

Simple Individualization of Head Related Transform Functions

An Interactive Qualifying Project
submitted to the Faculty of
WORCESTER POLYTECHNIC INSTITUTE
in partial fulfillment of the requirements for the
Degree of Bachelor of Science

By
Haofan Zhang

Date:
June 2017

Submitted to:
Frederick Bianchi
Worcester Polytechnic Institute

Abstract

The purpose of this project is to provide a simpler and faster solution for calibration of Head Related Transform Function (HRTF) for individuals. The project developed a concept which should make it much easier than traditional method for average users to personalize their own HRTFs, as well as a Max patch based on this concept for testing.

Acknowledgements

I would like to thank Jakob Hougaard Andersen for sharing his *FFT-based binaural panner* on the cycling74 community tools section, which is what the entire project is based on. Also thanks Yuheng Huo, Mengwen Li, Erik Paulson, Yihong Zhou for their *Comparison of Surround Sound Perception with Headphone Perception using HRTF*, which gave me much intel about *FFT-based binaural panner*, HRTF and CIPIC database. Final thanks to Professor Fredrick Bianchi, the project advisor, for his awesome advices on the direction and approach for this project.

Table of Contents

Abstract.....	1
Acknowledgements	2
Table of Contents	3
Table of Figures.....	4
1. Introduction.....	5
2. Methodology	7
2.1 Potential Approach: Anthropometric Measurements	7
2.2 Potential Approach: Improve Traditional Method	9
2.3 Taken Approach: Model Based Calibration	10
2.4 Information About CIPIC Database	10
2.5 Model Based Matching	12
3. Prototyping and Testing	16
4. Conclusion and Discussion.....	25
5. Bibliography	28
6. Appendixes.....	29
A. ITD Raw Data	29
B. FFT Raw Data	31
C. ITD Models Chart	32
D. FFT Models Chart	33
E. Main Patch Instructions	36
F. Supportive Max Patches	38

Table of Figures

Figure 2.1.1: Head and torso measurements	7
Figure 2.1.2: Pinna measurements	8
Figure 2.1.3: A list of anthropometric parameters	8
Figure 2.4.1: Elevation and Azimuth of (a) front (b) side	10
Figure 2.4.2: A sample FFT	11
Figure 2.4.3: Convolution formula	11
Figure 2.5.1: Subject 8, 11, 17's FFT at 0/0	12
Figure 2.5.2: Subject 8, 11, 17, 10's FFT at 36/0	13
Figure 2.5.3: Left ear ITDs of 10 subjects at 7 directions	14
Figure 3.1: Loader	16
Figure 3.2: ITD matcher	18
Figure 3.3: ITD confirmation	19
Figure 3.4: FFT matcher	20
Figure 3.5: HRTF subject matcher	22
Figure 3.6: Audio player	23

1. Introduction

As higher quality content consumption is becoming more of a commonplace, people expect realistic surround sounds in video or audio contents. Home theater enthusiasts use 5.1 channel or even 7.1 channel systems in their homes to bring the realistic environment from movies.

However, multi-channel audio systems are limited in many ways. Since each channel can only produce sound from one direction, and it is impossible to create a system of infinite number of channels. Furthermore, multi-channel systems are expensive, and in most cases, of poor portability. Fortunately, researchers developed Head Related Transform Function (HRTF) to solve the problem.

HRTF is a set of algorithms which simulates the function of human ears in terms of modifying delay, amplitude and equalization of frequency domains of a sound to let people recognize the direction of a sound in 3D space. Theoretically by applying HRTF on a sound, it can appear to be coming from any direction, even on a stereo system, and the listener should be able to accurately track the exact location of the sound object.

Many stereo and multi-channel systems for home theater and VR, use technology like Dolby TrueHD¹, which implements HRTF, for spatially rich sound. HRTF is also widely used in contents like music, videos and games. Software like MNTN² allows content creators to place their audio tracks in separate spatial locations for interesting and realistic sonic experiences.

¹ "Dolby TrueHD." Wikipedia. Wikimedia Foundation, 10 June 2017. Web. 02 July 2017. <https://en.wikipedia.org/wiki/Dolby_TrueHD>.

² "What is Binaural 3D Audio?" Mntn.rocks. MNTN, 23 Mar. 2017. Web. 02 June 2017. <<https://mntn.rocks/blog/2017/3/6/what-is-binaural-3d-audio>>.

Games like Rainbow Six: Siege allows players to listen to footsteps as a valuable source of information to tell where the enemies are located.

However, HRTF can vary significantly between listeners based on physical parameters of the ears and heads. The perception improvement (of personalized HRTF) compared to non-personalized HRTF is massive³. Each listener has their own HRTFs because sound processed by standardized HRTF will be perceived differently by different people.

Traditional individualized HRTF measurement involves putting a microphone inside each of the listener's ears and play a standardized sound from several hundred spatial locations and constructing the HRTF by measuring the difference between the standardized sound and the microphone's input. This is a procedure which requires expensive and complicated equipment, and which can take upwards of two hours⁴, which is not at all suitable for average content consumers or stereo system users. This project is thus intended to provide a simpler, faster, cheaper and more accessible solution for average users to personalize their own HRTFs.

³ "Personal HRTFs." 3D Sound Labs. N.p., n.d. Web. 02 July 2017. <<http://pro.3dsoundlabs.com/category/personal-hrtfs/>>.

⁴ Bibliography no.4, page 3.

2. Methodology

Before the method used for this project came up, a few other existing methods are discovered and evaluated.

2.1 Potential Approach: Anthropometric Measurements

Three projects below took the approach of measuring body parts (anthropometric measurements) of users to generate HRTFs.

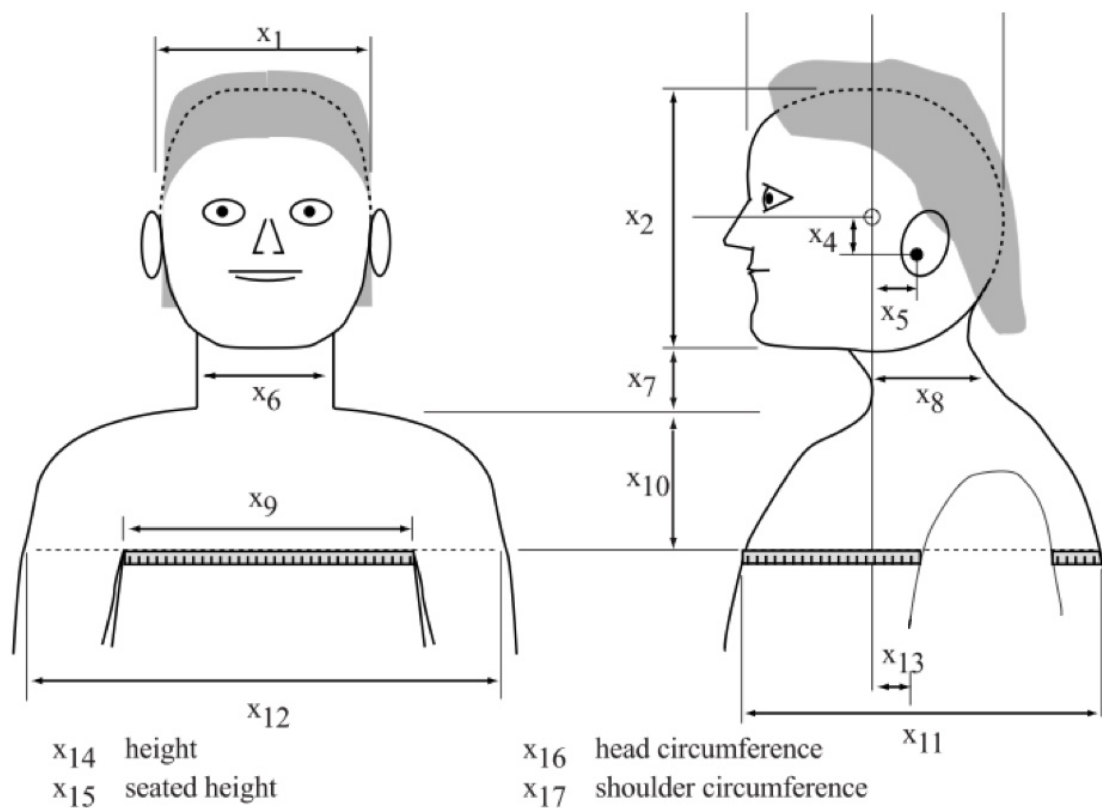


Figure 2.1.1: Head and torso measurements

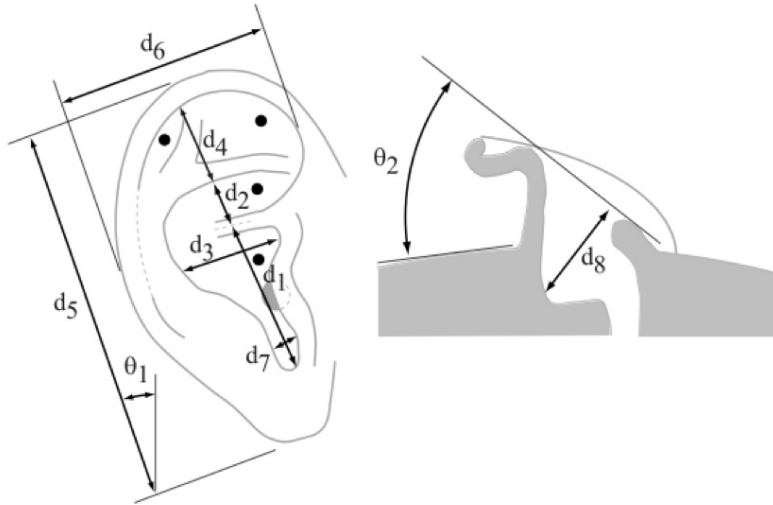


Figure 2.1.2: Pinna measurements

Var	Measurement
x_1	head width
x_2	head height
x_3	head depth
x_4	pinna offset down
x_5	pinna offset back
x_6	neck width
x_7	neck height
x_8	neck depth
x_9	torso top width
x_{10}	torso top height
x_{11}	torso top depth
x_{12}	shoulder width
x_{13}	head offset forward
x_{14}	height
x_{15}	seated height
x_{16}	head circumference
x_{17}	shoulder circumference
d_1	cavum concha height
d_2	cymba concha height
d_3	cavum concha width
d_4	fossa height
d_5	pinna height
d_6	pinna width
d_7	intertragal incisure width
d_8	cavum concha depth
θ_1	pinna rotation angle
θ_2	pinna flare angle

Figure 2.1.3: a list of anthropometric parameters

The HRTFSubjectMatcher⁵ patch in FFT-based binaural panner takes 27 different anthropometric measurements (Figure 2.1.3) including width, height and depth of user's

head, pinna, neck, torso and more. It compares data in CIPIC database with provided anthropometric parameters to find the closest match, without needing users to put microphones in their ears. But average users most likely won't have the equipment to accurately measure everything about their heads, nor would they have time to undertake the time consuming task.

*Improved method to individualize head-related transfer function using anthropometric measurements*⁶ uses and improves upon an existing method proposed by Kistler and Wightman, which performs a principal components analysis (PCA) on standardized HRTFs to detect basis

⁵ Bibliography no.1

⁶ Bibliography no.5

functions and generate users's HRTFs by multiple regression of their anthropometric measurements. Instead of comparing to existing database, it generates estimates of HRTFs, which should create less error. However, it still requires anthropometric data directly from users.

3D Sound Labs⁷ makes it easier for users by letting them take pictures of their ears, and uses image recognition and algorithms to reproduce the 3D model of users' ears. This eliminates having users measure everything by themselves and only takes seconds for a calibration to complete. In practice, a picture won't show any detail of in-ear depths, since light conditions can greatly vary the shadow, which determines how deep an object is. 3D Sound Labs takes far less anthropometric data than CIPIC database and other similar projects. These issues make the 3D models and generated HRTFs rather questionable.

2.2 Potential Approach: Improve Traditional Method

*Individualized Head-Related Transfer Functions: Efficient Modeling and Estimation from Small Sets of Spatial Samples*⁸ took the approach of improving the traditional method. It reduces the sample size of the traditional method and uses algorithms to estimate the rest of the HRTF, thus requiring less time for the calibration to complete.

⁷ "Personal HRTFs." 3D Sound Labs. N.p., n.d. Web. 02 July 2017. <<http://pro.3dsoundlabs.com/category/personal-hrtfs/>>.

⁸ Bibliography no.4

2.3 Taken Approach: Model Based Calibration

The method used by this project is inspired by the FFT-based binaural panner's matching solution, but instead of matching anthropometric measurements, it generates models from HRTFs in the CIPIC database and allows users to test each model and pick the best fit.

2.4 Information About CIPIC Database

Before showing the details of Model Based Calibration, some basic information about CIPIC database can be helpful.

Each subject's HRTF in the CIPIC database consists of two parts: Interaural Time Delay (ITD) and Fast Fourier Transformation (FFT). Each ITD and FFT is assigned to an azimuth, elevation, and ear (left or right).

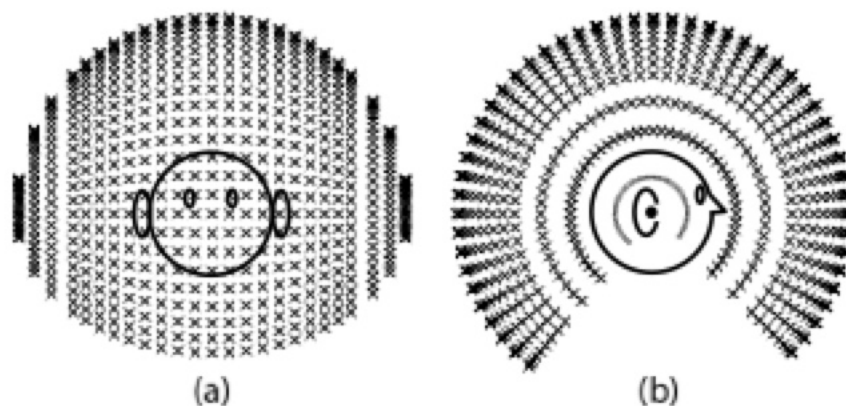


Figure 2.4.1: Elevation and Azimuth of (a) front (b) side

Azimuth is the horizontal direction of a sound (left/front/right), elevation is the vertical direction of a sound (front/top/back/bottom). Each HRTF has 73 (0 is left most, 72 is right most) azimuths and 128 elevations (0 is 45 degrees downward, 16 is front, 48 is top, 80 is back, 112 is bottom).

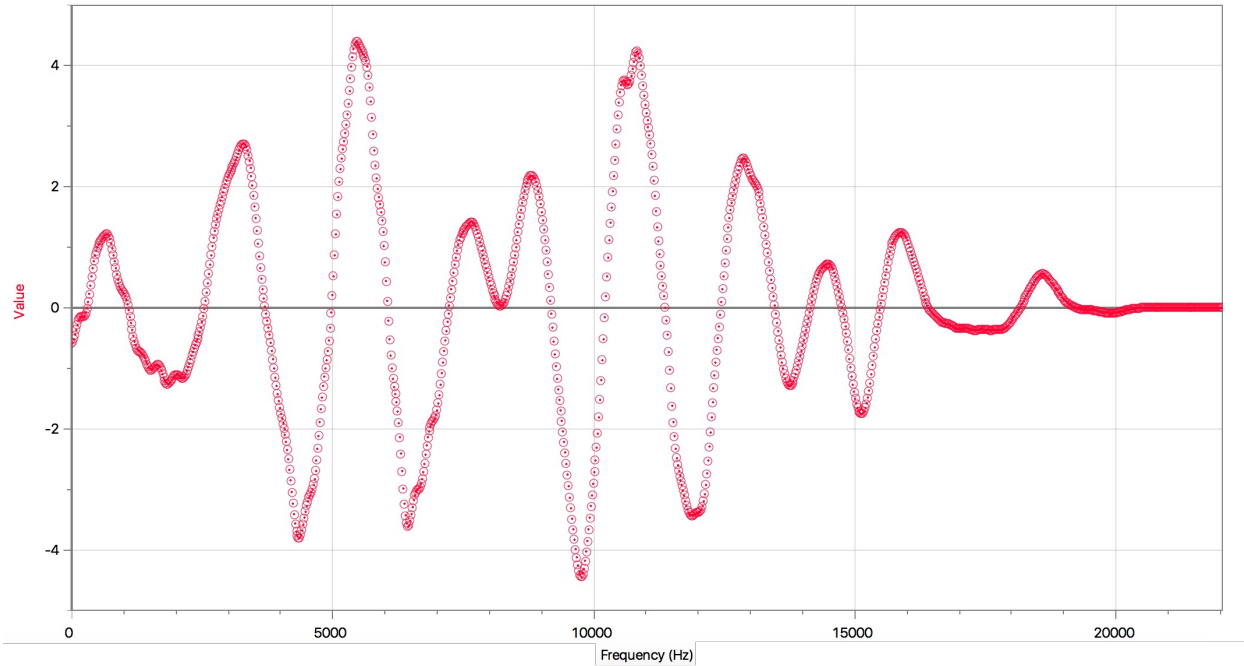


Figure 2.4.2: A sample FFT

The database is based on 44.1kHz recordings and each FFT consists of 2048 (1024 real and 1024 imaginary) samples. Therefore the first 1024 data points resembles the FFT on a 22.04kHz spectrum. ITD is stored as time in milliseconds.

$$\begin{aligned}
 (f * g)(t) &\stackrel{\text{def}}{=} \int_{-\infty}^{\infty} f(\tau) g(t - \tau) d\tau \\
 &= \int_{-\infty}^{\infty} f(t - \tau) g(\tau) d\tau.
 \end{aligned}$$

Figure 2.4.3: convolution formula

Convolution is a mathematical operation on two functions; it produces a third function, that is typically viewed as a modified version of one of the original functions, giving the integral of the pointwise multiplication of the two functions as a function of the amount that one of the original functions is translated⁹. By convoluting the FFT of a direction with the original sound and applying left ITD and right ITD of that direction on left and right channel, it will sound like it is coming from that direction.

2.5 Model Based Matching

By analyzing ITDs and FFTs of 10 subjects at characteristic azimuths and elevations, I discovered something interesting. Subjects can be split into several groups as their FFTs follow similar patterns, which can be observed visually through graphs. Those FFTs of the same group tend to fluctuate more or less at specific frequencies and have a very similar overall shape/trend. For example, subject 8's FFT has the same pattern with subject 11's and subject 17's at 0 azimuth and 0 elevation (from now on, 0/0) (Figure 2.5.1).

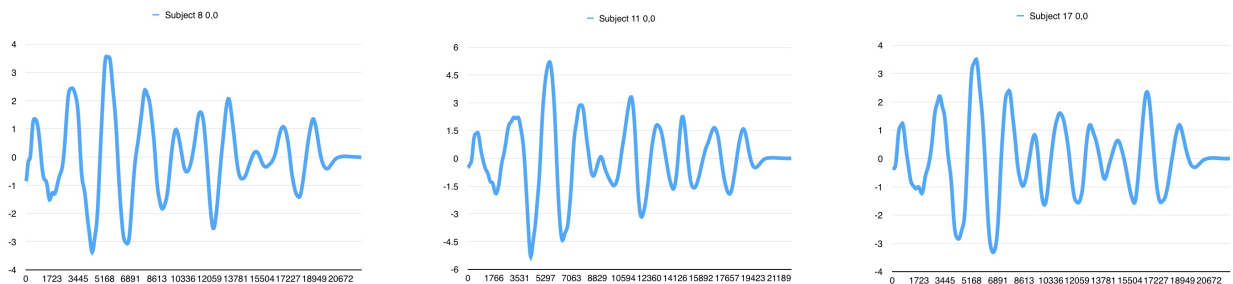


Figure 2.5.1: Subject 8, 11, 17's FFT at 0/0

⁹ "Convolution" Wikipedia. Wikimedia Foundation, 10 June 2017. Web. 02 July 2017. <<https://en.wikipedia.org/wiki/Convolution>>.

What's more interesting is that a subject's FFT can have a similar pattern with one group of subjects' FFTs of a direction, while having a similar pattern with another group of subjects' of different direction. For example, on Figure 2.5.2, subject 8's FFT does not have the same pattern with subject 11's and subject 17's at 36/0. Instead, it does with subject 10's.

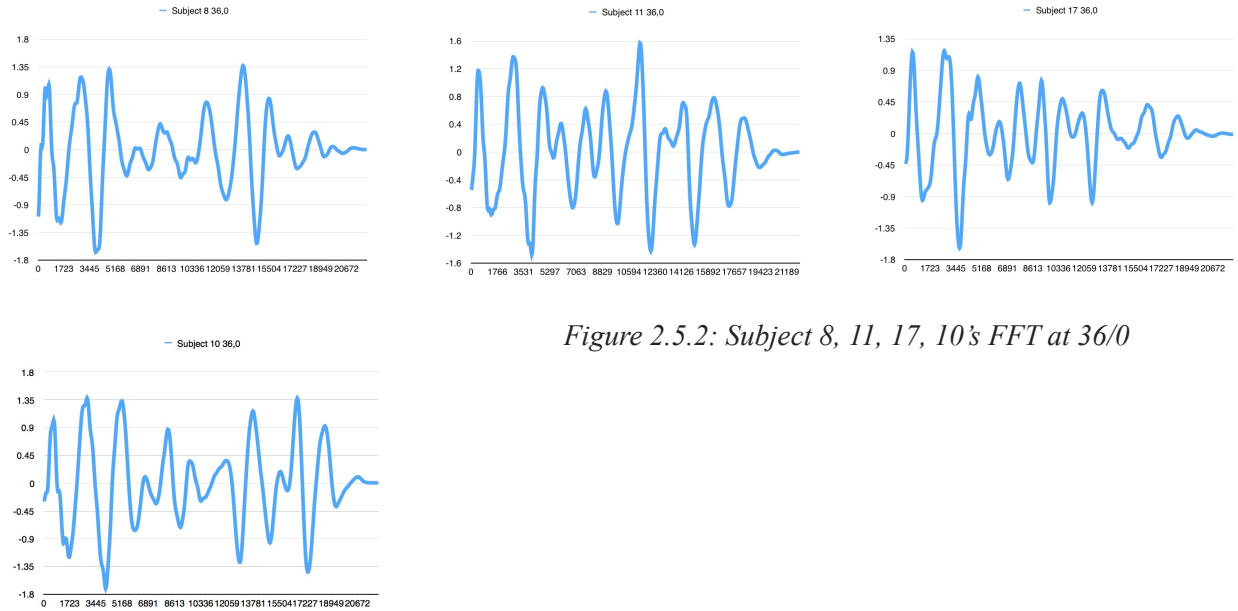


Figure 2.5.2: Subject 8, 11, 17, 10's FFT at 36/0

According to this, if we can analyze these patterns and make them into models, a combination of models for all directions represents one possible FFT set, and the entirety of such combinations covers all possible individualized FFT sets. If people can test and pick their own best fits of models for every direction, they will get their own personalized FFT sets.

Similarly, ITD for different subjects can also be made into models. Figure 2.5.3 shows subject 17 has relatively higher ITDs across the board, subject 10 has relatively lower ITDs across the board and subject 8 has average ITDs across the board. Some other subjects have pretty interesting ITDs. Subject 9 has higher ITDs for sound coming from left/right, and has a lower ITD for which coming from direct front. Subject 18 has lower ITDs for sound coming from left/right, while having an average ITD for which coming from direct front. Again, by acquiring the high, average and low points of ITDs from the database and letting users pick their best ITD fits, they can build their own ITD sets.

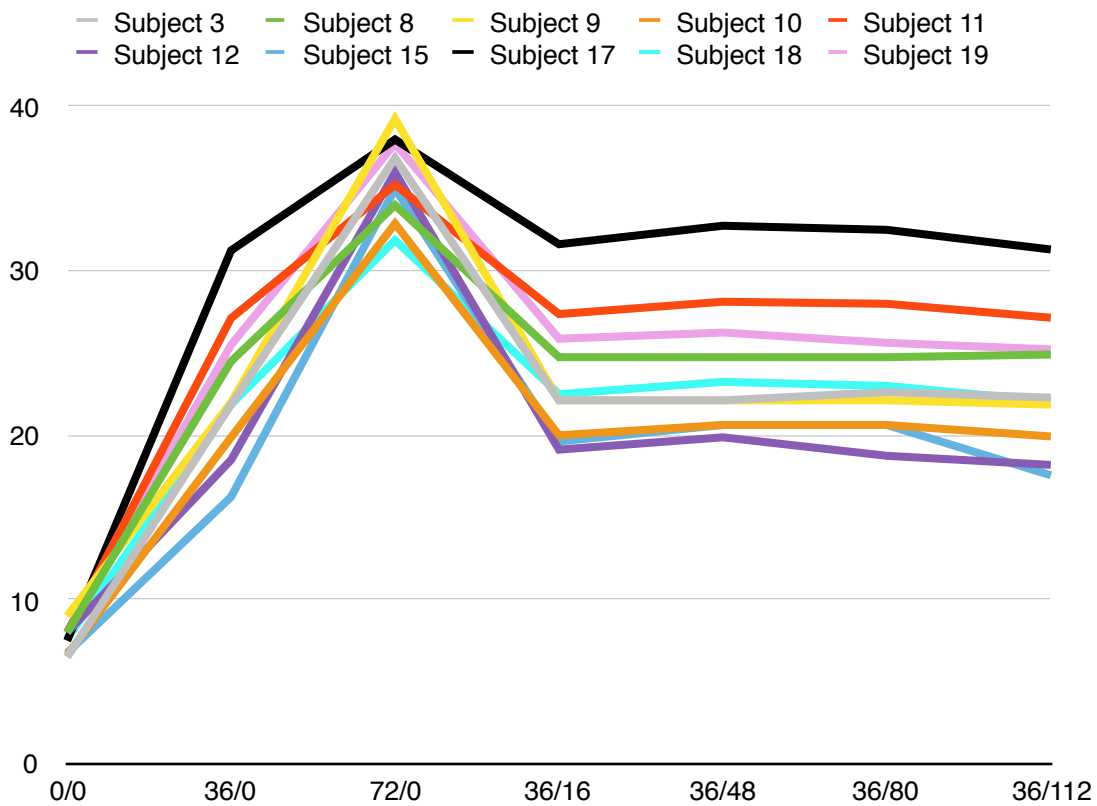


Figure 2.5.3: Left ear ITDs of 10 subjects at 7 directions

However, it is neither possible for users to manually match over 9000 possible directions (73 x 128) for both the FFT set and the ITD set, nor our goal to develop such a complicated method. Therefore 7 directions: 0/0, 36/0, 72/0, 36/16, 36/48, 36/80, 36/112 are used, as they cover 6 basic directions (front, back, left, right, top, bottom) in 3D space. Instead of building individual HRTFs from scratch, users can pick models which sound the closest to these 7 directions. By comparing their picks to the database, the best matching subject profile can be acquired.

Meanwhile, I observed that elevation change has a minor effect on ITD as long as azimuth is not changing. The average standard deviation for each subject's ITD at center front, top, back, bottom and 45° downward is 0.567 dB for the left ear, and 1.163 dB for the right ear, which is relatively minor compared to an average of 24.556 dB for the left ear, and 24.144 dB for the right ear. So instead of doing database matches, the rest of ITDs are estimated for better accuracy. ITDs of the same azimuth with all elevations will be assigned the same value, while ITDs of azimuth from 1-35 and 37-71 will be assigned values in between ITD of 0 and 36, and which of 36 and 72 respectively in a linearly increasing/decreasing fashion.

3. Prototyping and Testing

Following the design concept of model based matching, a MaxMSP patch was made as a prototype to test its effectiveness. It used a database of 10 subjects.

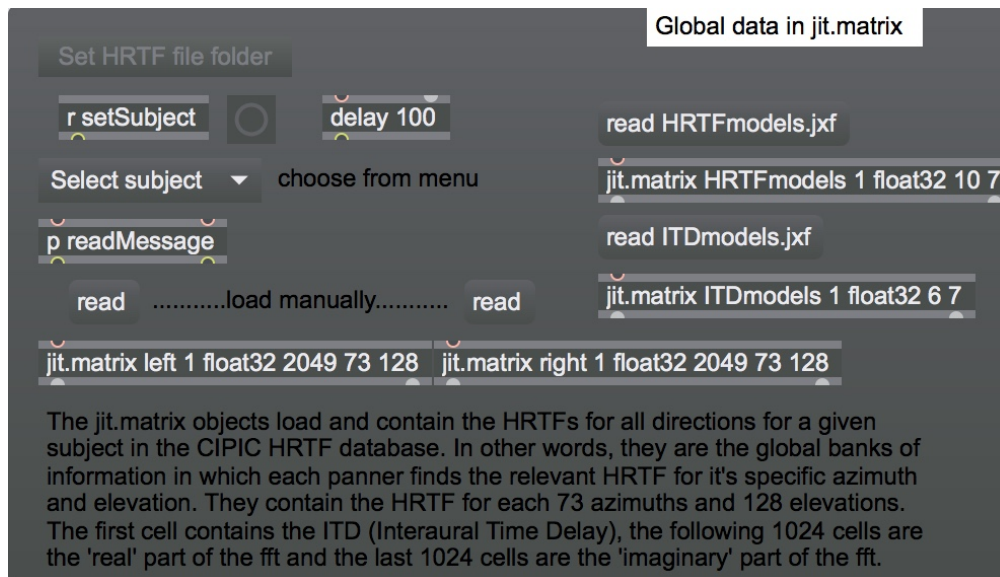


Figure 3.1: Loader

Loader reads pre-made JXF files into jitter matrices.

Set HRTF file folder and **Select subject** allows selection of subject database profiles by subject number. It loads existing HRTF into two matrices called left and right each for left ear and right ear.

Read HRTFmodels.jxf loads FFT model numbers into a 10 by 7 matrix, which 10 is the number of subjects and 7 is the 7 directions for model matching. FFT model numbers are manually dumped out, and mapped only by left ear FFTs.

Read ITDmodels.jxf loads ITD models into a 6 by 7 matrix, which 6 is the parameters for the model (left ITD left ear, front ITD left ear, right ITD left ear, left ITD right ear, front ITD right ear and right ITD right ear), and 7 is the 7 directions. ITD models are also manually analyzed and dumped out.

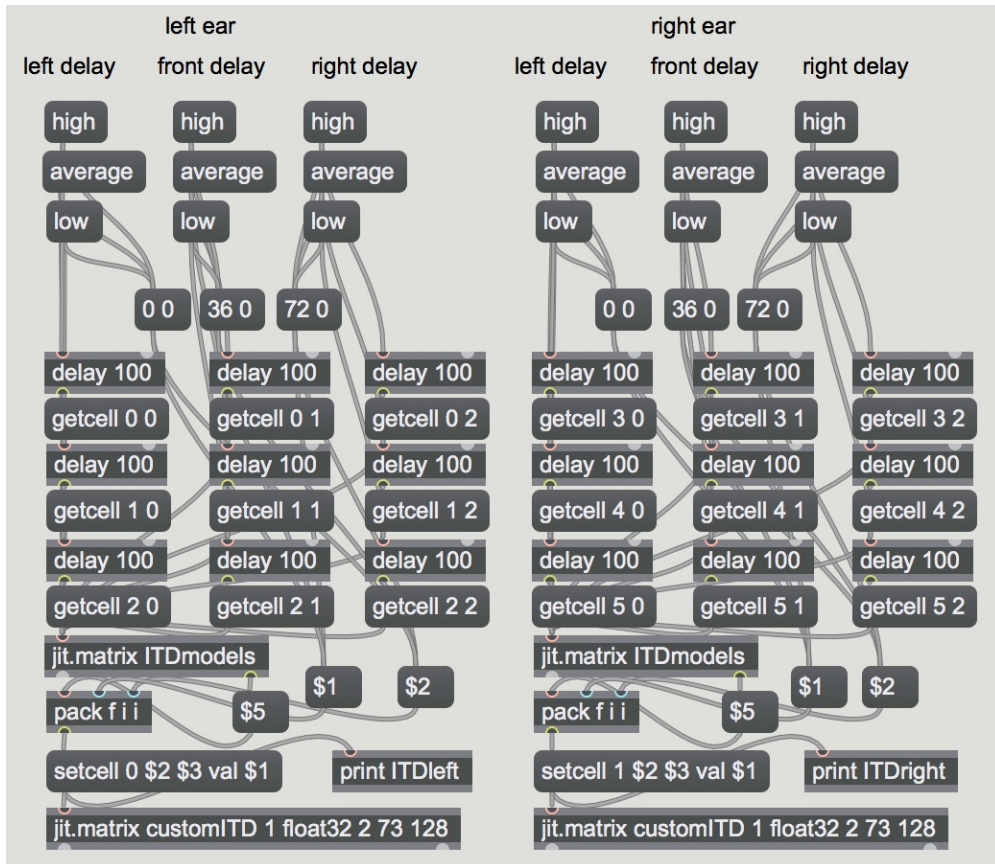


Figure 3.2: ITD matcher

ITD matcher takes high, average and low ITD values for 3 directions for each ear from the ITDmodels matrix, let users experiment with them, and store their picks into the customITD matrix.

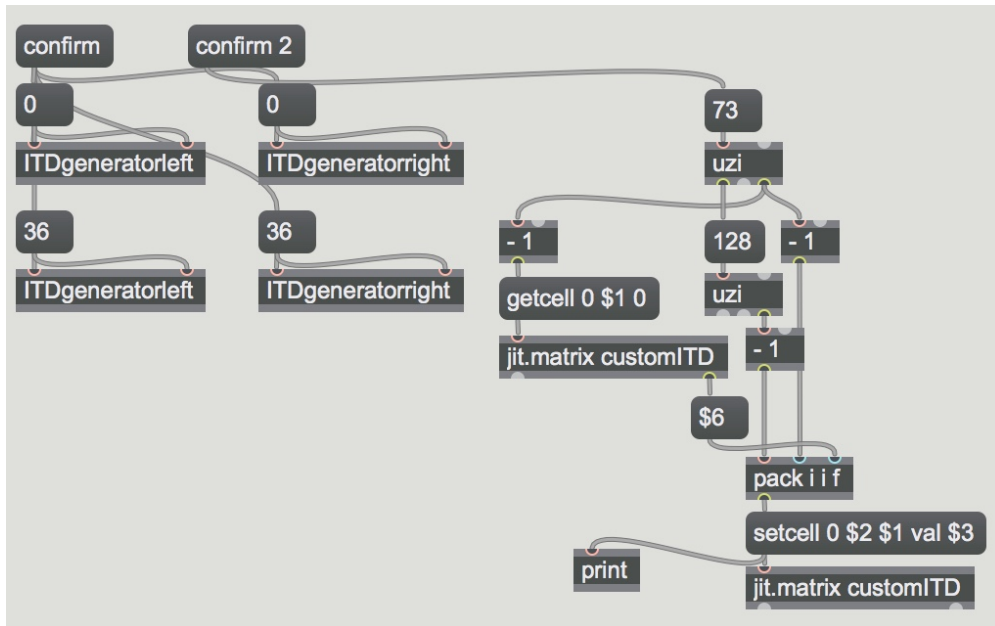


Figure 3.3: ITD confirmation

Confirm estimates ITDs of azimuth 1 to 35 (and 37 to 71) by making a linear function between the two.

Confirm 2 fill all ITDs of the same azimuth same value.

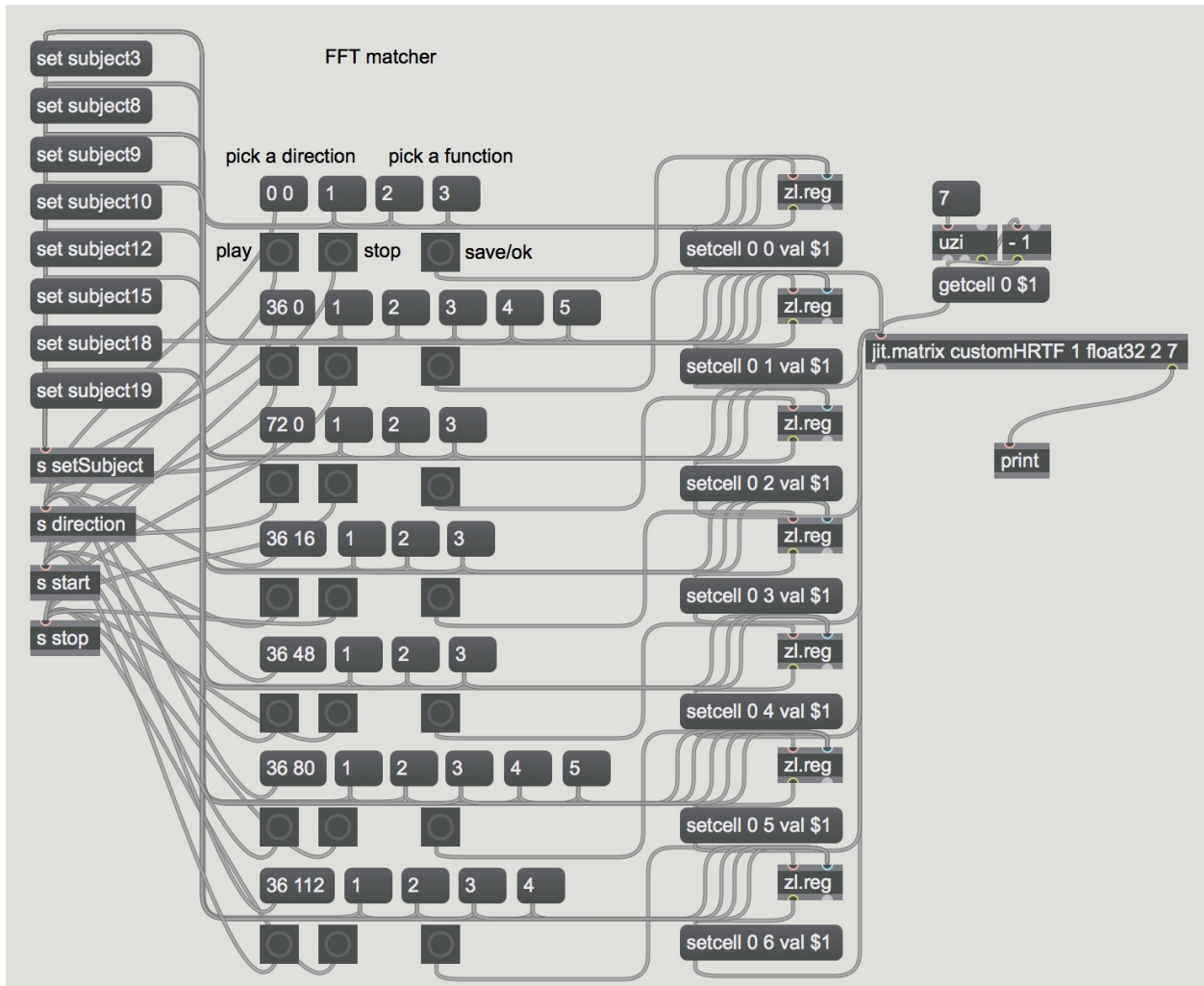


Figure 3.4: FFT matcher

FFT matcher let users test models with stimulus sound for 7 directions. Each model is represented by a number button at the right side of direction buttons. When a model button is clicked, it will load the subject profile with that model and hold the model number in `zl.reg` in case the user is not yet satisfied. When they are done they can release `zl.reg` and save the model number in the `customHRTF` matrix.

Model buttons and set subject buttons are manually linked and created. They are only effective for this database.

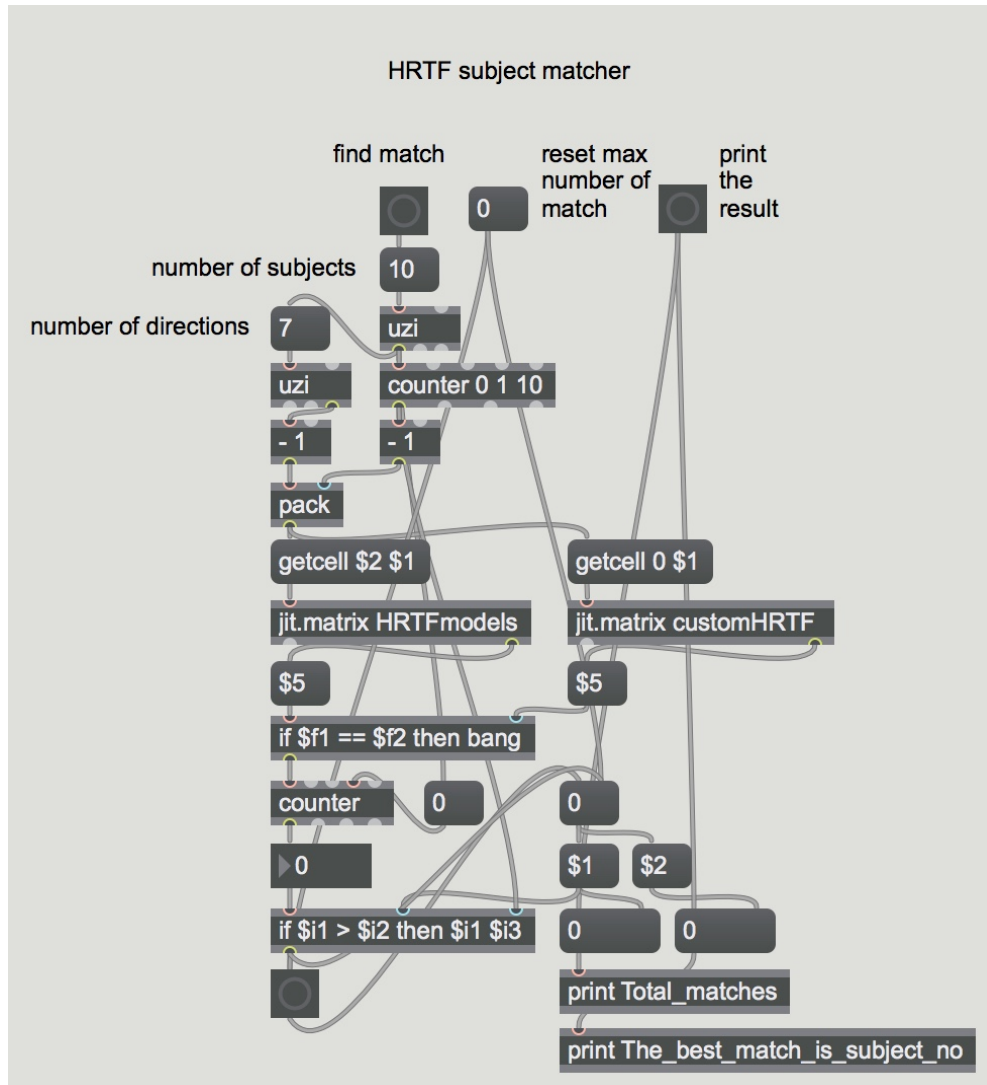


Figure 3.5: HRTF subject matcher

HRTF subject matcher takes model numbers out of customHRTF and compares it with each subject's model numbers in HRTFmodels. The best matching subject will be printed on the console.

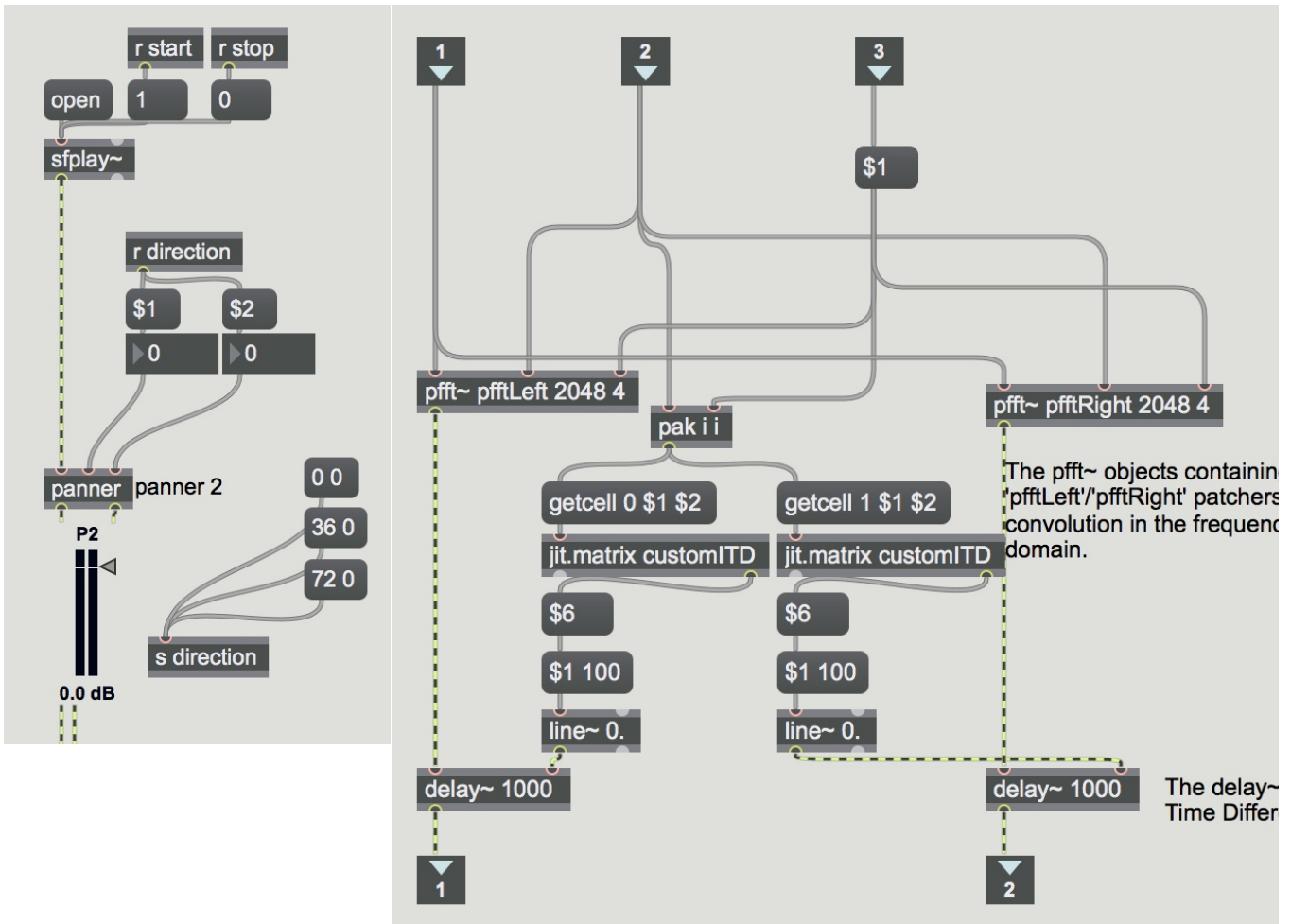


Figure 3.6: Audio player

Audio player uses `customITD` for delay and performs convolution on audio using FFTs from left and right matrices.

The result of the first round of testing is 2232152 for the FFT model numbers and average, low, average for left ear ITD and average high, average for right ear ITD. Therefore the best match is subject 8, which has 4 matches out of 7 total. During this test, it shows that changing ITD is not changing much of horizontal direction. Moreover, generated HRTF is not the most accurate while comparing the preset spatial location of a stimulus to where the test subject, whom hear the stimulus, thought it is coming from. The subject can get the elevate right every time, but always get the azimuth too right. This is probably because the effect of FFTs in frequency domain and equalization is overwhelmingly greater than which of delays. If FFTs favor higher right ear amplitude, our ears might not notice the subtle difference caused by balancing delays.

The result of the second round of testing is 3232152 for the FFT and the same ITD as the first round. When I specifically compared model no.3 and no.2 for 0/0, I found it difficult to choose between the two as they sound pretty much the same. While for that specific location it is fine if a user finds two models are indistinguishable, but the HRTF subject matcher may return an unreliable result because of this. A less accurate result can have the same number of matches as another more accurate result, simply because the user chooses a very similar model over another. Think of a scenario which other than 0/0, subject 1's profile has 1 more match than subject 2's, and for 0/0 they have indistinguishable FFT models. It should be obvious that subject 1 is the better result. However, if a subject picks subject 2's model at 0/0, subject 1's profile will have the same number of matches as subject 2's, so there is a 50% chance for HRTF subject matcher to return subject 2's HRTF, even though it is the inferior result.

4. Conclusion and Discussion

Model based matching is a simple and fast method to acquire individualized HRTFs.

It does not need any additional equipment other than a regular headphone. For two of my tests, it takes less than 10 minutes for each to complete. The prototype is not the easiest to use, but is far better than most of other methods, and can be vastly improved by a more carefully designed user interface.

However, being a very minimalistic approach, it has some major flaws.

Matching FFTs with the database means the acquired FFTs will almost never be perfectly accurate. Even if every model matches, subtle differences between standardized models and each person's FFTs should not be neglected. The only case for matches to be perfect is that the users' profiles already exist in the database, but then it doesn't make sense for them to use calibration.

It also means this method requires a large database for the match to be as accurate as possible, and collecting these data is a long term and expensive task. Meanwhile, to deploy this method in a business/consumer scale, it requires complicated algorithms to map each FFT to a model, hence put a heavy load on developers.

To improve on this, larger databases and model matching with more directions can reduce errors.

The more parameters (directions), the more accurate a match will be. The more existing profiles, the better chance for the user to find a perfectly matching profile. But as mentioned, larger database requires work and finance and will generate more models, user will find a hard time to

try everything out and eventually pick the best one. Using more directions will increase the time the calibration takes, which might diminish the main advantage of this method.

The calibration process is greatly dependent on user's sensation, sensitivity of hearing, or even feeling. Not only the problem mentioned in testing exists, if the user can't tell the difference between multiple models at multiple directions, then it might try to match some profile that simply won't exist. The amount of human errors varies greatly between individuals, and will result in inconsistent experiences.

Different headphones may need different sets of calibration. For the two tests I used two different headphones, and it can be a contributing factor for me to prefer model 3 over 2 in the second round of testing. According to *The role of individualized headphone calibration for the generation of high fidelity virtual auditory space*¹⁰, the use of nonindividualized HpTFs creates considerable distortions in the 4-10 kHz range, meaning that the headphone transfer function (HpTF) of each headphone is a part of HRTF and should also be calibrated. Using a different headphone can reduce the effectiveness of an accurate calibration, but that is an issue every calibration method is facing.

Future researches on this topic should focus on testing. To improve this specific research, more sets of tests should be done by different subjects with more standardized testing methodologies. Rather than simply letting subjects tell what they are hearing, place microphone in subjects' ears

¹⁰ Bibliography no.7

to calculate the actual error for the calibration. Test stimulus (audio sound for calibration) should be carefully chosen and explore on whether it will affect accuracy of the calibration. Left ear FFTs and right ear FFTs should be modeled separately instead of only using left ear FFTs.

5. Bibliography

1. Andersen, Jakob Hougaard. "Tool: FFT-based Binaural Panner." *Cycling '74*. N.p., 12 Dec. 2011. Web. 02 June 2017. <<https://cycling74.com/tools/fft-based-binaural-panner/#.WLdmRCMrJFQ>>.
2. Blauert, Jens. *The Technology of Binaural Listening*. Berlin: Springer Berlin, 2014.
3. Gilkey, Robert, and Timothy R. Anderson. *Binaural and Spatial Hearing in Real and Virtual Environments*. Place of Publication Not Identified: Psychology, 2015.
4. Romigh, Griffin D. "Individualized Head-Related Transfer Functions: Efficient Modeling and Estimation from Small Sets of Spatial Samples." *Cs.cmu.edu*. Carnegie Mellon University, 1 May 2012. Web. 21 June 2017. <<http://www.cs.cmu.edu/afs/cs/user/robust/www/Thesis/RomighThesis.pdf>>.
5. Xu, Song, Zhizhong Li, and Gavriel Salvendy. "Improved Method to Individualize Head-related Transfer Function Using Anthropometric Measurements." *Jstage.jst.go.jp*. The Acoustical Society of Japan, 25 Mar. 2008. Web. 3 June 2017. <https://www.jstage.jst.go.jp/article/ast/29/6/29_6_388/_pdf>.
6. Ghorbal, Slim, Renaud Séguier, and Xavier Bonjour. "Process of HRTF Individualization by 3D Statistical Ear Model." *Audio Engineering Society* (n.d.): n. pag. [Http://pro.3dsoundlabs.com/](http://pro.3dsoundlabs.com/). Audio Engineering Society, 29 Sept. 2016. Web. 1 July 2017. <http://pro.3dsoundlabs.com/wp-content/uploads/aes_3d_sound_labs_hrtf_personalization_brief.pdf>.
7. Pralong, Danièle, and Simon Carlile. "The Role of Individualized Headphone Calibration for the Generation of High Fidelity Virtual Auditory Space." *The Journal of the Acoustical Society of America* 100.6 (1996): 3785-793. Web.

6. Appendixes

A. ITD Raw Data

Left Ear ITD

Azimuth, Elevation	Subject 3 (dB)	Subject 8 (dB)	Subject 9 (dB)	Subject 10 (dB)	Subject 11 (dB)
0, 0	6.500000	8.000000	9.000000	6.625000	8.000000
36, 0	21.875000	24.500000	22.000000	19.875000	27.125000
72, 0	36.875000	34.000000	39.250000	32.875000	35.250000
36, 16	22.125000	24.750000	22.125000	20.000000	27.375000
36, 48	22.125000	24.750000	22.125000	20.625000	28.125000
36, 80	22.625000	24.750000	22.125000	20.625000	28.000000
36, 112	22.292318	24.912445	21.862518	19.928938	27.151970
Azimuth, Elevation	Subject 12 (dB)	Subject 15 (dB)	Subject 17 (dB)	Subject 18 (dB)	Subject 19 (dB)
0, 0	8.000000	6.750000	7.500000	7.625000	8.000000
36, 0	18.500000	16.250000	31.250000	21.875000	25.500000
72, 0	36.000000	35.000000	38.000000	31.875000	37.750000
36, 16	19.125000	19.625000	31.625000	22.500000	25.875000
36, 48	19.875000	20.625000	32.750000	23.250000	26.250000
36, 80	18.750000	20.625000	32.500000	23.000000	25.625000
36, 112	18.195990	17.565203	31.303938	22.073013	25.225037

Right Ear ITD

Azimuth, Elevation	Subject 3 (dB)	Subject 8 (dB)	Subject 9 (dB)	Subject 10 (dB)	Subject 11 (dB)
0, 0	36.875000	36.625000	39.250000	32.125000	35.875000
36, 0	22.000000	23.875000	20.875000	19.750000	26.500000
72, 0	6.375000	7.250000	6.875000	0.000000	7.250000
36, 16	22.375000	24.125000	20.875000	20.625000	26.250000
36, 48	22.875000	25.375000	19.500000	22.125000	27.125000
36, 80	23.000000	26.250000	18.500000	20.500000	28.500000
36, 112	22.251125	25.619844	19.045588	19.644709	27.160976
Azimuth, Elevation	Subject 12	Subject 15	Subject 17	Subject 18	Subject 19
0, 0	34.375000	34.875000	36.750000	33.000000	37.500000
36, 0	18.250000	16.750000	29.875000	21.250000	25.750000
72, 0	8.500000	6.500000	8.375000	6.250000	8.500000
36, 16	21.375000	21.625000	31.000000	22.000000	26.000000
36, 48	22.125000	21.875000	33.625000	24.750000	26.375000
36, 80	13.500000	19.000000	34.000000	24.000000	25.875000
36, 112	16.349331	17.285656	30.423418	21.656155	25.750000

B. FFT Raw Data

See FFT Raw Data in Simple HRTF Individualization.

C. ITD Models Chart

Right ear

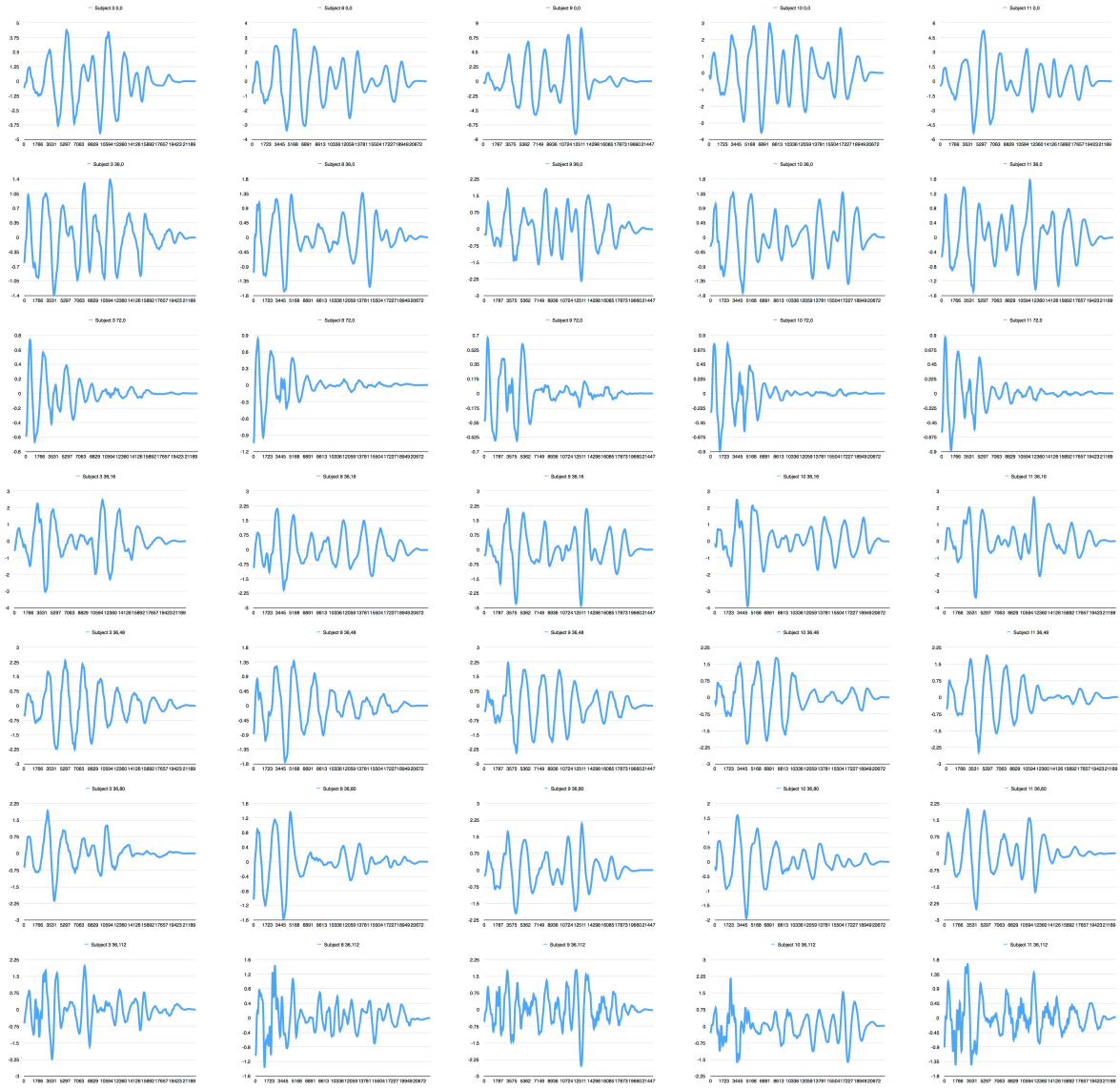
	High	Average	Low
0, 0	39.25	35.725	32.125
36, 0	29.875	22.4875	16.75
72, 0	8.5	6.5875	6.25
36, 16	31	23.625	20.625
36, 48	33.625	24.575	19.5
36, 80	34	23.3125	13.5
36, 112	30.423418	22.5186802	16.349331

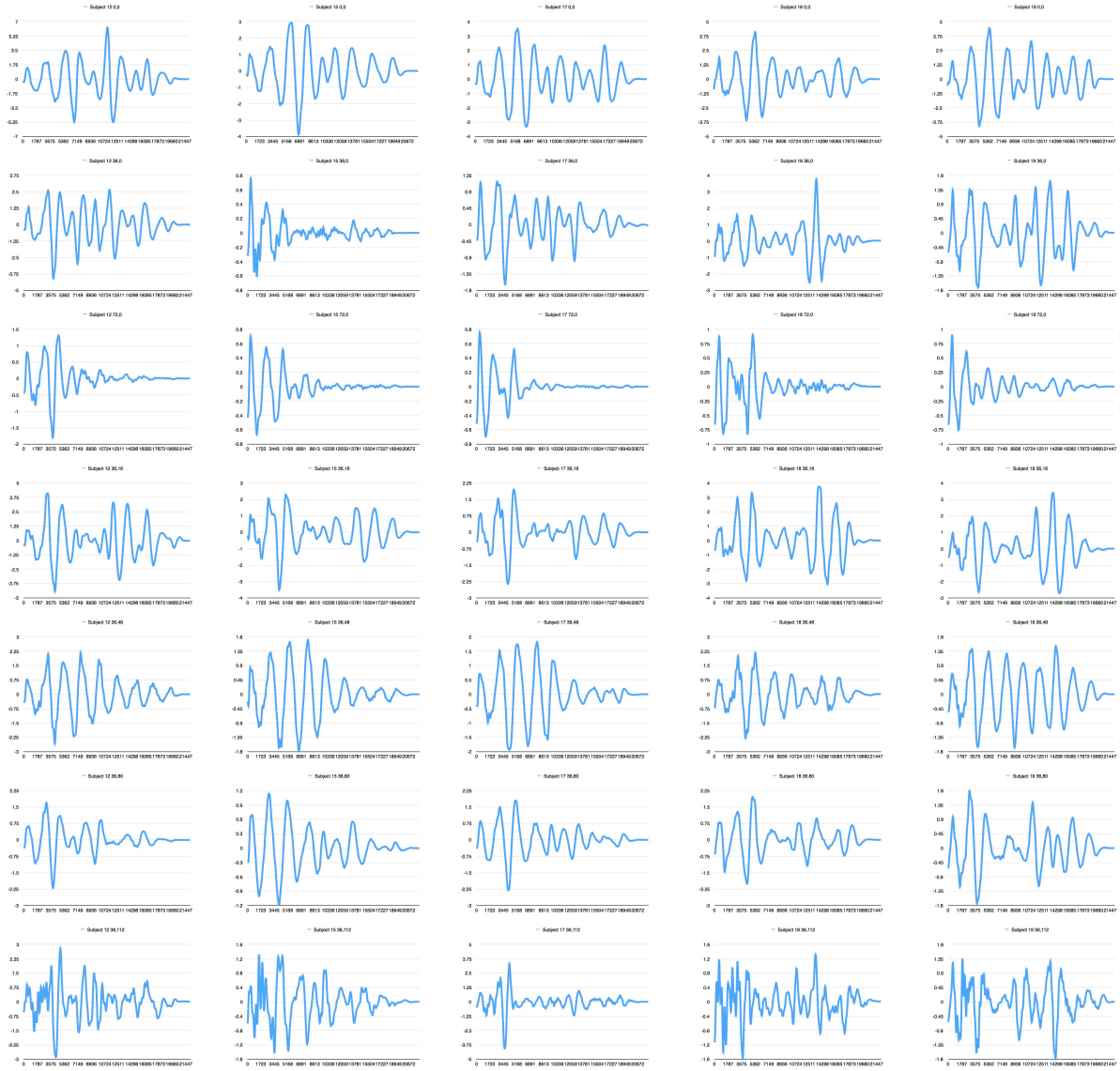
Left ear

	High	Average	Low
0, 0	9	7.6	6.5
36, 0	31.25	22.875	16.25
72, 0	39.25	35.6875	31.875
36, 16	31.625	23.5125	19.125
36, 48	32.75	24.05	19.875
36, 80	32.5	23.8625	18.75
36, 112	31.303938	23.051137	17.565203

D. FFT Models Chart

FFT graphs for left ear





Model Numbers

Subject\Direction	0, 0	36, 0	72, 0	36,16	36, 48	36, 80	36, 112
3	2	1	1	1	1	1	1
8	1	2	1	2	1	2	2
9	3	3	2	3	1	3	3
10	1	2	1	2	2	1	4
11	1	1	1	1	2	1	2
12	2	3	3	1	1	1	1
15	1	4	1	2	2	4	2
17	1	3	1	2	2	1	4
18	1	5	2	3	1	1	2
19	1	1	1	3	3	5	3

E. Main Patch Instructions

Also in readme.txt in Simple HRTF Individualization.

The main file is Simple HRTF Individualization.maxpat

About elevation and azimuth

Azimuth is from 0 (direct left) to 72 (direct right), so the center is 36. Elevation is from 0 (45 degrees downward) to 127 (46 degrees downward), so front is 16, top is 48, back is 80 and bottom is 112.

Preparation

Click on "Set HRTF file folder" at the top of the grey panel, and set the folder to "All HRTF matrix files". This loads the HRTF files of all subjects. You can download it from <http://jakobhandersen.dk/projects/fft-based-binaural-panner/download/>

Click on "open" button on top of sfplay~ on the right, and select the stimulus sound file (in this case, cowbell.aiff). This loads the stimulus for testing

Click on every "average" button in the ITD matcher (purple panel area). This initializes the environment with average ITDs across the board

Click "confirm" and then "confirm 2" below ITD matcher

First start with the FFT matcher (blue panel area)

Click on a direction button (i.e.: "0 0", "36 80"). The first number is azimuth, second is elevation

Pick a function model at the right side of the direction button (number buttons)

Click on the play button below to test it out

If what you hear does not match the direction, pick another function model and play

Repeat 2- 4 until you find the best match

Click save/ok button right to the play button

Repeat 1 - 6 until all 7 directions are calibrated

You can always check out your calibrated result by clicking the "7" button at the top right corner, result will be printed on the console

Then use the HRTF subject matcher (orange panel area)

Click on "find match" button

Click on "print the result" button

The best match subject number and the number of matches will be printed on the console

Notice: Subject number is not the name of the subject. i.e.: Subject no. 3 doesn't mean subject3, it means the third subject which is subject9

You can now load the subject to the best match subject on the grey panel by clicking "choose from menu" umenu

Optional: Anytime you feel unbalanced delay in your ears, use the ITD matcher (purple panel area)

Each ear's left/front/right delay refers to each ear's delay in ms of sound coming from direct left/front/right.

For example, if you are listening from direct front, but you think it is a little to the right, it means either your left ear's front delay is too high, or your right ear's front delay is too low, or both. You can calibrate it by setting left ear's front delay to low, and/or right ear's front delay to high.

Another example, if you are listening from direct right, but you think it is a bit to the left, it means either your left ear's right delay is too low, or your right ear's right delay is too high, or both. You can calibrate it by setting left ear's right delay to high, and/or right ear's right delay to low.

You can always test out the effect by clicking on any direction button on the left of the ITD matcher and "1" on top of sfplay~ to play

After you are done with ITD matcher, always click "confirm" and then "confirm 2" button below to apply it to all directions.

Details about this patch

ITD matcher:

The high/average/low is calculated based on the ITDs of the first ten subjects in the database.

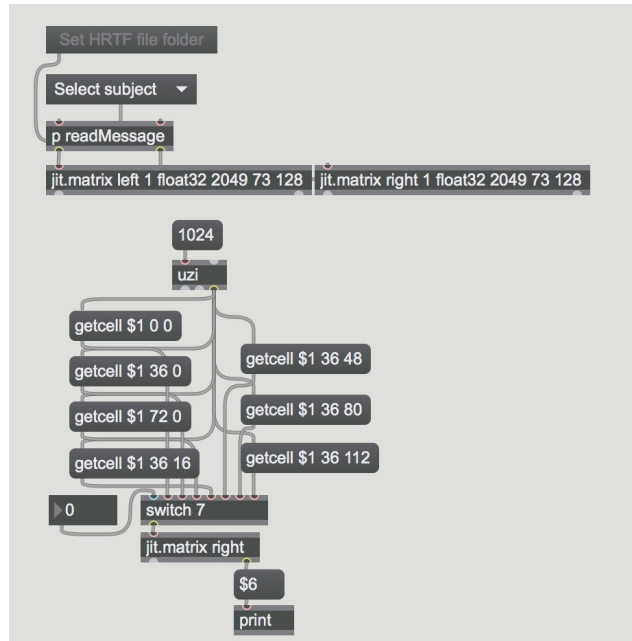
The ITD matcher assumes ITD has no relationship with elevation. So ITDs of the same azimuth will always be the same no matter how elevation is changed. "Confirm" estimates ITDs between azimuth 0 and 36 (and 36 to 72) by making a linear function between the two.

"Confirm 2" fill all ITDs of the same azimuth same value.

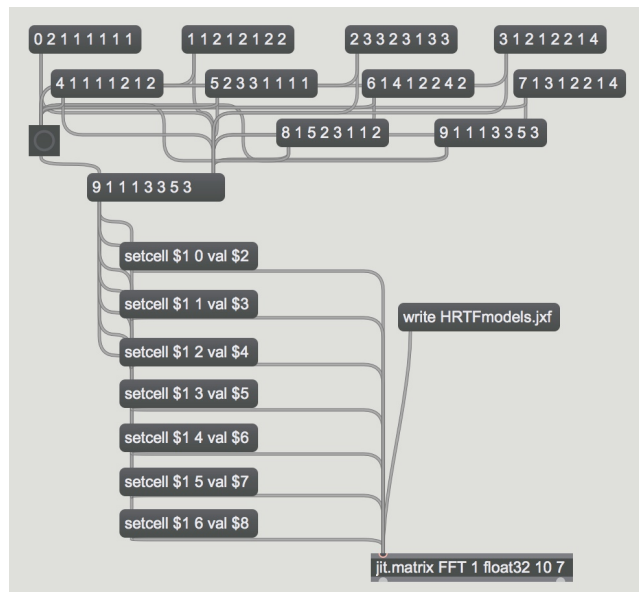
FFT macher:

The FFT matcher does not create a custom HRTF. It only helps find the best match of the database.

F. Supportive Max Patches



All FFT dump



FFT models dump



ITD models dump