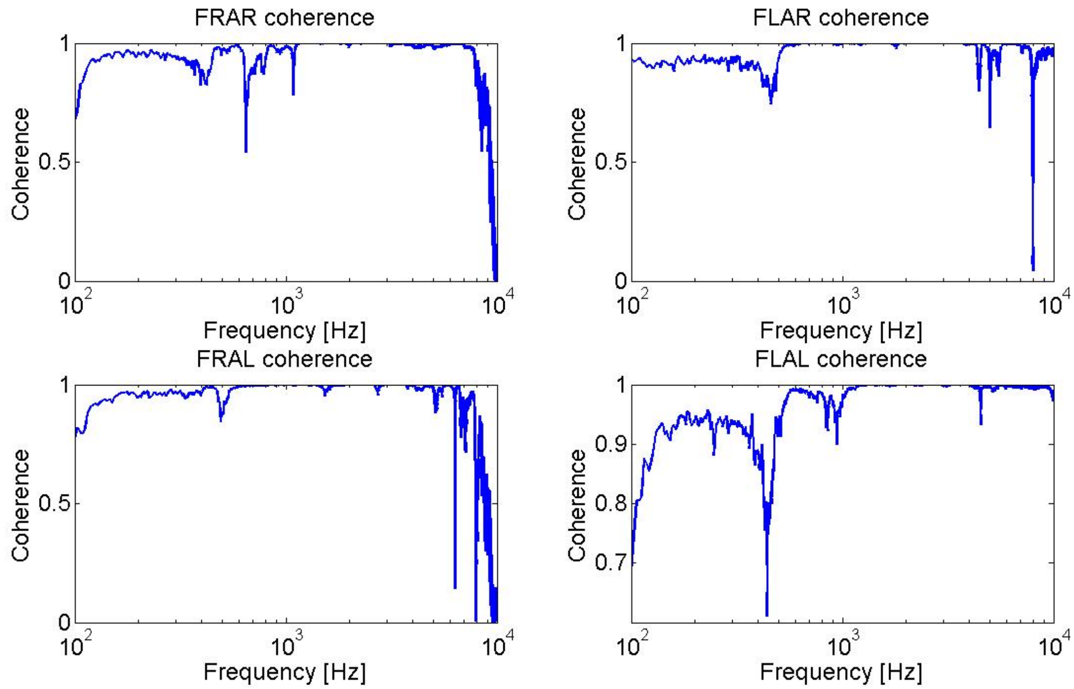


Figure 6.4.: Coherence

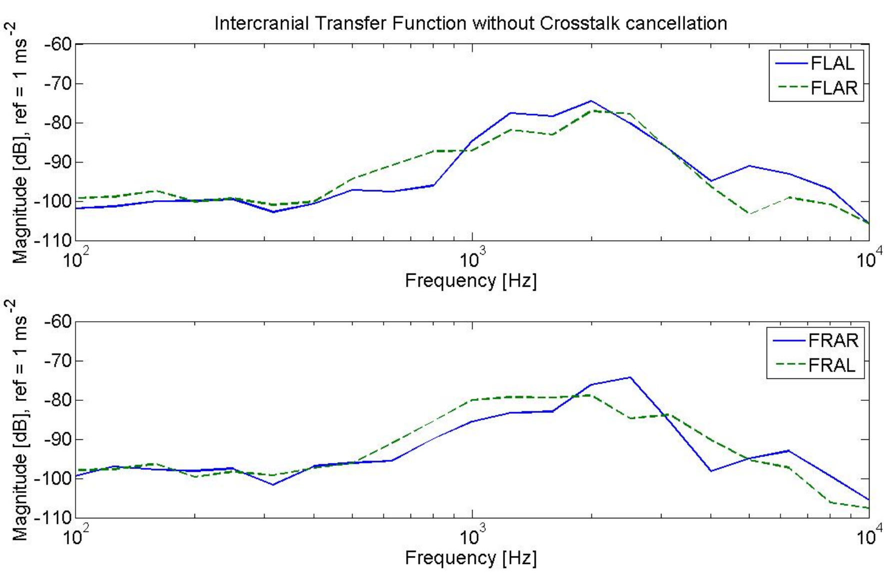


Figure 6.5.: Intercanial transfer function obtained by the recording method

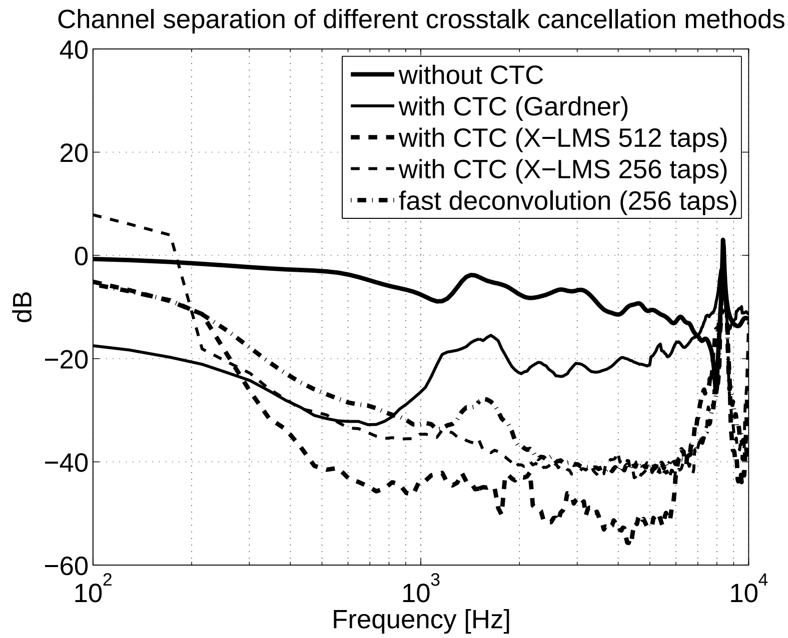


Figure 6.6.: Comparison of crosstalk cancellation algorithms

mostly in the anechoic chamber, so to say, there is no reverberation nor echo in the sound field. On contrary, if one imagines the skull is a room, it is not as damped as an anechoic chamber, but more like a room with reverberation. Therefore, to design a crosstalk cancellation filter to deal with the vibration in the skull is much more troublesome, especially doing the inversion of non-minimum phase transfer function.

The multi-error least mean square (ME-LMS) method, there are several tricky problems to consider first, one should have the feedback signal to adapt the filter coefficient. For the conventional application of stereo reproduction, two microphones are placed in the ear canals of the dummy head. The application of the BAHAs, one should think how to obtain the vibration signals of the cochlea. Second is that how to get an exact inverse of the transfer function with this method, especially in the reverberant field. Nelson et al [21] claimed that the combination of ME-LMS and multiple-input/output inverse filtering theorem (MINT) can get exact inverse in a reverberant room. To apply MINT, one should have at least one extra output signal compared to the number of input signals. In 3D audio application, people use three loudspeakers instead of the conventional two speakers to reach a better channel separation in the anechoic chamber and in a reverberant room as well. If one wants to apply this method in the skull with two BAHA, in order to get an exact inversion of the transfer function, one needs three BAHAs. That can be a challenge from the esthetic point of view and the complexity of surgery. One may prefer wearing hearing aids without crosstalk cancellation.

For the application of IIR method, one can see that the inversion of the transfer function is essential in this method to obtain A, B, C and D in Equation (3.6) and (3.8). Jot (1992) applied zero-phase low pass filter to the transfer impulse response and added some suitable delay to get B and C. Low pass filtering is the way to get rid of the non-minimum phase part of the transfer function, therefore being in the feedback path will not cause the divergent problem. Unfortunately, this method is only approved working in the anechoic chamber, it is more complex to get B and C blocks of the reverberant room, because of the non-minimum phase problem [12].

The last method has been tried in this project is the fast deconvoluton method. This method is so far feasible and can avoid the problems or limitations that the others methods have. Therefore it will be adapted to this project. There are more discussion in the following section.

6.2.2. Design the crosstalk cancellation filter coefficients with fast deconvolution method

After the intercranial transfer functions were measured, the ideal crosstalk cancellation filter coefficient was be calculated from the Equation (3.4); and the result is shown in Figure 6.7. In the figure, there are ill-condition problem at low frequency and high frequency range. At low frequencies, the ill-condition is caused because of that the skull vibrates as a rigid body, the two cochleae have same level vibration signal; at high frequencies, it could be caused by the same small level at both cochleae. The shape function is therefore designed as in Figure 6.8. By trial and error, one good value was found in this application, 0.0001. The length of the filter is set to 256 taps. By substituting the shape function and the gain factor to the Equation (3.25), the crosstalk cancellation filters are calculated and are shown in Figure 6.9. One can compare the amplitude of the filters calculated by the ideal crosstalk cancellation method and the fast deconvolution method. The dynamic range of crosstalk cancellation filter amplitude is lower than by using fast deconvolution method, so that is reasonable to be applied.

The intercranial transfer function with the crosstalk cancellation is first simulated in MATLAB. Figure 6.10 shows the simulated results of the ideal crosstalk cancellation method and the one with fast deconvolution method. It is a trade-off of using fast deconvolution method. One can see the crosstalk signal was not perfectly cancelled by the fast deconvolution method, especially the frequencies where the signal value was lowered by the shape function. The information has been distorted during calculation.

6.3. Crosstalk cancellation effect

By implementing the crosstalk cancellation filter in the DSP board, the transfer functions were measured and are shown in Figure 6.11. Compared to the ones without crosstalk cancellation in Figure 6.5, at low frequencies range, from about 250 Hz to about 1.5 kHz,

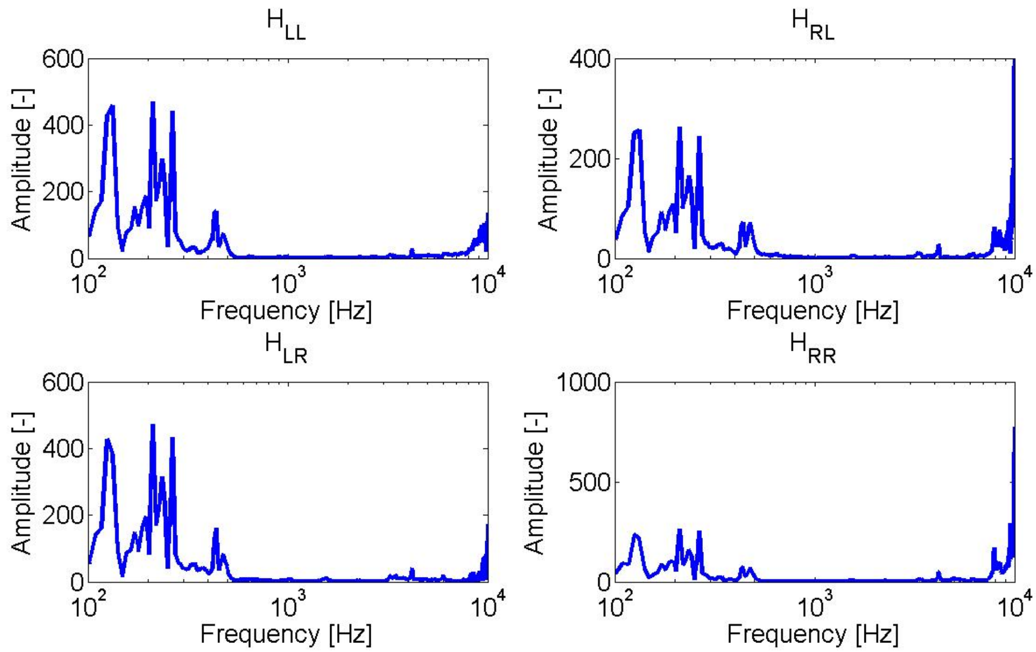


Figure 6.7.: Magnitude of the filters of ideal crosstalk cancellation for the application of two force transducers on the skull

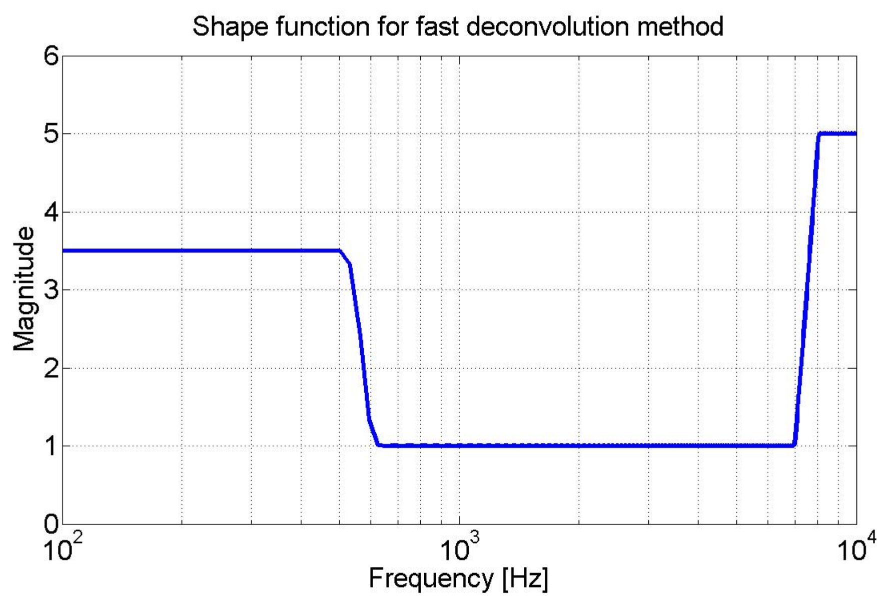


Figure 6.8.: Shape function

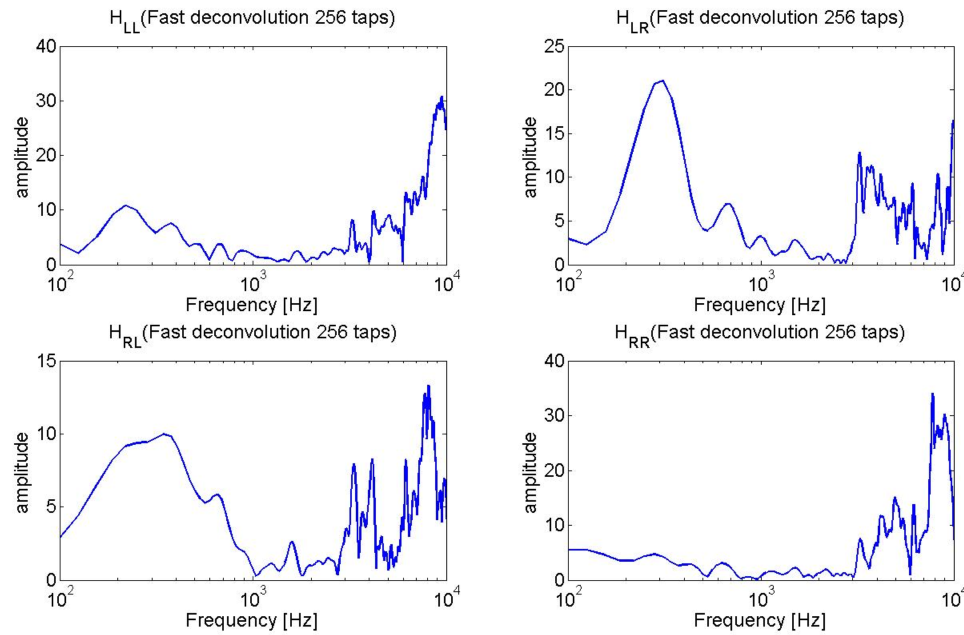


Figure 6.9.: Crosstalk cancellation coefficient from fast deconvolution method

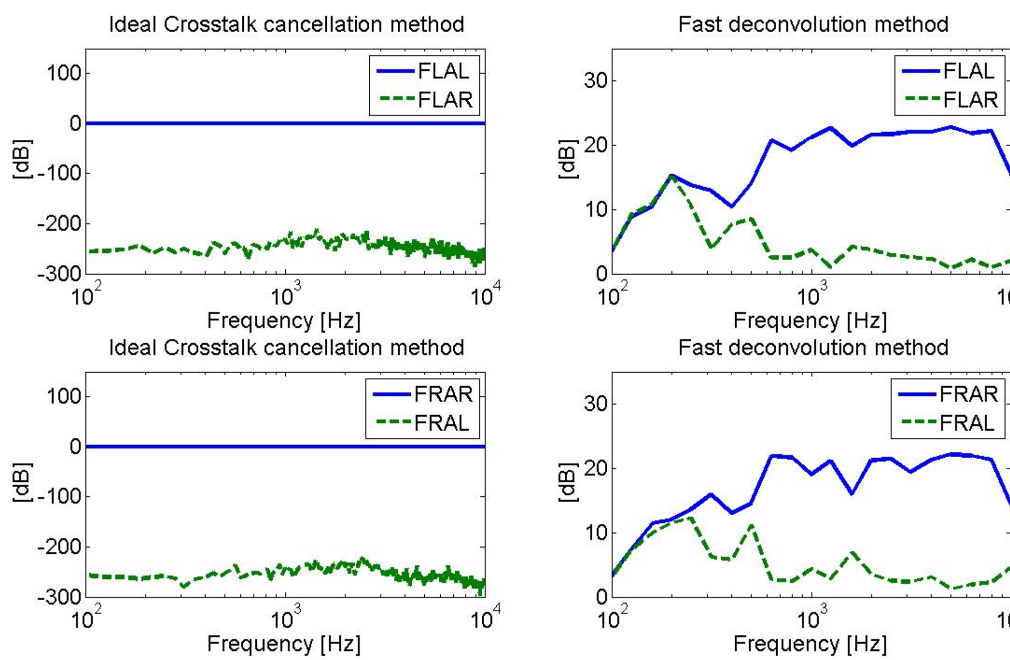


Figure 6.10.: Crosstalk cancellation simulation results by Ideal method and fast deconvolution

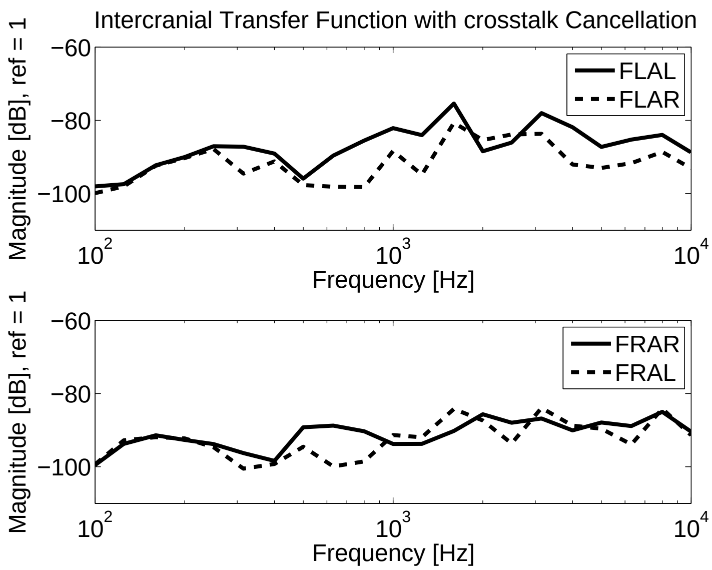


Figure 6.11.: Measured intercranial transfer function with crosstalk cancllation when implemented in DSP

one can see the crosstalk cancellation effect. But the attenuation of the crosstalk signal is not as large as the ones of simulation, especially above 1.5 kHz. There are two reasons: first, the system can be seen as a linear system at low frequencies, therefore it fits the model of crosstalk cancellation system shown in Figure 3.3. Once it goes to the higher frequency, it becomes a non-linear system and the model is not suitable anymore. Second, the wave types in the skull become complicated at higher frequencies [9]. As mentioned in chapter two, it is the combination of longitudinal and transverse wave. There is only one force transducer installed on the skull at each side which is not sufficient to cancel all the wave types happening at the same time. Furthermore, at lower frequencies, it is sure that the dominating force at the accurate eminence is the direction that the force transducer vibrates; but one is not sure what kind of wave type that dominates at the cochlea at high frequencies, since the force transducer is attached on the surface of the skull bone, but not on the cochlea directly [9].

The crosstalk cancellation effect can be seen as an artificial attenuation of the crosstalk signals. As discussed before, the crosstalk cancellation effect is only working below 1.5 kHz, so the discussion will be limited below 1.5 kHz.

The attenuation with and without crosstalk cancellation effect is shown in Figure 6.12. The dashed line shows the natural attenuation of the skull, with negative or zero dB attenuation below 1 kHz. The negative attenuation means that the signal is not decreased, but on contrary, is enlarged. It is because of that the body moves as whole body and the antiresonance phenomenon introduced in chapter two.

The solid line shows the attenuation with the crosstalk cancellation. It attenuates the

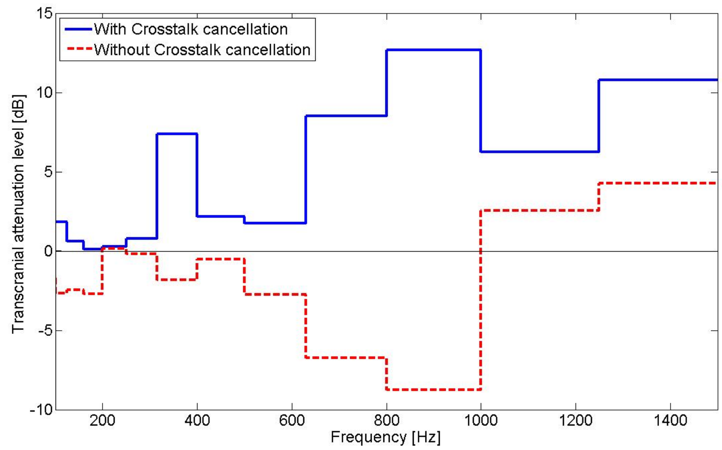


Figure 6.12.: Transcranial attenuation with and without crosstalk cancellation

signal much more than the natural one above 1 kHz, and prevents the antiresonance between 500 Hz to 1 kHz. At lower frequencies, it makes less attenuation, but still performs better than without crosstalk cancellation.

The attenuation level of the crosstalk signal to cochlea does not provide the information if one is benefit by it. One is interested how much the artificial attenuation helps one who wears BAHAs to improve their directional hearing abilities. The ILD is an indicator that shows the effect, and can be calculated from Equation (5.3). In order to simplify the system, the transfer function $FLAL$ and $FLAR$ are used as the TF_{IPSI} and TF_{CONTRA}. In another word, this system is assumed symmetric in the following discussion. Besides this, there are few more assumptions:

1. The delay time for signal transferring to the other side of BAHA to do the crosstalk cancellation is assumed to be zero, i.e. the wire connection is used. From Equation (3.3), one knows that the crosstalk cancellation matrix works with knowing the both the input signals, X_L and X_R. Using today's wireless technology, two devices without wire connection, it will take minimum 20 ms to do the data processing and transferring by Bluetooth or similar wireless technology. Such a delay could too large cause this algorithm to fail. Therefore, it is simplified into the best condition that no delay between two BAHAs.

2. The vibration measured at the petrous part of the temporal bone represents the vibration of the cochlea. As mentioned before, one can't measure the vibration of the cochleae nor can't know how much the person can really sense the sound signal when the petrous vibrates. Therefore it is assumed that the vibration level at the petrous part represents the signal that one perceived at the cochlea, and is proportional to the sound level one can perceived.

3. The cochleae are working well.

4. The system is Linear Time Invariant (LTI). The intercranial transfer function can be changed for a living person. For example, when the mouth is tightly closed, the jaw is connected to the head bone; or when the mouth is open, these are two different structures; the intercranial transfer functions will be definitely changed. Therefore, when a person is talking or is chewing, the whole system is time variant. But the system in this project is time invariant. The jaw was removed, and it is a dead human skull.

With all the assumptions, one can discuss how much contribution the crosstalk cancellation gives to improve ILD between two cochleae.

In Figure 6.13 shows four different ILD curves in different conditions at the mid frequency of 315 Hz. The x axis represents the sound source at the azimuth angle, as shown in Figure 5.1. The solid line, HRTF, is the ILD of two ears of the KEMAR, the data was from MIT Media Lab measurement. It is apparently that when the sound source incidents from 90 degree of the azimuth angle, it causes the largest ILD. It is because the position of the ear is at the side of the head, approximately 90 degree from the middle of the face. The 270 degree is just the opposite side of the ear; minus value of dB means the left side ear/microphone/cochlea gets larger signal.

The dash line is the ILD of two BAHAs, the largest ILD happens at 135 degree. This is because the BAHA is positioned a bit behind the ear. The other lines presents the ILD of the cochleae with and without crosstalk cancellation; the "with crosstalk cancellation" was from simulation. The curve, without crosstalk cancellation, has different trend from others; the cochlea opposites of the sound source gets higher level in this case. This antiresonance was introduced above. When the crosstalk cancellation is implemented, the ILD curves match people's expectation.

The ILD at 315 Hz calculated from measurement results is shown in Figure 6.14. The crosstalk cancellation indeed helps the ILD trend. The Figure 6.15 show the ILD at 1/3 octave frequency bands from 400 Hz until 1250 Hz. The crosstalk cancellation indeed improves the ILD.

The ITD inside of the skull is very difficult to estimate. Therefore it can't be interpreted more in this project.

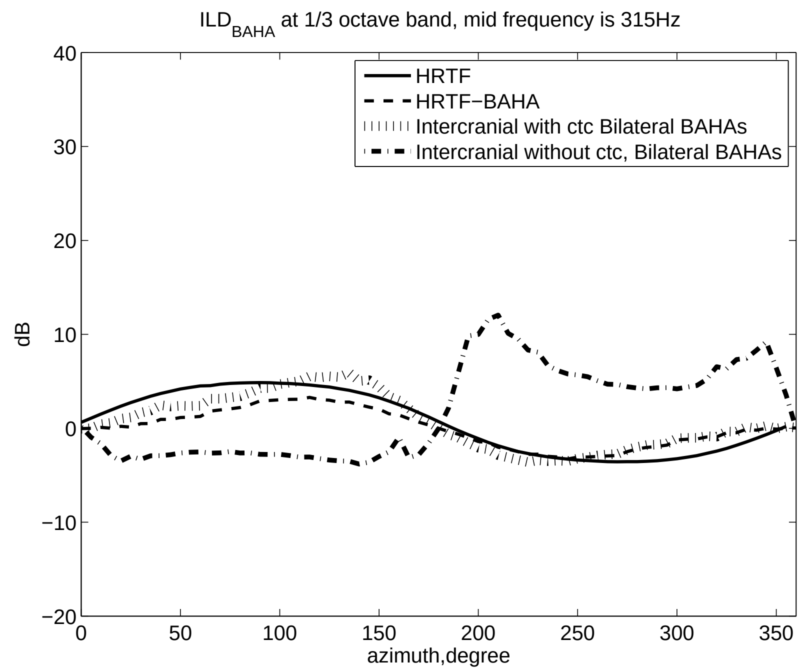


Figure 6.13.: ILD at mid frequency 315Hz

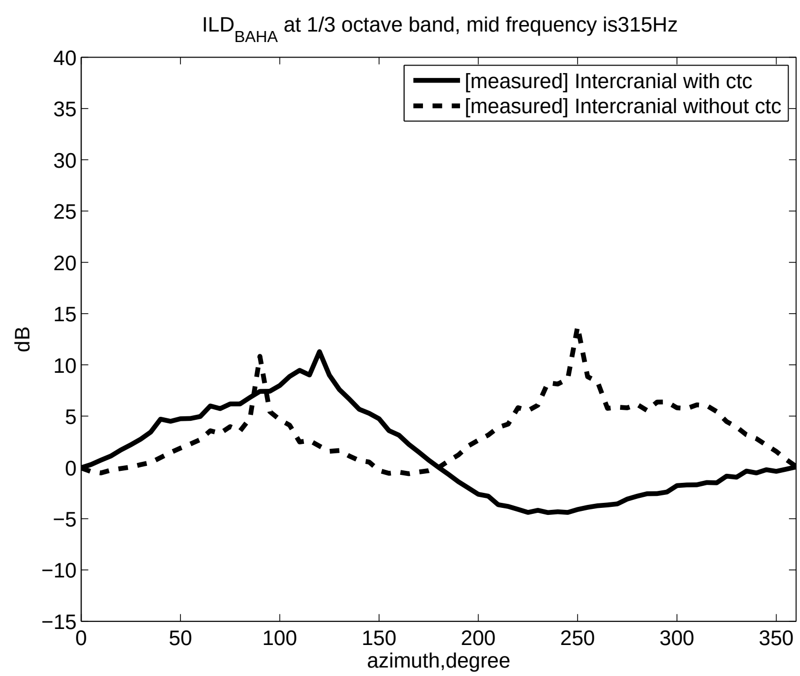
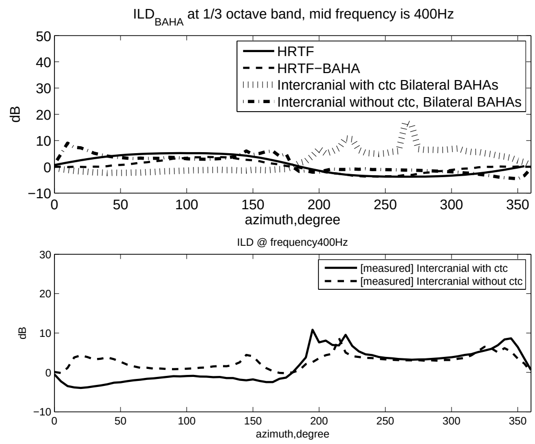
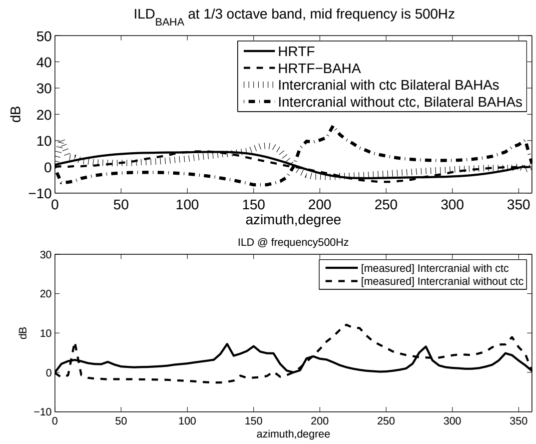


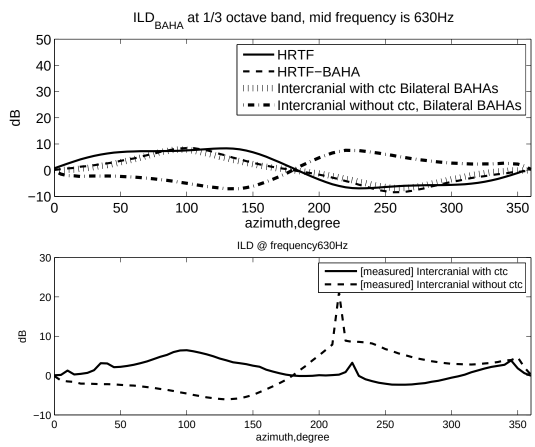
Figure 6.14.: Measured ILD at mid frequncy 315Hz



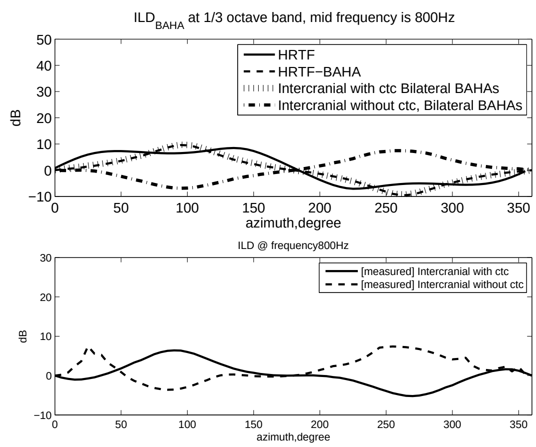
(a) 400Hz



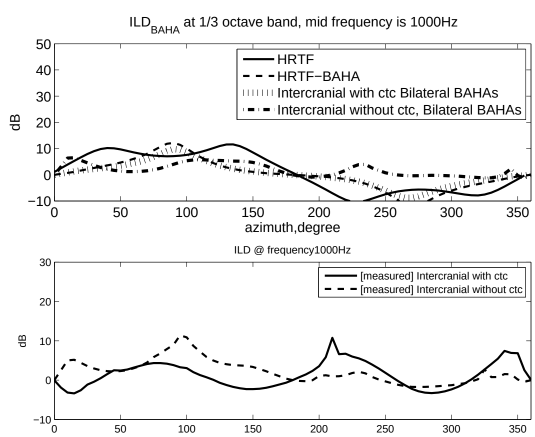
(b) 500Hz



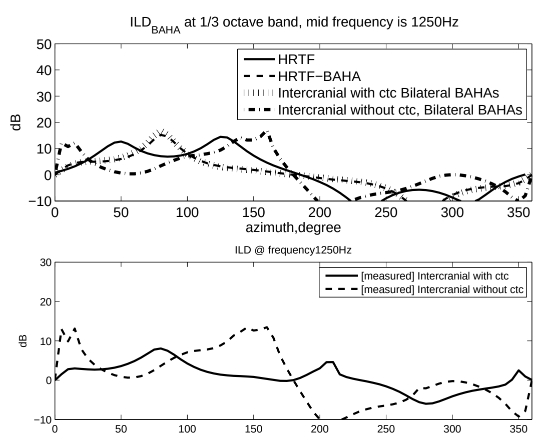
(c) 630Hz



(d) 800Hz



(e) 1000Hz



(f) 1250Hz

Figure 6.15.: ILD of cochleae at different frequencies

7. Conclusions and future work

7.1. Conclusion

For the implementation of crosstalk cancellation on the dry skull, there are two conclusions:

1. Fast deconvolution method of crosstalk cancellation works in the skull from 250 Hz to around 1.5 kHz and the attenuation of the crosstalk signal is from 2 to 12 dB in different frequencies.

2. The crosstalk cancellation improves the binaural direction property (ILD) below 1 kHz.

The results are based on the measurement on only one dry skull. It can't represent the situation on every individual. Nevertheless, the results can be helpful for people to understand how much crosstalk cancellation that can be achieved, and how much it improves the binaural directional properties with the existed knowledge and technology.

7.2. Future work

There are still quite many interesting questions in this project. If the time and resources are sufficient, more ideas can be carried out that can definitely help people to understand the crosstalk cancellation effect to the binaural directional property completely.

1. Perform listening tests. The investigation in this project is mainly done and judged from the values and curves. One can't really know how much the ILD of the cochleae is enough or is good.

2. Use real BAHA. The input transducers used in this project are two B&K force transducers, which is different from using real BAHAs. Because of, first, the force transducers in the measurement were attached on the surface of the skull instead of implant to the head. The coupling and mechanical properties will not be the same. Second, how the BAHA vibrates could be different from the force transducer. There could be more component of force or even rotating forces. One can't see the effect by using the force transducer.

3. Try to use the other algorithms. There are although some drawbacks or limitations for the algorithms introduced in the report, still more algorithms can be tried. For

example, the warped filter makes the filter length much shorter. It may reduce the calculating time and reduce the loading to the hardware, and therefore save the cost.

4. Use adaptive filters. As mentioned in chapter five, the system is assumed to be LTI. When it is applied to a real person, the system properties are changing all the time. The adaptive filter can therefore help to maintain the task of doing crosstalk cancellation. As to how to get the feedback signal to update the filter coefficient, it is still an open and interesting question. But it is still possible to conquer one day in the future.

5. Perform measurements with the jaw moving. Since the crosstalk cancellation filter coefficient is fixed, it is very interesting to know how much it can work in the time variant system.

6. Use wireless connection and include the time delay. It takes time to transmit the signal to opposite side of BAHA. It is interesting to know how much time delay the algorithm can manage and how much delay the user can stand.

Bibliography

[1] C. Priwin, S. Stenfelt, G. Granström, A. Tjellström, and B. Håkansson, "Bilateral bone-anchored hearing aids(bahas): An audiometric evaluation," *The Laryngoscope*, vol. 114, pp. 77–84, 2004.

[2] S. Stenfelt, "Bilateral fitting of bahas and baha fitted in unilateral deaf persons: Acoustical aspects," *International Journal of Audiology*, vol. 44, no. 3, pp. 178–189, 2005.

[3] A. J. Bosman, A. F. M. snik, C. T. M. van der Pouw, E. A. M. Mylanus, and C. W. J. Cremers, "Audiometric evaluation of bilaterally fitted bone-anchored hearing aids," *International Journal of Audiology*, vol. 40, no. 3, pp. 158–167, 2001.

[4] S. Stenfelt, "Hearing by bone conduction physical and physiological aspects," Ph.D. dissertation, Chalmers University of Technology(CTH), Göteborg, Sweden, 1999.

[5] P. U. Carlsson, "On direct bone conduction hearing device- advances in transducer technology and measurement method," Ph.D. dissertation, Chalmers University of Technology(CTH), Göteborg, Sweden, 1990.

[6] [Online]. Available: http://www.csmc.edu/14576.html

[7] B. C. Moore, Ed., *An Introduction to the Psychology of Hearing*.

[8] [Online]. Available: http://www.ent.uci.edu/BAHA.htm

[9] S. Stenfelt and R. L. Goode, "Transmission properties of bone conducted sound: Measurements in cadaver heads," *The Journal of the Acoustical Society of America*, vol. 118, no. 4, pp. 2373–2391, 2005.

[10] S. Stenfelt, B. Håkansson, and A. Tjellström, "Vibration characteristics of bone conducted sound in virto," *The Journal of the Acoustical Society of America*, vol. 107, no. 1, pp. 422–431, 2000.

[11] B. Håkansson, "The bone anchored hearing aid-engineering aspects," Ph.D. dissertation, Chalmers University of Technology(CTH), 1984.

[12] W. G. Gardner, Ed., *3-D AUDIO USING LOUDSPEAKERS*.

[13] P. A, Nelson, H. Hamada, and S. J. Elliot, "Adaptive inverse filters for stereophonic sound reproduction," *IEEE Transactions on signal processing*, vol. 40, no. 7, pp. 1621–1632, 1992.

[14] O. Kirkeby, P. A. Nelson, H. Hamada, and F. Orduna-Bustamante, "Fast deconvolution of multichannel systems using regularization," *IEEE Transactions on speech and audio processing*, vol. 6, no. 2, pp. 189–195, 1998.

[15] O. Kirkeby, P. Rubak, P. A. Nelson, and A. Farina, "Design of cross-talk cancellation networks by using fast deconvolution," in *AES 106th Convention*, Munich, Germany, May 1999.

[16] O. Kirkeby, P. A. Nelson, and H. Hamada, "Digital filter design for virtual source imaging systems," in *AES 104th Convention*, Amsterdam, May 1998.

[17] P. Andersson, "Vibroacoustic measurement equipment at applied acoustics," 2007.

[18] D. Bell, "Applications using the TMS320C6000 Enhanced DMA," Texas Instruments, Application Report SPRA636A, October 2001.

[19] J. Eriksson and J. Sverker, "Head related transfer function with the bone anchored hearing aids: Benefits in directionality of bilateral application compared to unilateral application," Göteborg, Sweden, 2003.

[20] S. Kleiven, "Sound and vibration measurement laboratory work4: Signal analysis," 2007.

[21] P. A. Nelson, F. Orduna-Bustamante, and H. Hamada, "Inverse filter design and equalization zones in multichannel sound reproduction," *IEEE Transactions on speech and audio processing*, vol. 3, no. 3, pp. 185–192, 1995.

A. ME-LMS crosstalk cancellation algorithm

ME-LMS has been introduced in chapter three. The detail of this algorithm will be interpreted here. In Figure 3.6 shows a 2X2 system which has two loudspeakers and two microphones. The mathematical expression can be written as Equation (A.1) below:

$$\begin{bmatrix} \hat{Z}_R \\ \hat{Z}_L \end{bmatrix} = \begin{bmatrix} C_{RR} & C_{RL} \\ C_{LR} & C_{LL} \end{bmatrix} \begin{bmatrix} H_{RR} & H_{RL} \\ H_{LR} & H_{LL} \end{bmatrix} \begin{bmatrix} X_R \\ X_L \end{bmatrix} \tag{A.1}$$

Equation (A.1) can be rewritten as Equation (A.2)

$$\begin{bmatrix} \hat{Z}_R \\ \hat{Z}_L \end{bmatrix} = \begin{bmatrix} X_R C_{RR} & X_R C_{RL} & X_L C_{RR} & X_L C_{RL} \\ X_R C_{LR} & X_{LC_{LL}} & X_L C_{LR} & X_L C_{LL} \end{bmatrix} \begin{bmatrix} H_{RR} \\ H_{LR} \\ H_{RL} \\ H_{LL} \end{bmatrix} \tag{A.2}$$

The signals $X_k C_{lm}$ are defined as the "filtered reference signals" R_{lmk}. Equation (A.2) can be therefore redefined as Equation (A.3):

$$\begin{bmatrix} \hat{Z}_R \\ \hat{Z}_L \end{bmatrix} = \begin{bmatrix} R_{RRR} & R_{RLR} & R_{RRL} & R_{RLL} \\ R_{LRR} & R_{LLR} & R_{LRL} & R_{LLL} \end{bmatrix} \begin{bmatrix} H_{RR} \\ H_{LR} \\ H_{RL} \\ H_{LL} \end{bmatrix} \tag{A.3}$$

The filtered reference signals R_{lmk} in frequency domain can be defined as in Equation (A.4):

$$R_{lmk} = C_{lm} X_k \tag{A.4}$$

In frequency domain, can be defined as Equation (A.5):

$$r_{lmk}(n) = \sum_{j=0}^{J-1} c_{lm}(j)\, x_k(n-j) \tag{A.5}$$

It is assume the C matrix has J coefficients and the crosstalk cancellation filter h as I coefficients. The filter h can be defined as:

$$h^T = \left[\begin{array}{ccc|ccc|ccc|ccc} h_{RR}(0) & \cdots & h_{RR}(I-1) & h_{LR}(0) & \cdots & h_{LR}(I-1) & h_{RL}(0) & \cdots & h_{RL}(I-1) & h_{LL}(0) & \cdots & h_{LL}(I-1 \end{array} \right. \tag{A.6}$$

The resulting output can be written as:

$$
\begin{aligned}
\hat{z}_R &= r_R^T(n)\,h \\
\hat{z}_L &= r_L^T(n)\,h
\end{aligned}
\tag{A.7}
$$

Where

$$
\begin{aligned}
r_R^T(n) &= \begin{bmatrix} r_{RRR}^T(n) & r_{RLR}^T(n) & r_{RRL}^T(n) & r_{RLL}^T(n) \end{bmatrix} \\
r_L^T(n) &= \begin{bmatrix} r_{LRR}^T(n) & r_{LLR}^T(n) & r_{LRL}^T(n) & r_{LLL}^T(n) \end{bmatrix}
\end{aligned}
\tag{A.8}
$$

Each component in Equation (A.8) is defined as

$$
r_{lmk}^T(n) = \begin{bmatrix} r_{lmk}(0) & r_{lmk}(n-1) & r_{lmk}(n-2) & \cdots & r_{lmk}(n-I+1) \end{bmatrix}
\tag{A.9}
$$

The net output can be expressed as the composite vector

$$
\hat{z}(n) = \mathbf{R}(n)\,h
\tag{A.10}
$$

Where

$$
\begin{aligned}
\hat{z}(n) &= \begin{bmatrix} \widehat{z}_R(n) \\ \widehat{z}_L(n) \end{bmatrix} \\
\mathbf{R}(n) &= \begin{bmatrix} \widehat{r}_R(n) \\ \widehat{r}_L(n) \end{bmatrix}
\end{aligned}
\tag{A.11}
$$

There are always deviations between the desired signal z and the actual desire reach the ears $\hat{z}$. In order to minimize the error value, the least mean square method is therefore introduced here.

A variant J is defined as below:

$$
J = E\left[(z(n) - \hat{z}(n))^T (z(n) - \hat{z}(n))\right]
\tag{A.12}
$$

Where E denotes the expectation operator and $z(n)$ is the vector defined by $Z^T(n) = [z_1(n)z_2(n)]$. Therefore Equation (A.12) can be rewritten as:

$$
J = E\left[(z(n) - R(n)\,h)^T (z(n) - R(n)\,h)\right]
\tag{A.13}
$$

After expansion and use of the expection operator reduces to the quadratic form:

$$
J = E\left[z^T(n)\,z(n)\right] - 2E\left[z^T(n)\,\mathbf{R}(n)\right]h + h^T E\left[\mathbf{R}^T\mathbf{R}(n)\right]h
\tag{A.14}
$$

To minimum Equation (A.14), one can take the gradient of it, and it is shown in Equation (A.15). By doing iteratively update of the value of the composite tap weight vector by an amount proportional to the negative of the gradient of the quadratic performance surface.

$$\frac{\partial J}{\partial h} = -2E\left[\mathbf{R}^T(n)\,z(n)\right] + 2E\left[\mathbf{R}^T(n)\,\mathbf{R}(n)\right]h \tag{A.15}$$

Equation (A.15) can be rewritten as:

$$\frac{\partial J}{\partial h} = -2E\left[\mathbf{R}^T(n)\,e(n)\right] \tag{A.16}$$

Where the instantaneous error vector is:

$$e(n) = z(n) - \mathbf{R}(n)\,h \tag{A.17}$$

The filter h is updated every step according to the equation below, $\alpha$is the convergence coefficient of the algorithm.

$$h(n+1) = h(n) + \alpha R^T(n)\,e(n) \tag{A.18}$$