


Article

# Monaural Sound Localization Based on Reflective Structure and Homomorphic Deconvolution

Yeonseok Park <sup>1</sup>, Anthony Choi <sup>2</sup> and Keonwook Kim <sup>1,\*</sup> 

<sup>1</sup> Division of Electronics & Electrical Engineering, Dongguk University-Seoul, Seoul 04620, Korea; dustjrdk@dongguk.edu

<sup>2</sup> Department of Electrical & Computer Engineering, Mercer University, 1501 Mercer University Drive, Macon, GA 31207, USA; choi\_ta@mercer.edu

\* Correspondence: kwkim@dongguk.edu; Tel.: +82-2-2260-3334

Received: 16 August 2017; Accepted: 19 September 2017; Published: 23 September 2017

**Abstract:** The asymmetric structure around the receiver provides a particular time delay for the specific incoming propagation. This paper designs a monaural sound localization system based on the reflective structure around the microphone. The reflective plates are placed to present the direction-wise time delay, which is naturally processed by convolutional operation with a sound source. The received signal is separated for estimating the dominant time delay by using homomorphic deconvolution, which utilizes the real cepstrum and inverse cepstrum sequentially to derive the propagation response's autocorrelation. Once the localization system accurately estimates the information, the time delay model computes the corresponding reflection for localization. Because of the structure limitation, two stages of the localization process perform the estimation procedure as range and angle. The software toolchain from propagation physics and algorithm simulation realizes the optimal 3D-printed structure. The acoustic experiments in the anechoic chamber denote that 79.0% of the study range data from the isotropic signal is properly detected by the response value, and 87.5% of the specific direction data from the study range signal is properly estimated by the response time. The product of both rates shows the overall hit rate to be 69.1%.

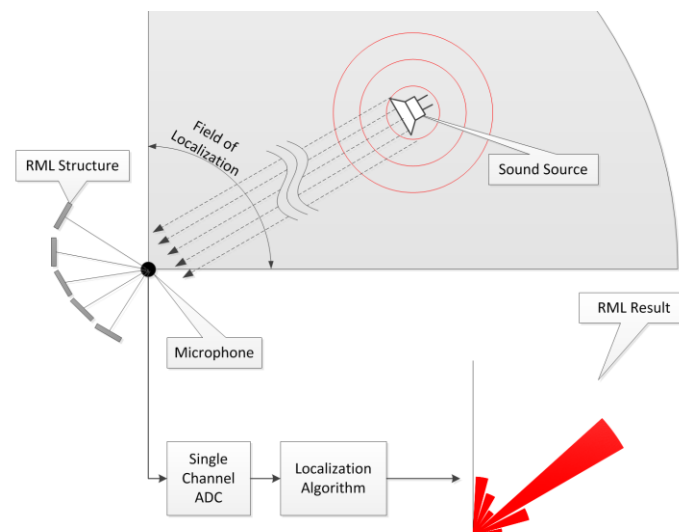
**Keywords:** sound localization; angle of arrival; monaural localization; acoustic reflection; time delay; homomorphic deconvolution; cepstrum; single microphone; far-field; 3D printer

## 1. Introduction

Humans localize sound sources in three-dimensional (3D) space by using the binaural correlation and structure profile. The unique shapes of the head and pinna modify the propagated sound properties in the magnitude, phase and spectrum. The horizontal plane over the human head provides the binaural sound localization environment due to both ear positions. In an asymmetric manner, the direct and indirect distances from the source to the ears deliver clues for estimating the angle of arrival (AoA) in the plane. However, the median plane provides the symmetric space to the ears; therefore, the vertical position variation cannot be recognized by the simple correlation information between the ears. The pinna shape presents the acoustic variation of multiple reflections in the structure to identify the AoA over the median plane. Numerous papers have described the role of the pinna for median plane sound localization comprehensively [1–11]. Especially, Batteau [11] suggested the Laplace transform-based parametric model to understand the acoustic propagation over the pinna structure for extension to reverberation and other facets of human hearing. Further exploration of the pinna and head-related transfer function can be found by the recent papers [12,13]. Sound localization based on a single receiver, known as monaural localization (ML), is inspired by median-plane localization.

ML cannot be realized in the isotropic and far-field condition. The structure around the receiver should be placed for particular modification over the propagation path similar to the pinna structure. The receiver identifies the variation to estimate the AoA in the ML system. The asymmetric structure maximizes the acoustic property difference for individual localization angles. Numerous investigations have been conducted for a structure-related localization system and are summarized below. The designed analog circuit estimated the time difference between the direct and indirect propagation for ML [14]. The binaural system was extended for an extra dimension by using the pinna-like reflector and corresponding estimation algorithm [15–19]. The various structures around the microphone were explored to improve the directivity pattern of the head-related transfer function [20]. Based on the hidden Markov model and signal moment, the machine learning is proposed to approach the monaural sound localization with an artificial pinna [21]. The actively deformable pinna system is proposed and analyzed for sound localization in the application of a mobile robot [22]. From the artificial pinna, the sound source elevation is estimated by using the propagation transfer function and the neural network classifier [23]. The characteristics of indoor speech propagation were utilized for non-structural ML within a limited situation [24–26]. The parabolic structure with cepstral speech parameters was explored for position-dependent indoor ML [27]. A hybrid ML system based on the audio-visual method with cepstral parameters was proposed by Friedland et al. [28].

This paper proposes a novel reflective monaural localization (RML) system for the far-field condition with the structure around the receiver. The structure provides the distinctive reflection times for individual angles to implant the direction information over the sound propagation. The received single-channel information is decoded to estimate the induced time delay by using homomorphic deconvolution (HD). The estimated delay is linearly mapped to the corresponding AoA in the final stage. Figure 1 shows the overall functional diagram. The incoming signal is discretized by the analog-to-digital converter (ADC), and the outcomes in terms of the likelihood of each direction are computed by the discrete process of the RML algorithm. A higher value represents an elevated possibility of an AoA. The designed structure should have at least one open face to the receiver and source to generate the proper reflections; therefore, the field of localization is limited to a certain range.



**Figure 1.** Functional diagram of the reflective monaural localization (RML) system.

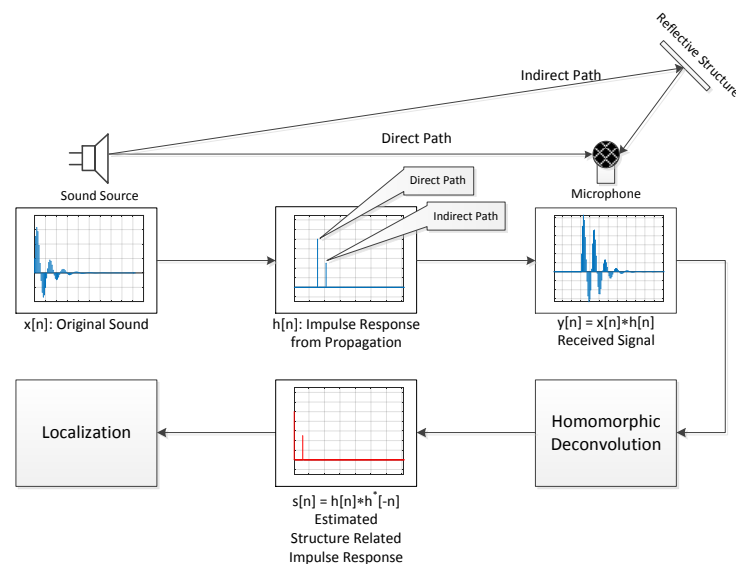
Along with the structure design, the time delay produced by the reflection is computed by the HD algorithm, which is established through the homomorphic system [29,30]. The HD removes or alters one of the components of a convolutional operation by using the homomorphic property. The HD is used for a variety of areas for single- and multi-dimensional signals, such as image, audio and seismic [31–36]. This paper realizes the HD based on the real cepstrum [37–39], lifting and inverse

cepstrum procedure to estimate the propagation information, which delivers the time difference between the signal arrivals at the receiver. The source signal is eliminated by the lifting procedure, and the derived propagation information corresponds to the incoming AoA. Note that lifting is the term used by the cepstral analysis for filtering, and the cepstrum indicates the real cepstrum in this paper.

This paper accomplishes the work proposed by the authors' previous ML publications. The fundamental frequencies induced by the asymmetric horizontal pyramidal horns were arranged for the far-field ML system by utilizing cepstral parameters [40]. The small-profile near-field ML system was realized by the asymmetric vertical cylindrical pipes around a single microphone [41]. The RML system of this paper improves the overall size from the pyramidal horn structure method, as well as the working range from the cylindrical pipe structure method. Other localization works on the subject by the authors are also related to and expanded during the research, such as azimuthal movement detection based on binaural architecture [42] and a target localization algorithm over a distributed acoustic sensor network [43]. Observe that the RML experiments are performed and evaluated within an identical anechoic chamber [44] to that used in the previous works.

## 2. Methodology

Multi-path propagation can be described as a linear time-invariant (LTI) system with convolutional operation. The direct and indirect arrivals at the receiver are represented by the corresponding temporal Kronecker delta functions (or delta functions) at the impulse response. The received output is developed by the convolution output between the source signal and the impulse response. Once the impulse response is derived from the received signal, the propagation path can be calculated by the response. The RML system is initiated from the fundamental idea that the reflective structure is placed and the deconvolution operation is performed to find the signal's AoA. The careful structural layout provides the linear variation of the reflection time between the arrivals over the signal AoAs. The calculated impulse response contains arrival time information that can be translated into the AoA based on the reflection model. Figure 2 denotes the overall system architecture of the proposed RML system.



**Figure 2.** System architecture of the proposed RML system.

The source waveform  $x[n]$  propagates over the physical space characterized by the environment and structure. The impulse response  $h[n]$  of the space delivers the path and distance information of the propagation. The received information  $y[n]$  from the single receiver can be depicted by the convolution (shown as \* operator) between the waveform  $x[n]$  and impulse response  $h[n]$ . Note that

the impulse response  $h[n]$  cannot be measured directly without using the waveform  $x[n]$  information which is not available. The HD estimates  $h[n]$  in the form of the structure-related impulse response (SRIR)  $s[n]$ , which is the autocorrelation of  $h[n]$ . The SRIR  $s[n]$  represents the time difference between the sound arrivals due to the nature of the autocorrelation shown as convolution (\*) between  $h[n]$  and time reversed & conjugated  $h^*[-n]$ . The time information from the SRIR is translated into the AoA by applying the localization model that shows the relationship between the arrival angle and dominant reflection equivalently distinctive time difference. Note that the model is devised by the physical simulation and optimized by the acoustic experiments.

The equations below provide the HD computational procedure for the SRIR. The HD utilizes the real cepstrum, which applies the logarithm to the absolute discrete Fourier transform (DFT) outcome. If it is desired that the system obtains the original waveform or impulse response, the HD should employ the complex cepstrum that uses the logarithm for the raw DFT with extra complexity induced from phase unwrapping. The complex cepstrum preserves the phase information; therefore, no ambiguity can be experienced in the inverse procedure, such as maximum and minimum phase realization. The HD from the real cepstrum generates the autocorrelation of the impulse response, named the SRIR, because of the absolute operation on the DFT. The window function  $w[n]$  separates the SRIR from the received signal by utilizing the logarithm property. The index of the minimum value within the interesting range signifies the time difference between the first and second arrivals to the receiver. The corresponding time is computed by sampling the frequency parameter. The details of the HD derivations can be found in Appendix A. Note that the indexes for the time and cepstrum domains are equalized for convenience.

$$\begin{array}{l} \text{Received signal} \\ \rightarrow \end{array} y[n] = x[n] * h[n] \quad (1)$$

$$\begin{array}{l} \text{Real cepstrum} \\ \rightarrow \end{array} c_y[n] = \text{IDFT}\{\log|\text{DFT}\{y[n]\}|^2\} \quad (2)$$

$$\begin{array}{l} \text{Window} \\ \rightarrow \end{array} c_u[n] = c_y[n]w[n] \quad (3)$$

$$\begin{array}{l} \text{Inverse cepstrum} \\ \rightarrow \end{array} s[n] = \text{IDFT}\{e^{\text{DFT}\{c_u[n]\}}\} \approx h[n] * h^*[-n] \quad (4)$$

$$\begin{array}{l} \text{Find minimum location} \\ \rightarrow \end{array} r = \underset{n=\{n_{\min} \leq n \leq n_{\max}\}}{\text{argmin}} s[n] \quad (5)$$

$$\begin{array}{l} \text{Compute the reflection flight time} \\ \rightarrow \end{array} t_d = \frac{r}{f_s} \quad (6)$$

The equations below indicate the mathematical derivations for the HD algorithm on direct and indirect sound propagation.  $x[n]$  is the original waveform generated by the sound source, and  $x[n - r_1]$  and  $x[n - r_2]$  denote the first and second sound arrivals to the receiver, respectively. The second arrival is the reflected sound to the hard boundary; thus, the signal presents a phase reversal ( $\pi$  radian difference).  $-\alpha$  represents the phase reversal with the reflection level. The HD output shows that  $\delta[n - (r_2 - r_1)]$  demonstrates the time difference between the arrivals with a minimum value. Observe that  $\delta[n - (N - (r_2 - r_1))]$  is the shadow value created by the autocorrelation and DFT circular property. The HD outcome magnitude is independent of the signal amplitude and correlates to the reflection level  $\alpha$ .

Received:

$$y[n] = x[n - r_1] - \alpha x[n - r_2] \text{ where } r_1 < r_2 \text{ and } 0 < \alpha < 1 \quad (7)$$

Real cepstrum:

$$\begin{aligned} \begin{array}{l} \text{DFT} \\ \rightarrow \end{array} Y[k] &= X[k]e^{-j\frac{2\pi}{N}kr_1} - \alpha X[k]e^{-j\frac{2\pi}{N}kr_2} \\ &= X[k]\left(e^{-j\frac{2\pi}{N}kr_1} - \alpha e^{-j\frac{2\pi}{N}kr_2}\right) \end{aligned} \quad (8)$$



$$\begin{aligned}
& \xrightarrow{\text{Squared Magnitude}} |Y[k]|^2 = Y[k]Y^*[k] \\
& = |X[k]|^2 \left( e^{-j\frac{2\pi}{N}kr_1} - \alpha e^{-j\frac{2\pi}{N}kr_2} \right) \left( e^{j\frac{2\pi}{N}kr_1} - \alpha e^{j\frac{2\pi}{N}kr_2} \right) \\
& = |X[k]|^2 \left( 1 + \alpha^2 - 2\alpha \cos\left(\frac{2\pi}{N}k(r_2 - r_1)\right) \right)
\end{aligned} \tag{9}$$

$$\begin{aligned}
& \xrightarrow{\text{Logarithm}} \log|Y[k]|^2 \\
& = \log|X[k]|^2 + \log\left(1 + \alpha^2 - 2\alpha \cos\left(\frac{2\pi}{N}k(r_2 - r_1)\right)\right)
\end{aligned} \tag{10}$$

$$\begin{aligned}
& \xrightarrow{\text{IDFT}} \text{IDFT}\{\log|Y[k]|^2\} \\
& = \text{IDFT}\{\log|X[k]|^2\} + \text{IDFT}\left\{\log\left(1 + \alpha^2 - 2\alpha \cos\left(\frac{2\pi}{N}k(r_2 - r_1)\right)\right)\right\}
\end{aligned} \tag{11}$$

Window:

$$\begin{aligned}
& \xrightarrow{\text{Window}} c_u[n] = \text{IDFT}\{\log|Y[k]|^2\}w[n] \\
& \approx \text{IDFT}\left\{\log\left(1 + \alpha^2 - 2\alpha \cos\left(\frac{2\pi}{N}k(r_2 - r_1)\right)\right)\right\}
\end{aligned} \tag{12}$$

Inverse cepstrum:

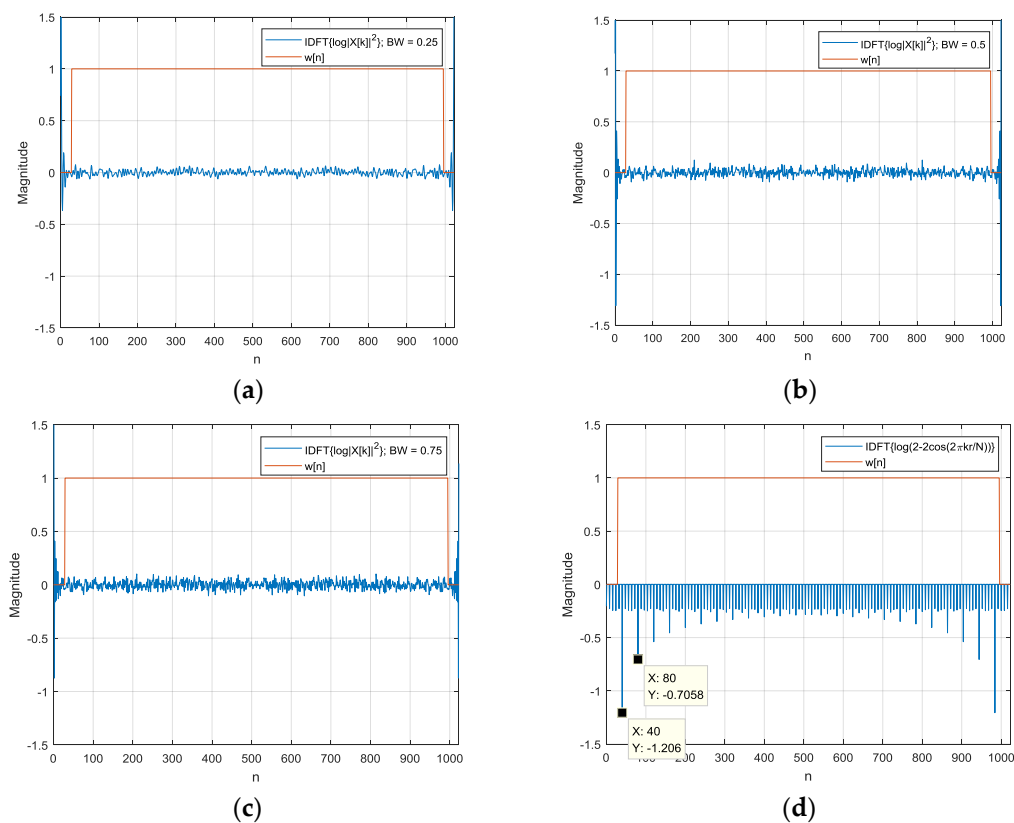
$$\begin{aligned}
& \xrightarrow{\text{DFT}} \text{DFT}\{c_u[n]\} \approx \log\left(1 + \alpha^2 - 2\alpha \cos\left(\frac{2\pi}{N}k(r_2 - r_1)\right)\right)
\end{aligned} \tag{13}$$

$$\begin{aligned}
& \xrightarrow{\text{Exponential}} e^{\text{DFT}\{c_u[n]\}} \approx 1 + \alpha^2 - 2\alpha \cos\left(\frac{2\pi}{N}k(r_2 - r_1)\right)
\end{aligned} \tag{14}$$

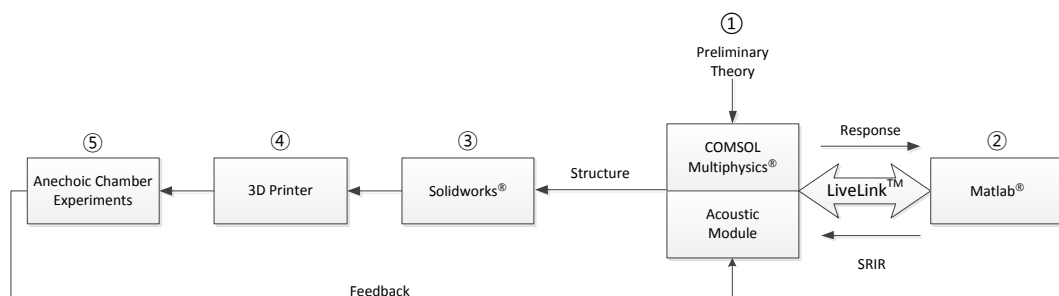
$$\begin{aligned}
& \xrightarrow{\text{IDFT}} \text{IDFT}\left\{e^{\text{DFT}\{c_u[n]\}}\right\} \\
& \approx \left(1 + \alpha^2\right)\delta[n] - \alpha\delta[n - (r_2 - r_1)] - \alpha\delta[n - (N - (r_2 - r_1))]
\end{aligned} \tag{15}$$

$\text{IDFT}\{\log|Y[k]|^2\}$  is the linear combination of the signal stand delay components in the above equations. The window function  $w[n]$  separates the delay part from the inverse DFT operation. An example is shown in Figure 3 with  $\alpha = 1$  magnitude and  $(r_2 - r_1) = 40$  samples. The signal is generated by the white noise with a 10th-order Butterworth low-pass filter for 0.25, 0.5 and 0.75 normalized frequency bandwidths. Note that the 1 normalized frequency indicates half of the sampling frequency. For the increased bandwidth,  $\text{IDFT}\{\log|X[k]|^2\}$  exhibits the amplified values in the overall range; however, a significant portion of the energy is concentrated on the edges. In contrast, the delay part  $\log(2 - 2\cos(2\pi k(r_2 - r_1)/N))$  demonstrates the strong and damped values over the harmonics of the time difference. The delay part logarithm can be approximated by the Newton–Mercator series [45] for converging harmonics. The window  $w[n]$  including the interesting time delay can effectively extract the delay part for further processing. Note that the circular property of the real cepstrum requires the window in a circular manner. Figure A1 denotes the rest of the process for the HD and its outcome.

The structure of the RML consists of the combination of multiple plates to produce distinct time delays from individual directional signals. This paper comprehensively employs the procedure of modeling, simulation and experimentation to minimize design and experiment iterations. To create and evaluate the RML structure, Figure 4 presents the overall procedure, which follows an identical process to that of the previous study [41]. The initial structure is assessed by the COMSOL Multiphysics software to compute the signal propagation. MATLAB provides parametric variations and receives the temporal response to calculate the time delay from the HD algorithm. Procedures ① and ② in Figure 4 signify the simulation workflow to propose the structure shape. The candidate structure from the simulation is realized by the design program (SolidWorks) and a 3D printer for acoustic experiments in the anechoic chamber. The actual shape is placed in the anechoic chamber to analyze the received signal and examine the localization performance. Procedures ③, ④ and ⑤ in Figure 4 represent the experiment workflow for comprehensive structure feedback. Based on the acoustic performance, further modifications could be necessary for several iterations.



**Figure 3.** The signal portion  $\text{IDFT}\{\log|X[k]|^2\}$  for the (a) 0.25, (b) 0.5 and (c) 0.75 normalized frequency bandwidths and (d) the delay part  $\log(2 - 2 \cos(2\pi k(r_2 - r_1)/N))$  for 40 sample delays.



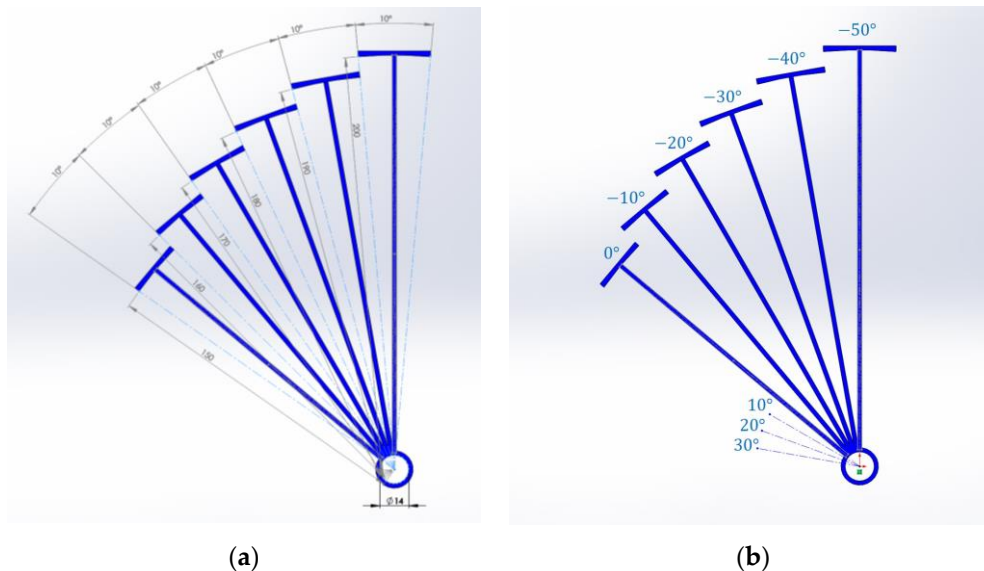
**Figure 4.** Workflow of the RML system development [41].

The acoustic experiments are executed and analyzed in an anechoic chamber that has been verified to exhibit partial conformance with ISO 3745 [46] for the 250 Hz–16 kHz one-third octave band in a free-field chamber and for the 1 kHz–16 kHz one-third octave band in a hemi-free-field chamber [44]. The RML system is evaluated with the free-field chamber mode, which contains fully-covered surfaces for all directions with acoustic wedges. Note that the RML structure indicates the physical structure, and the RML algorithm denotes the HD algorithm with the direction model from time delay to AoA. In addition, the RML system is the combination of the structure and the algorithm.

### 3. Structure Design and Simulation

The designed RML structure consists of six plates, which involves  $60^\circ$  physical coverage, as shown in Figure 5. The plate is the arc of the circle with an individual radius to the center. The arc shows a fixed  $10^\circ$  central angle and a distance range from 150 mm to 200 mm in every 10 mm variation. The arc center also includes a concentric circle with a 7 mm radius for the receiver microphone. The azimuthal

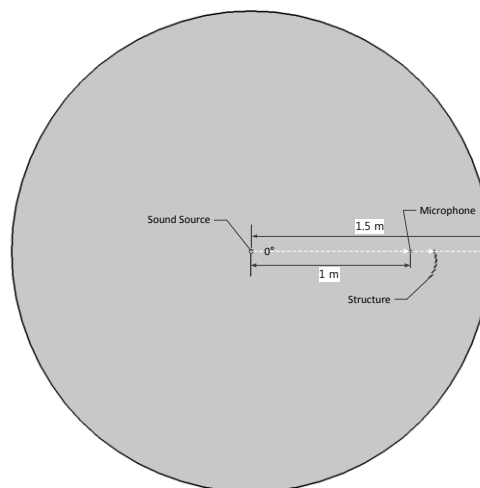
angle is measured over the RML structure plane from the leftmost arc center to the counterclockwise direction. Figure 5b demonstrates the corresponding angles for each arc. Note that the arc height is 100 mm, which is not illustrated in the figure below. To avoid acoustical distraction, the connection line from the arc to the center is developed as a low profile.



**Figure 5.** The designed RML structure for (a) dimensions in millimeters and (b) angles in degrees.

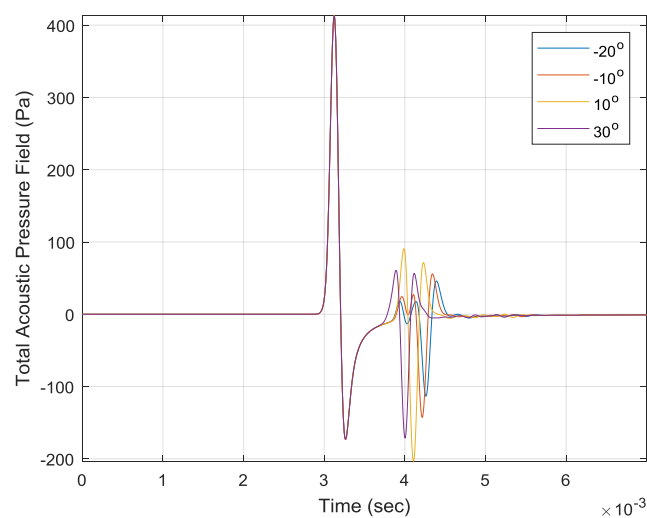
The proposed RML structure is evaluated by the COMSOL Multiphysics simulator for optimal parameters over multiple iterations. The simulation space is described by the two-dimensional (2D) space for a 1.5 m radius cylindrical free field as shown Figure 6. The principal interest of the analysis is acoustic propagation over the RML structure plane; thus, the 2D configuration provides sufficient information for parametric search. The circle boundary is established as a perfectly matched layer to represent an open and non-reflecting infinite domain for all wave types. The sound source is located at the circle center, and the receiver is placed 1 m away from the source in the eastbound direction for far-field provision. The RML structure is initiated from the microphone position, which is the 7 mm radius circle. The connection lines between the microphone and plates are not studied and shown in the simulation process. The rigid body of the structure is defined by the sound hard wall in the simulation for zero normal components of the velocity. The simulation space is filled with the air provided by the default simulator parameters.

The sound source excites the air medium through a point Gaussian pulse in terms of amplitude, frequency bandwidth and pulse peak time. The complete parameters for the COMSOL simulation are described in Table A1 (Appendix B). With a  $2 \text{ m}^2/\text{s}$  amplitude, 4 kHz bandwidth and  $2.5 \times 10^{-6} \text{ s}$  peak time, the microphone receives the signals for  $-20^\circ$ ,  $-10^\circ$ ,  $10^\circ$  and  $30^\circ$ , as shown in the figure below. The first sound arrivals are identical to all AoAs; however, the second arrivals denote a phase reversal from the reflection with various magnitudes and delays. The minimum value at the second arrival corresponds to the maximum value at the first arrival due to the  $\pi$  radian phase difference. The times of the second impact minimum illustrate the inversely-proportional relationship with AoAs in a consistent manner. Therefore, the precise measurement of the time difference between the arrivals can be used to estimate the signal AoA.

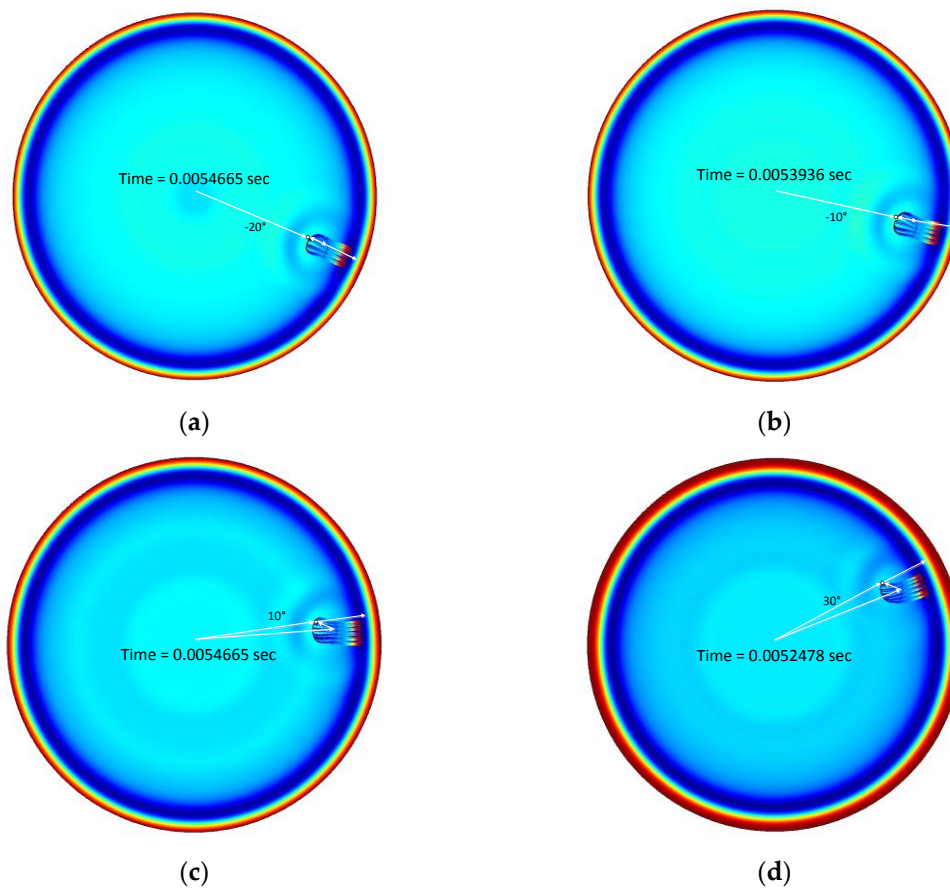


**Figure 6.** COMSOL simulation geometry for  $0^\circ$  angle of arrival (AoA). The RML structure is located in the air space (circle) with a perfectly absorbing boundary.

Because of the structure configuration shown in Figure 5b, the positive AoAs do not exhibit the normal reflection, which has a perpendicular incident and reflected angle to the plate. The propagation line does not connect the source, receiver and plate in the direct path; however, the reflection strength is relatively strong in Figure 7 for the  $10^\circ$  and  $30^\circ$  AoAs. The point Gaussian pulse collectively arrives in the direct path at approximately 3 ms; however, the reflected signals individually reach the receiver with a time difference between the 3.5 ms and 4.5 ms in Figure 7. Figure 8 denotes the surface plot for the total acoustic pressure field, the streamline for local acceleration and the white line for the dominant propagation path over  $-20^\circ$ ,  $-10^\circ$ ,  $10^\circ$  and  $30^\circ$  AoAs. The surface plot animations for successive time flow can be found with this *Sensors* paper on the MDPI website as Supplementary Materials. Figure 8 shows the simulation of the specific time indicated at the plot center. The Gaussian pulse is reached at the simulation boundary at the noted time, and a portion of the wave is reflected to other directions illustrated by the streamline. The normal reflections in Figure 8a,b are represented by the relatively straight streamline and propagation path. The non-perpendicular reflections in Figure 8c,d are depicted by the curved streamline and indirect path. The structure plates are properly designed to reflect the incoming signal toward the receiver for a non-normal incident over certain positive angles. Therefore, the strong second arrivals result from the non-perpendicular reflections.

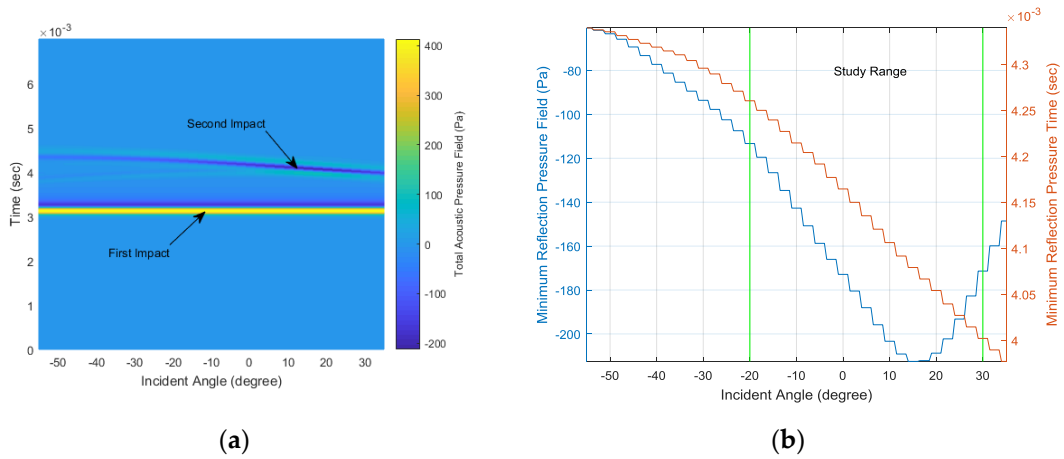


**Figure 7.** The received acoustic pressure field for various AoAs.



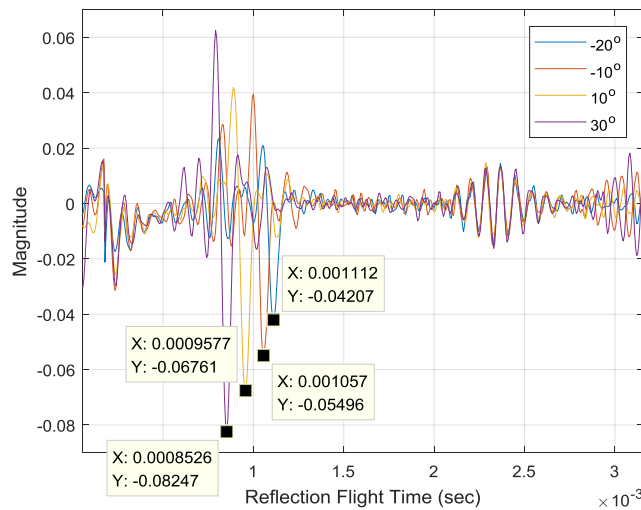
**Figure 8.** Total acoustic pressure field (surface plot) and local acceleration (streamline) for (a)  $-20^\circ$ , (b)  $-10^\circ$ , (c)  $10^\circ$  and (d)  $30^\circ$  AoA. Note that the white lines are intentionally inserted for visualizing the dominant propagation paths.

The LiveLink connection between the COMSOL simulator and MATLAB software provides a batch simulation process with continuous parameter variation. Figure 9 represents the simulation outcomes from the  $-55^\circ$  to  $35^\circ$  AoA range in every  $2.5^\circ$  step. The surface plot in Figure 9a distinctly demonstrates the constant first impact at around 3 ms and the linear second impact in the time range. Figure 9b shows the second arrival magnitude in the left y-axis and corresponding time in the right y-axis. Note that the time is the absolute value and not the relative time between the first and second arrivals. The structure is devised to deliver the strong second arrival near  $15^\circ$  and the linear arrival time from  $-20^\circ$  to  $30^\circ$  for optimal performance through a parametric search in the simulation. The designated scope is symbolized as the study range in Figure 9b. For a given AoA, the time delay induced by the structure is estimated by the HD algorithm and corresponds to the linear second arrival time shown in Figure 9b. Theoretically, the arrivals cannot be measured in the absolute time scale by using the single receiver; however, the time difference can be derived by using the HD algorithm.



**Figure 9.** (a) Simulated received signals with various angles at a  $2.5^\circ$  step and (b) the second arrival magnitude (left y-axis) and time (right y-axis).

The simulated Gaussian pulse response can be delivered to MATLAB for HD computation. The SRIRs from the HD algorithm are shown in Figure 10 for the  $-20^\circ$ ,  $-10^\circ$ ,  $10^\circ$  and  $30^\circ$  AoAs. Since the SRIR is the autocorrelation of the propagation response, the SRIR time represents the relative time scale equivalent to the time difference between the first and second impacts as the reflection flight time (RFT). The minimum value time in the SRIR indicates the highest likelihood for RFT that can be used for AoA computation. For increasing the AoA, the minimum time decreases at around 1 ms, as shown in the figure below with labels. The inversely-proportional relationship corresponds to the Gaussian pulse response plot in Figure 9b.

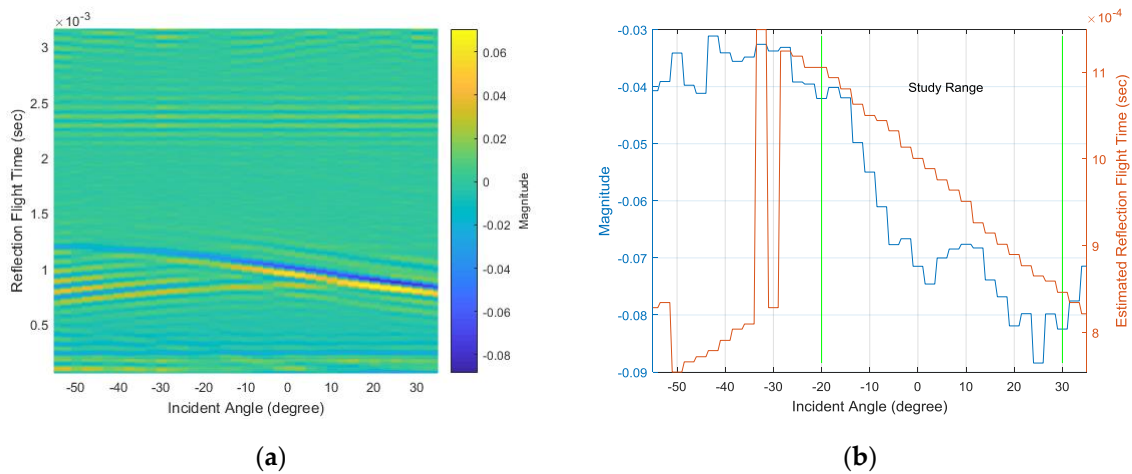


**Figure 10.** Simulated homomorphic deconvolution (HD) outputs for structure-related impulse response (SRIR).

The simulated SRIR from the  $-55^\circ$  to  $35^\circ$  AoA range is presented in Figure 11a with a  $2.5^\circ$  resolution. The minimum values denoted by the blue line gradually show the linear movement and intense depth within the  $-20^\circ \sim 30^\circ$  AoA range. The corresponding value and time are illustrated in Figure 11b in the left and right y-axes, respectively. The magnitude in Figure 11b signifies the reflection magnitude  $-a$  in Equation (7) and is related to the Figure 9b magnitude. The reflection magnitude decreases overall and indicates the minimum value at around  $25^\circ$ . The RFT also linearly diminishes within the study range and can be convertible to the Figure 9b time by adding the first impact time. Due to the weak reflection, the range outside of the study range provides an elevated magnitude,



as well as a fluctuating RFT. The consistency in the RFT is crucial information for determining the AoA with a simple linear model; therefore, the SRIR in the study range appropriately delivers clues for estimating the incident AoA.

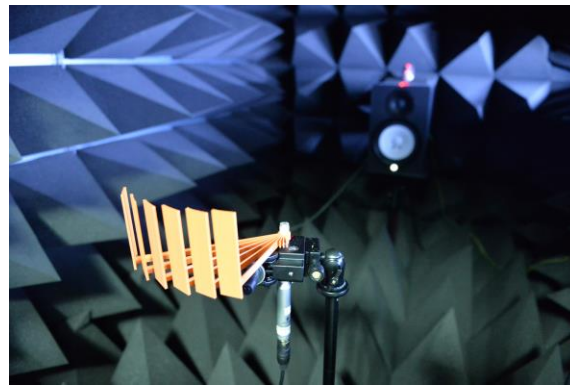


**Figure 11.** (a) Simulated SRIR with various angles at a  $2.5^\circ$  step and (b) the minimum magnitude (left y-axis) and reflection flight time (RFT) (right y-axis).

This section presents the RML structure design and simulation results related to the HD algorithm to determine the SRIR capability. The optimal RML structure presented above is derived from the parametric search through an extensive simulation. The Gaussian pulse response clearly shows the first and second impacts induced by the propagation and reflection, respectively. The linear time difference between the arrivals is denoted by the SRIR from HD computation. Due to the phase reversal from the reflection, the minimum value and its time represent the reflection magnitude and flight time, correspondingly. The designed RML structure provides the linear RFT variation within the  $-20^\circ \sim 30^\circ$  AoA range. Thus, the simple first-order model can convert the RFT to the incident AoA in the next section.

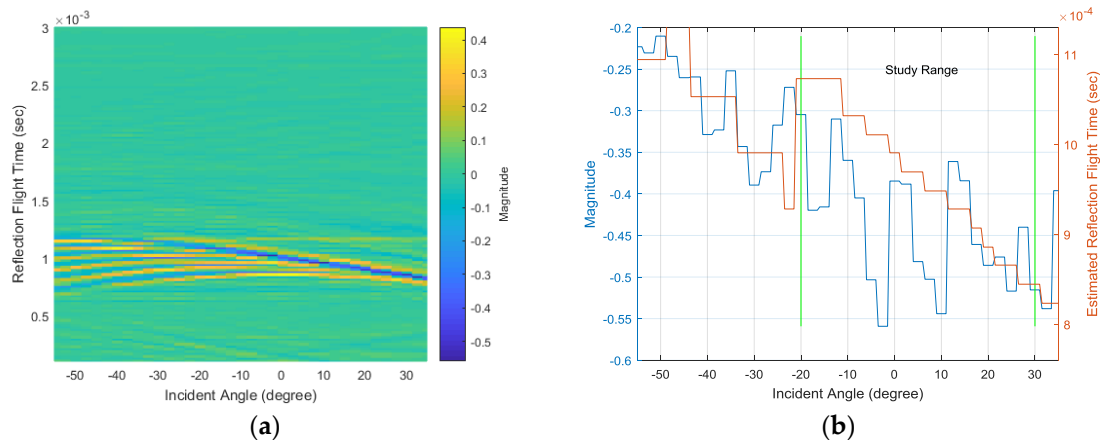
#### 4. Results

The RML structure is realized by the 3D printer (Replicator 2, MakerBot, Brooklyn, NY, USA) based on the polylactic acid (PLA) filament and is illustrated in Figure 12. The acoustic experiments are performed and analyzed in an anechoic chamber, described in Section 2, with a free-field condition [44]. The structure is located in the direct-front direction 1.38 m away from the speaker. The small hole in the structure is loaded with the measurement microphone vertically, as shown in Figure 12. The MATLAB programming controls the microphone (ECM8000, Behringer, Tortola, British Virgin Islands), computer-connected audio device (Quad-Capture, Roland, Hamamatsu, Japan) and speaker (HS80M, Yamaha, Hamamatsu, Japan) simultaneously. The MATLAB system object with the audio stream input/output (ASIO) driver processes the real-time audio in terms of generation, reception and execution. Thus, the single-process iteration generates the SRIR outcome from the HD computation for the given structure and AoA.



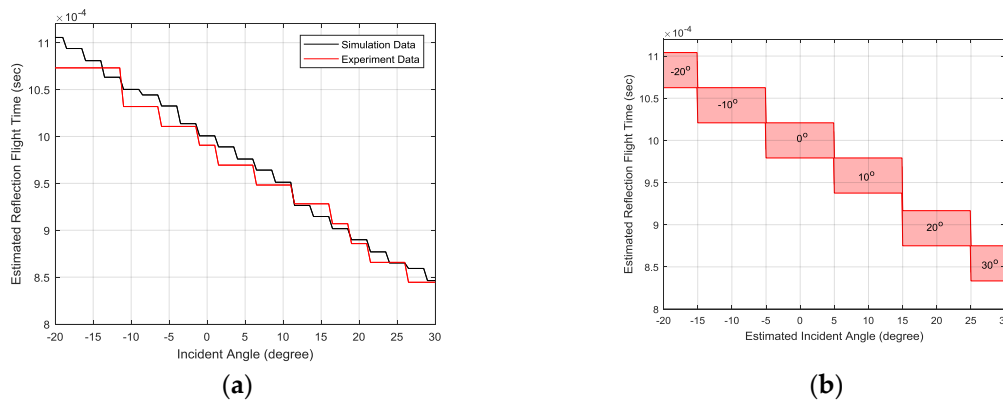
**Figure 12.** Acoustic experiment in the anechoic chamber with the 3D-printed RML structure.

The white noise with a normal distribution generates the full uniform spectrum signal from the speaker. The data window length for the HD parameter is 1024 samples, and the 30 second data processing with a truncating-transition head and tail portion are ensemble-averaged to present the dominant SRIR. Note that the sampling frequency is 48 kHz for all experiments. The SRIR experiment results from the  $-55^\circ$  to  $35^\circ$  AoA range are demonstrated in Figure 13a with a  $2.5^\circ$  resolution. Similar to the simulation results in Figure 11a, the blue line for the minimum values denotes the linear movement and intense depth within the  $-20^\circ$ ~ $30^\circ$  AoA range. The corresponding value and time are plotted in Figure 13b in the left and right y-axes, respectively. The reflection magnitude decreases overall with higher variance than the simulation results shown in Figure 11b. The RFT consistently diminishes within the study range and harmonizes well with the simulation counterpart.



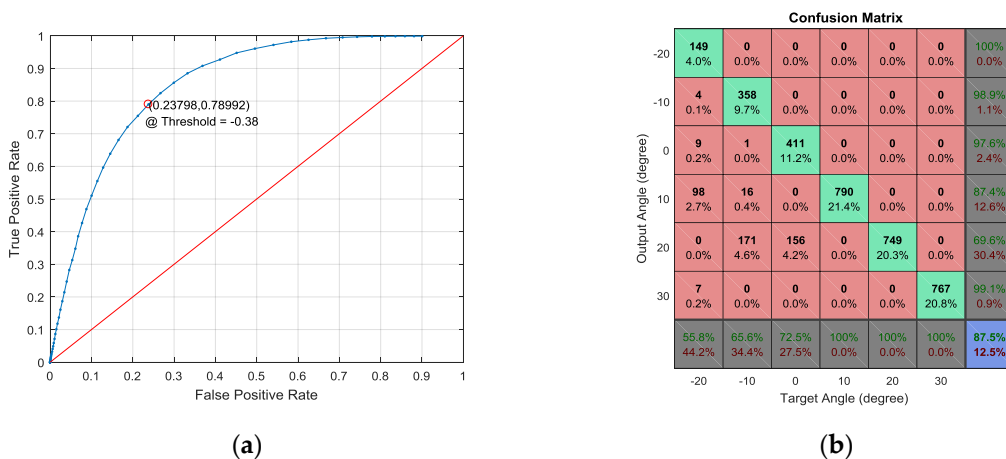
**Figure 13.** (a) SRIR experiment results for various angles at a  $2.5^\circ$  step and (b) the minimum magnitude (left y-axis) and RFT (right y-axis).

The consistent linearity of the RFT is important information for determining the RML system feasibility. Once the linearity is preserved, the individual RFT range corresponds to the specific AoA by using first-order arrangement. The simulation and experimentation results for the RFT are illustrated in Figure 14a. The experiment RFT outcome follows the simulation counterpart well in terms of linearity, as well as value. The AoA estimation model for the given experiment is derived in Figure 14b. The output AoAs are divided into six angles from  $-20^\circ$  to  $30^\circ$  with a  $10^\circ$  resolution, and the corresponding RFT range is illustrated in the figure. For example, the RFT scope from 1.06 ms to 1.10 ms represents the  $-20^\circ$  AoA. From 0.83 ms to 1.10 ms, the six discrete RFT scopes are continuously distributed for six AoAs. Note that the RFT decision gap located between  $10^\circ$  and  $20^\circ$  does not influence the estimation performance because of the RFT distribution.



**Figure 14.** (a) The simulation and experimentation results for RFT from SRIR within the study range and (b) the discrete linear model for estimating AoA from RFT. Each red box represents the single 10-degree angle, such as  $-20^\circ$  and  $-10^\circ$ .

The RML system is designed to estimate the limited field of localization from the  $-20^\circ$  to  $30^\circ$  AoA range. Angles outside this range are discriminated by the reflection magnitude to exclude the given AoA from the localization process. The optimal threshold for the reflection magnitude is derived from the receiver operating characteristic (ROC) curve shown in Figure 15a. A total of 33,768 data frames (16,884 within and 16,884 outside range) are evaluated for the ROC curve, and the individual frames are evenly distributed over the designated AoA range. The single frame is equivalent to the 1024 sample data window. The true positive rate (TPR) is the ratio of the number of true positives to the number of positive conditions. The false positive rate (FPR) is derived by dividing the number of false positives by the number of negative conditions. A positive condition indicates the data frame from the study range, and a negative condition specifies a frame from outside of the range. As the decision threshold is changed from minimum to maximum, the TPR and FPR change the value in a complementary manner, as shown in Figure 15a. The area near the upper-left corner shows the best detection performance due to the perfect TPR and zero FPR of the corner; therefore, the decision threshold of  $-0.38$ , which is the nearest point to the corner, produces the highest statistical performance (78.99% TPR and 23.80% FPR).



**Figure 15.** (a) The receiver operating characteristic (ROC) curve for intrinsic AoA range detection and (b) the confusion matrix for given discrete AoA.

Below the  $-0.38$  reflection magnitude, the RML system considers the given signal as data from the localization field. Below the threshold, the RML algorithm performs the HD for the SRIR and finds the minimum value for the RFT over the AoA. According to the structure angular distance between the

plates, the resolvable AoA is determined to be  $10^\circ$  for the study range. Thus, the acoustic experiment is executed for the AoA range from  $-20^\circ$  to  $30^\circ$  with a  $10^\circ$  resolution. Figure 15b demonstrates the confusion matrix for the designated angles. The individual data frame for the angle is selected randomly without prior knowledge of the AoA; therefore, the number of frames for the dataset is distributed with variance. The total data frames for the confusion matrix is 3686. The number of data frames for  $-20^\circ$ ,  $-10^\circ$ ,  $-0^\circ$ ,  $10^\circ$ ,  $20^\circ$  and  $30^\circ$  is 267, 546, 567, 790, 749 and 767, respectively. Note that data used in the confusion matrix are assumed to belong to the study range AoA.

The AoA for each column vector of the matrix represents the target AoA (condition) in which the signal is incident. The AoA for each row vector indicates the output AoA (test outcome) that the RML system determines. The green and red rates on the last row signify the true-positive rate (hit rate) and false-negative rate (miss rate) for each AoA, respectively. In the bottom-right-hand corner, the overall values of the hit and miss rates are numerated. Except the bottom row and rightmost column, the confusion matrix elements denote the decision counts and overall percentile. The diagonal elements with a bright green color demonstrate the counts and percentile for the true-positive case. Observe that the percentile shows the ratio to total events. The performance for an individual AoA is represented in the last row as the hit rate.

As the AoA increases, the hit rate increases substantially (i.e., from 55.8% to 100%). The negative and zero angles that propagate the signal over the single direct line with an identical reflection path show the deteriorated hit rate performance. However, the positive angles that deliver the signal over the independent path for the propagation and reflection demonstrate the perfect hit rate. The  $-20^\circ$ ,  $-10^\circ$  and  $0^\circ$  AoAs are dominantly misinterpreted for  $10^\circ$ ,  $20^\circ$  and  $20^\circ$ , correspondingly. The structure around the major normal reflection plate creates early bounces that cause decision confusion. The more plates near the direct angle present a lower hit rate overall. For example, the  $-20^\circ$  AoA has two plates on one side and three plates on the other. The  $0^\circ$  AoA has zero plates on one side and five plates on the other. Further plates on both sides show the deteriorated decision accuracy performance. The positive AoAs do not include the plate on the direct propagation line, which provides a straight connection between the source and receiver. The reflection is induced by the obliquely-located plates and distributed widely in direction to avoid AoA decision ambiguity. The focused reflection to the receiver provides the dominant temporal information for the RML algorithm. The overall hit rate for the AoA study range is 87.5% and 69.1% for all ranges ( $0.875 \times 0.790$ ) according to the ROC TPR 79.0%.

## 5. Conclusions

This paper presents a novel localization method for finding the arrival angle of far-field sound propagation with a single microphone. The reflection structure over the microphone produces direction-wise temporal variation, which can be estimated by the structure-related impulse response from homomorphic deconvolution. The structure consists of six vertical rectangular plates with a  $10^\circ$  angular (from  $0^\circ$  to negative angles) and a 1 cm radius (from 15 cm to increasing distance) difference. Depending on the incident angle, the reflection signal experiences a distinctive propagation path for a unique flight time. According to the simulation and experiment, the proposed structure provides the linear variation of the time difference between the first and second arrivals for consistent movement. The homomorphic deconvolution used in this paper utilizes the real cepstrum and inverse cepstrum sequentially to derive the spatial propagation response's autocorrelation, named the structure-related impulse response. The reflection physics is represented by the minimum value and time at the response in the algorithm. The simulation and experiment demonstrate that the prominent value and linear time can be observed from the  $-20^\circ$  to  $30^\circ$  scope; therefore, the angle span is specified as the study range. The acoustic experiments in the anechoic chamber denote that 79.0% of the study range data from the isotropic signal is properly detected by the response value, and 87.5% of the specific direction data from the study range signal is properly estimated by the response time. The product of both rates ( $0.875 \times 0.790$ ) show the overall hit rate to be 69.1%.

The novelty of this paper can be found in the structure, algorithm and their connection. The structure is devised from the extensive parametric search simulations for optimal reflections in magnitude and time. The conventional homomorphic deconvolution employs the complex cepstrum to estimate the input signal or the propagation response. This proposition includes the homomorphic deconvolution for the autocorrelation of the propagation response based on the real cepstrum. The mutual direction optimizations between the structure and algorithm are exercised for the increased hit rate and decreased miss rate, as well. The fine-grained simulation is also included in the design procedure along with the preliminary modeling of the acoustic reflection; thus, the localization accuracy is improved considerably for the far-field sound source. This article is part of the future research described in the previous paper [41] as the near-field monaural localization extension to the far-field system. Future work will offer structure and algorithm enhancement for the comprehensive field of localization. The selection potential in terms of acoustic structures and estimation algorithms is considerable. Together with the continuous structure for high-resolution localization, future work will include the utilization of various structure architectures. The algorithm will be devised as a mathematical model to represent the received information by its coefficients as a parametric method. The similarity between the consecutive datasets will be explored by temporal post-processing in a statistical manner. With all of the above, the development of three-dimensional monaural localization for the azimuthal and elevation directions is the final objective of this research.

**Supplementary Materials:** The following are available online at [www.mdpi.com/1424-8220/17/10/2189/s1](http://www.mdpi.com/1424-8220/17/10/2189/s1): Video S1: The  $-20^\circ$  incident angle simulation video, Video S2: The  $-10^\circ$  incident angle simulation video, Video S3: The  $10^\circ$  incident angle simulation video, Video S4: The  $30^\circ$  incident angle simulation video.

**Acknowledgments:** This work was supported by the Dongguk University Research Fund of 2017.

**Author Contributions:** Y.P. and K.K. initiated and realized the fundamental design of the system. A.C. suggested the experimental methods. Y.P. performed the extensive experiments on the structure in the anechoic chamber. K.K. was responsible for the algorithm and simulation, including the data analysis and representation. K.K. wrote the entire manuscript. Y.P. and A.C. revised the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript (listed by order of appearance in the text).

3D	Three-dimensional
AoA	Angle of arrival (AoAs, angles of arrival)
ML	Monaural localization
RML	Reflective monaural localization
HD	Homomorphic deconvolution
ADC	Analog-to-digital converter
LTI	Linear time-invariant
SRIR	Structure-related impulse response
DFT	Discrete Fourier transform
IDFT	Inverse discrete Fourier transform
2D	Two-dimensional
RFT	Reflection flight time
PLA	Polylactic acid
ASIO	Audio stream input/output
ROC	Receiver operating characteristic
TPR	True positive rate
FPR	False positive rate

## Appendix A

The equations below illustrate the homomorphic deconvolution procedure based on the real cepstrum. The received signal  $y[n]$  is propagated through the two cascade linear time-invariant systems as  $h_1[n]$  and  $h_2[n]$ . The sound source is  $x[n]$ . In Equation (A9),  $\hat{x}[n] + \hat{x}^*[-n]$  is represented by  $c_x[n]$ , etc. We assume that the window function  $w[n]$  can successfully separate  $c_{h_1}[n] + c_{h_2}[n]$  from  $c_y[n]$  in Equation (A10). The final output of the homomorphic deconvolution is the autocorrelation of  $h_1[n] * h_2[n]$ , as shown in Equation (A13).

$$y[n] = x[n] * h_1[n] * h_2[n] \quad (\text{A1})$$

$$\xrightarrow{\text{DFT}} Y[k] = \sum_{n=0}^{N-1} y[n] e^{-j\frac{2\pi}{N}kn} = X[k]H_1[k]H_2[k] \quad (\text{A2})$$

$$\xrightarrow{\text{Squared Magnitude}} |Y[k]|^2 = |X[k]|^2 |H_1[k]|^2 |H_2[k]|^2 \quad (\text{A3})$$

$$\xrightarrow{\text{Logarithm}} 2\log|Y[k]| = 2\log(|X[k]| |H_1[k]| |H_2[k]|) \\ = 2\log|X[k]| + 2\log|H_1[k]| + 2\log|H_2[k]| \quad (\text{A4})$$

$$\xrightarrow{\text{IDFT}} c_y[n] = \frac{2}{N} \sum_{n=0}^{N-1} \log|Y[k]| e^{j\frac{2\pi}{N}kn} \quad (\text{A5})$$

$$= \frac{2}{N} \sum_{n=0}^{N-1} \log|X[k]| e^{j\frac{2\pi}{N}kn} + \frac{2}{N} \sum_{n=0}^{N-1} \log|H_1[k]| e^{j\frac{2\pi}{N}kn} + \frac{2}{N} \sum_{n=0}^{N-1} \log|H_2[k]| e^{j\frac{2\pi}{N}kn} \quad (\text{A6})$$

$$= \frac{1}{N} \sum_{n=0}^{N-1} \log(X[k]X^*[k]) e^{j\frac{2\pi}{N}kn} + \frac{1}{N} \sum_{n=0}^{N-1} \log(H_1[k]H_1^*[k]) e^{j\frac{2\pi}{N}kn} \\ + \frac{1}{N} \sum_{n=0}^{N-1} \log(H_2[k]H_2^*[k]) e^{j\frac{2\pi}{N}kn} \quad (\text{A7})$$

$$= \frac{1}{N} \sum_{n=0}^{N-1} (\log X[k] + \log X^*[k]) e^{j\frac{2\pi}{N}kn} + \frac{1}{N} \sum_{n=0}^{N-1} (\log H_1[k] + \log H_1^*[k]) e^{j\frac{2\pi}{N}kn} \\ + \frac{1}{N} \sum_{n=0}^{N-1} (\log H_2[k] + \log H_2^*[k]) e^{j\frac{2\pi}{N}kn} \quad (\text{A8})$$

$$= \hat{x}[n] + \hat{x}^*[-n] + \hat{h}_1[n] + \hat{h}_1^*[-n] + \hat{h}_2[n] + \hat{h}_2^*[-n] \\ = c_y[n] = c_x[n] + c_{h_1}[n] + c_{h_2}[n] \quad (\text{A9})$$

$$\xrightarrow{\text{Window}} c_u[n] = c_y[n]w[n] \approx c_{h_1}[n] + c_{h_2}[n] \\ = \frac{1}{N} \sum_{n=0}^{N-1} \log|H_1[k]|^2 e^{j\frac{2\pi}{N}kn} + \frac{1}{N} \sum_{n=0}^{N-1} \log|H_2[k]|^2 e^{j\frac{2\pi}{N}kn} \quad (\text{A10})$$

$$\xrightarrow{\text{DFT}} U[k] = \sum_{n=0}^{N-1} c_u[n] e^{-j\frac{2\pi}{N}kn} \approx \log|H_1[k]|^2 + \log|H_2[k]|^2 \quad (\text{A11})$$

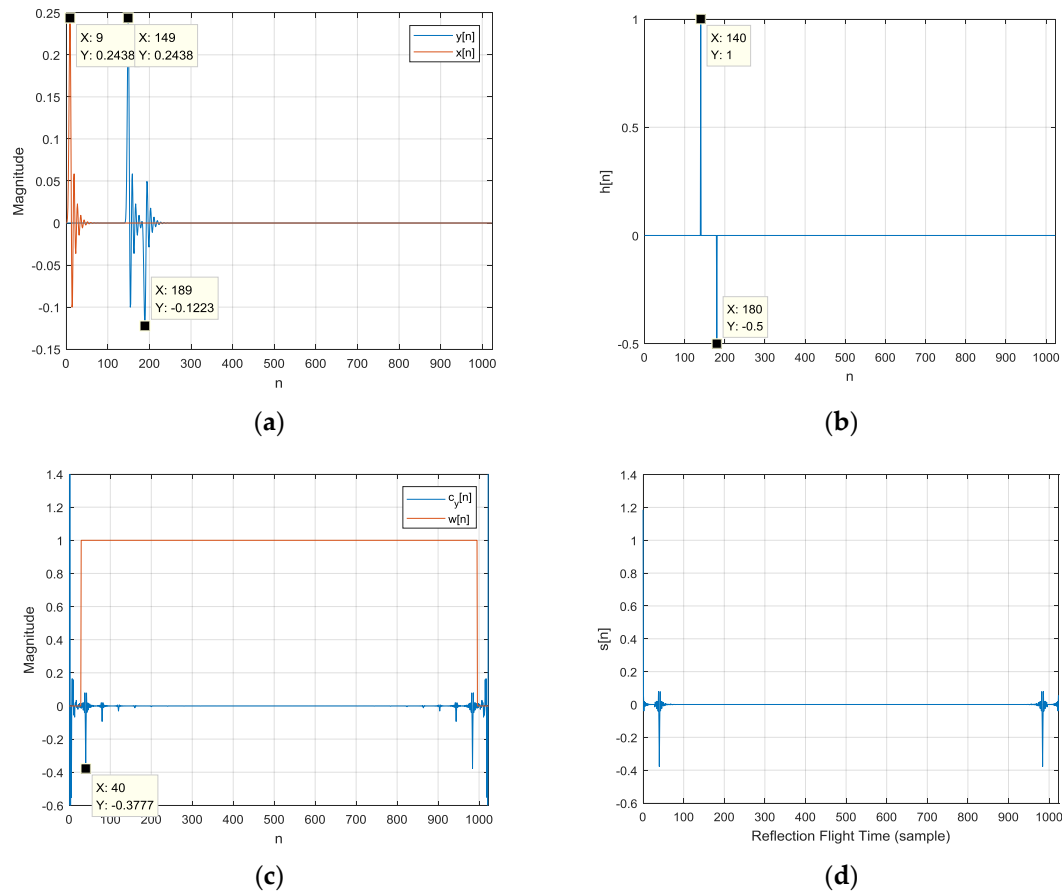
$$\xrightarrow{\text{Exponential}} e^{\log|H_1[k]|^2 + \log|H_2[k]|^2} = e^{\log|H_1[k]|^2} e^{\log|H_2[k]|^2} = |H_1[k]|^2 |H_2[k]|^2 \quad (\text{A12})$$

$$\xrightarrow{\text{IDFT}} \frac{1}{N} \sum_{n=0}^{N-1} |H_1[k]|^2 |H_2[k]|^2 e^{j\frac{2\pi}{N}kn} = \frac{1}{N} \sum_{n=0}^{N-1} H_1[k]H_1^*[k]H_2[k]H_2^*[k] e^{j\frac{2\pi}{N}kn} \\ = h_1[n] * h_1^*[-n] * h_2[n] * h_2^*[-n] = \hat{s}[n] \quad (\text{A13})$$

The subsequent example demonstrates the homomorphic deconvolution for a low-frequency signal. The signal is derived from the impulse response of the Butterworth filter with the 0.25 normalized bandwidth. Figure A1a shows the sound source  $x[n]$  (red plot) and received signal  $y[n]$  (blue plot). Figure A1b represents the impulse response of the propagation as  $h_1[n] * h_2[n]$ . Figure A1c denotes the



real cepstrum outcome  $c_y[n]$  and window function  $w[n]$ . The significant power of  $x[n]$  is concentrated on the low index portion of the real cepstrum; therefore, the window function can separate the impulse response exhibited by the harmonics of the delta function. The final output  $\hat{s}[n]$  of the algorithm is presented by Figure A1d as the autocorrelation of the impulse response. The difference between the first and second arrivals, 40 samples (180–140), is derived in the result.



**Figure A1.** Homomorphic deconvolution for low-frequency signal based on the impulse response of the Butterworth filter (BW: 0.25). (a) Sound source  $x[n]$  and received signal  $y[n]$ ; (b) impulse response of propagation; (c) real cepstrum output  $c_y[n]$  and window function  $w[n]$ ; and (d) the output of the homomorphic deconvolution as  $\hat{s}[n]$ .

## Appendix B

**Table A1.** COMSOL Multiphysics simulator parameters.

Name	Value	Description
f0	4000 Hz	Frequency Bandwidth <sup>1</sup>
A	2 m <sup>2</sup> /s	Amplitude <sup>1</sup>
t0	2.5 × 10 <sup>-6</sup> s	Peak Time <sup>1</sup>
N	6	Mesh Elements/λ
C	343 m/s	Sound Speed
h_max	0.014292 m	Maximum Mesh Size
fs	4.8 × 10 <sup>5</sup> Hz	Sampling Frequency
CFL	0.05	Relationship between Mesh Size and Time-step
Delt	2.0833 × 10 <sup>-6</sup> s	Maximum Time-step

<sup>1</sup> Gaussian pulse.

## References

1. Iida, K.; Itoh, M.; Itagaki, A.; Morimoto, M. Median plane localization using a parametric model of the head-related transfer function based on spectral cues. *Appl. Acoust.* **2007**, *68*, 835–850. [[CrossRef](#)]
2. Wightman, F.L.; Kistler, D.J. Monaural sound localization revisited. *J. Acoust. Soc. Am.* **1997**, *101*, 1050–1063. [[CrossRef](#)] [[PubMed](#)]
3. Chen, J.; Van Veen, B.D.; Hecox, K.E. External ear transfer function modeling: A beamforming approach. *J. Acoust. Soc. Am.* **1992**, *92*, 1933–1944. [[CrossRef](#)] [[PubMed](#)]
4. Asano, F.; Suzuki, Y.; Sone, T. Role of spectral cues in median plane localization. *J. Acoust. Soc. Am.* **1990**, *88*, 159–168. [[CrossRef](#)] [[PubMed](#)]
5. Musicant, A.D.; Butler, R.A. The influence of pinnae-based spectral cues on sound localization. *J. Acoust. Soc. Am.* **1984**, *75*, 1195–1200. [[CrossRef](#)] [[PubMed](#)]
6. Searle, C.L.; Braid, L.D.; Cuddy, D.R.; Davis, M.F. Binaural pinna disparity: Another auditory localization cue. *J. Acoust. Soc. Am.* **1975**, *57*, 448–455. [[CrossRef](#)] [[PubMed](#)]
7. Wright, D.; Hebrank, J.H.; Wilson, B. Pinna reflections as cues for localization. *J. Acoust. Soc. Am.* **1974**, *56*, 957–962. [[CrossRef](#)] [[PubMed](#)]
8. Hebrank, J.; Wright, D. Spectral cues used in the localization of sound sources on the median plane. *J. Acoust. Soc. Am.* **1974**, *56*, 1829–1834. [[CrossRef](#)] [[PubMed](#)]
9. Gardner, M.B.; Gardner, R.S. Problem of localization in the median plane: Effect of pinnae cavity occlusion. *J. Acoust. Soc. Am.* **1973**, *53*, 400–408. [[CrossRef](#)] [[PubMed](#)]
10. Shaw, E.A.G.; Teranishi, R. Sound pressure generated in an external-ear replica and real human ears by a nearby point source. *J. Acoust. Soc. Am.* **1968**, *44*, 240–249. [[CrossRef](#)] [[PubMed](#)]
11. Batteau, D.W. The role of the pinna in human localization. *Proc. R. Soc. Lond. B Biol. Sci.* **1967**, *168*, 158–180. [[CrossRef](#)] [[PubMed](#)]
12. Takemoto, H.; Mokhtari, P.; Kato, H.; Nishimura, R.; Iida, K. Mechanism for generating peaks and notches of head-related transfer functions in the median plane. *J. Acoust. Soc. Am.* **2012**, *132*, 3832–3841. [[CrossRef](#)] [[PubMed](#)]
13. Spagnol, S.; Geronazzo, M.; Avanzini, F. On the relation between pinna reflection patterns and head-related transfer function features. *IEEE Trans. Audio Speech Lang. Process.* **2013**, *21*, 508–519. [[CrossRef](#)]
14. Harris, J.G.; Pu, C.-J.; Principe, J.C. A monaural cue sound localizer. *Analog. Integr. Circuits Signal Process.* **2000**, *23*, 163–172. [[CrossRef](#)]
15. Nakashima, H.; Mukai, T. 3d sound source localization system based on learning of binaural hearing. In Proceedings of the 2005 IEEE International Conference on Systems, Man and Cybernetics, Waikoloa, HI, USA, 10–12 October 2005; Volume 3534, pp. 3534–3539.
16. Lee, S.; Park, Y.; Choi, J.-S. Estimation of multiple sound source directions using artificial robot ears. *Appl. Acoust.* **2014**, *77*, 49–58. [[CrossRef](#)]
17. Kumon, M.; Shimoda, T.; Kohzawa, R.; Mizumoto, I.; Iwai, Z. Audio servo for robotic systems with pinnae. In Proceedings of the 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems, Edmonton, AB, Canada, 2–6 August 2005; pp. 1881–1886.
18. Hwang, S.; Park, Y.; Park, Y.-S. Sound direction estimation using an artificial ear for robots. *Robot. Auton. Syst.* **2011**, *59*, 208–217. [[CrossRef](#)]
19. Tomoko, S.; Toru, N.; Makoto, K.; Ryuichi, K.; Ikuro, M.; Zenta, I. Spectral cues for robust sound localization with pinnae. In Proceedings of the 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems, Beijing, China, 9–15 October 2006; pp. 386–391.
20. Schillebeeckx, F.; De Mey, F.; Peremans, H. Bio-inspired sonar antennae: Enhancing directivity patterns for localization. In Proceedings of the 2008 2nd IEEE RAS & EMBS International Conference on Biomedical Robotics and Biomechatronics, Scottsdale, AZ, USA, 19–22 October 2008; pp. 366–371.
21. Saxena, A.; Ng, A.Y. Learning sound location from a single microphone. In Proceedings of the IEEE International Conference on Robotics and Automation, ICRA'09, Kobe, Japan, 12–17 May 2009; pp. 1737–1742.
22. Kumon, M.; Noda, Y. Active soft pinnae for robots. In Proceedings of the 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Francisco, CA, USA, 25–30 September 2011; pp. 112–117.

23. Murray, J.C.; Erwin, H.R. A neural network classifier for notch filter classification of sound-source elevation in a mobile robot. In Proceedings of the 2011 International Joint Conference on Neural Networks, San Jose, CA, USA, 31 July–5 August 2011; pp. 763–769.
24. Takashima, R.; Takiguchi, T.; Ariki, Y. Dimensional feature weighting utilizing multiple kernel learning for single-channel talker location discrimination using the acoustic transfer function. *J. Acoust. Soc. Am.* **2013**, *133*, 891–901. [[CrossRef](#)] [[PubMed](#)]
25. Takiguchi, T.; Sumida, Y.; Takashima, R.; Ariki, Y. Single-channel talker localization based on discrimination of acoustic transfer functions. *Eurasip J. Adv. Signal Process.* **2009**. [[CrossRef](#)]
26. Takashima, R.; Takiguchi, T.; Ariki, Y. Hmm-based separation of acoustic transfer function for single-channel sound source localization. In Proceedings of the 2010 IEEE International Conference on Acoustics, Speech, and Signal Processing, ICASSP 2010, Dallas, TX, USA, 14–19 March 2010; pp. 2830–2833.
27. Takashima, R.; Takiguchi, T.; Ariki, Y. Monaural sound-source-direction estimation using the acoustic transfer function of a parabolic reflection board. *J. Acoust. Soc. Am.* **2010**, *127*, 902–908. [[CrossRef](#)] [[PubMed](#)]
28. Friedland, G.; Yeo, C.; Hung, H. Dialocalization: Acoustic speaker diarization and visual localization as joint optimization problem. *ACM Trans. Multimed. Comput. Commun. Appl.* **2010**, *6*. [[CrossRef](#)]
29. Oppenheim, A.V. Generalized superposition. *Inf. Control* **1967**, *11*, 528–536. [[CrossRef](#)]
30. Gold, B.; Rader, C.M. *Digital Processing of Signals*; McGraw-Hill: New York, NY, USA, 1969.
31. Oppenheim, A.; Schaffer, R.; Stockham, T. Nonlinear filtering of multiplied and convolved signals. *IEEE Trans. Audio Electr.* **1968**, *16*, 437–466. [[CrossRef](#)]
32. Stockham, T. The application of generalized linearity to automatic gain control. *IEEE Trans. Audio Electr.* **1968**, *16*, 267–270. [[CrossRef](#)]
33. Stockham, T.G., Jr. Image processing in the context of a visual model. *Proc. IEEE* **1972**, *60*, 828–842. [[CrossRef](#)]
34. Oppenheim, A.V. Speech analysis-synthesis system based on homomorphic filtering. *J. Acoust. Soc. Am.* **1969**, *45*, 458–465. [[CrossRef](#)] [[PubMed](#)]
35. Murthy, I.S.N.; Rangaraj, M.R.; Jayaram Udupa, K.; Goyal, A.K. Homomorphic analysis and modeling of ECG signals. *IEEE Trans. Biomed. Eng.* **1979**, *BME-26*, 330–344. [[CrossRef](#)]
36. Tribolet, J.M.; Oppenheim, A.V. Deconvolution of seismic data using homomorphic filtering. In Proceedings of the Joint Automatic Control Conference, Tuscaloosa, AL, USA, 7–9 March 1993.
37. Bogert, B.P.; Healy, M.J.R.; Tukey, J.W. The quefrency analysis of time series for echoes: Cepstrum, pseudo-autocovariance, cross-cepstrum, and saphe-cracking. In *Proc. Symp. Time Series Analysis*; Rosenblatt, M., Ed.; Wiley: New York, NY, USA, 1963; pp. 209–243.
38. Childers, D.G.; Skinner, D.P.; Kemerait, R.C. The cepstrum: A guide to processing. *Proc. IEEE* **1977**, *65*, 1428–1443. [[CrossRef](#)]
39. Oppenheim, A.V.; Schaffer, R.W. From frequency to quefrency: A history of the cepstrum. *Signal Proc. Mag. IEEE* **2004**, *21*, 95–106. [[CrossRef](#)]
40. Kim, K.; Kim, Y. Monaural sound localization based on structure-induced acoustic resonance. *Sensors* **2015**, *15*, 3872–3895. [[CrossRef](#)] [[PubMed](#)]
41. Kim, Y.; Kim, K. Near-field sound localization based on the small profile monaural structure. *Sensors* **2015**, *15*, 28742–28763. [[CrossRef](#)] [[PubMed](#)]
42. Kim, K.; Choi, A. Binaural sound localizer for a azimuthal movement detection based on diffraction. *Sensors* **2012**, *12*, 10584–10603. [[CrossRef](#)] [[PubMed](#)]
43. Kim, K. Lightweight filter architecture for energy efficient mobile vehicle localization based on a distributed acoustic sensor network. *Sensors* **2013**, *13*, 11314–11335. [[CrossRef](#)] [[PubMed](#)]
44. Kim, K. Design and analysis of experimental anechoic chamber for localization. *J. Acoust. Soc. Korea* **2012**, *31*, 10. [[CrossRef](#)]
45. Boyer, C.B. *A History of Mathematics*; Wiley: New York, NY, USA, 1968.
46. Standardization, I.O.F. Acoustics—Determination of sound power levels of noise sources using sound pressure—precision methods for anechoic and hemi-anechoic rooms. In *ISO 3745:2003*; ISO: Geneva, Switzerland, 2003.

