Simulation of Microphone Array for Sound Localization using Human Binaural Hearing Model

Viktor Hristov, Snejana Pleshkova

Abstract – Investigation of the human hearing system and attempts for its simulation are one of the most popular themes nowadays. If the sound signals are received from a person it is easy to determine the proper direction only with the help of the human ears and human brain. But if the human is replaced with a mobile robot - combination of a proper microphone array structure and the corresponding sound localization algorithm is needed. Both working together are important to achieve the appropriate accuracy of direction of sound arrival and the corresponding sound source localization. There are proposed many microphone array structures and algorithms, which are tested and can be characterized with advantages and disadvantages in some concrete applications. Here in this article is proposed to apply and test a combination of a simple two microphone array structure and algorithm for modelling and simulating binaural human perception. The most important thing in this proposal is the simulation of human ears with two microphones and to choose their type to have the beam forming diagram of sensitivity similar each other. Also in simulation algorithm is proposed to apply the elements of human binaural hearing to as much as is possible to known the complex and not fully understand the processes in the human brain on auditory perception and in particular human perception of the direction from which sound waves arrive. The results from simulations are presented to show the characteristics of sound source localization achieved with proposition to apply human binaural hearing model for the purposes of sound direction of arrival and sound source localization.

Keywords – Human binaural hearing model, Sound source localization, Microphone arrays.

I. INTRODUCTION

Sound source localization is a powerful approach of finding direction of sound arrival from sound source to a target. Usually it could be a person, but there are also the applications, where the target can be a mobile robot, a video conferencing audio system or something other like artificial audio perception system. If the sounds from the source are received from a person, the human hearing system is capable to determine properly direction of sound arrival from sound source. This is done not only with the help of the human ears, but with the leading role of the human brain, where the received from the two human ears binaural audio information is processed to form in human mind the impression of sound direction of arrival. Unfortunately, for mobile robots or for other auditory perception systems this is not possible. Reason for this is the absence in them of similar system as the human brain with such capabilities to analyze the received audio information and to determine the sound direction of arrival. Therefore, there is a need to develop some appropriate methods and algorithms to find the proper sound arrival direction without the presence of a system similar to the human brain. Sound source localization can be separated and described in three main groups: depending on the time of arrival, fixed beamforming and adaptive beamforming using correlation functions [1]. There are two scenarios depending on the place, where experiments was done. It could be in open space, without any reflections and in closed space, also called multipath propagation, because each reflection from the wall or object can be treated as a separate source [2]. A microphone array consist of a number of acoustic sensors positioned on the proper distance between each other in a way that spatial information can be captured [3]. As a result, the array outputs contain signal of interest, noise, interference and also propagation information that is represented by the acoustic impulse response from the radiating sources to the microphones. The combination of a microphone array and the corresponding sound localization algorithm can be used to simulate the human hearing system [4]. Both microphone array structure and chosen algorithm working together are important to achieve the appropriate accuracy of direction of sound wave arrival and the corresponding sound source localization. A number of factors affect the spatial aspects of how a sound is perceived. The “Duplex theory” was the first comprehensive analysis of the physics of binaural perception [5]. There is a different time of arrival of the sound signals to the receiving microphones (human ears) as illustrated in Fig.1.

![Fig.1. Interaural differences of time and intensity of impinging sound wave on an ideal spherical head from a distant source](image)

Unless a sound source is located directly in front of or behind the head, sound arrives slightly earlier in time at the ear that is physically closer to the source, and with somewhat greater intensity [6]. This interaural time difference (ITD) is produced because it takes longer for the sound to arrive at the...
ear that is farther from the source [7]. The interaural intensity difference (IID) is produced because the 'shadowing' effect of the head prevents some of the incoming sound energy from reaching the ear that is turned away from the direction of the source [8]. The estimation of ITD is the most critical aspect of the binaural processing. Many models are based on the cross-correlation of the signals to the two ears after processing by the auditory periphery [9]. Several series of measurements have been performed the time difference in receiving of the sound source [10]. If the imaginary or artificial head as model of the human head in a mobile robot human hearing model of mobile robot microphone array had completely spherical and uniform surface, as Fig.1, the interaural time difference (ITD) produced by a sound source that arrives from an azimuth of \( \theta \) radians can be approximately described with the following equation using diffraction theory [11, 12 and 13]:

\[
\tau = \left( \frac{a}{c} \right) \sin \theta, 
\]

where

- \( \tau \) is the interaural time difference (ITD);
- \( a \) - the radius of the imaginary or artificial model of human head;
- \( c = 343 \text{ m/s} \) - the speed of sound waves in the air.

There are proposed many microphone array structures and algorithms, which are tested and can be characterized with advantages and disadvantages in some concrete applications. The goal of this article is to propose, apply and test a combination of a simple two microphone array structure and algorithm for modelling and simulating binaural human hearing. The most important in this proposal is the arrangement of two microphones in such a way like the human ears and to choose their type with the proper sensitivity. Also, in the simulation algorithm, it is proposed to apply the elements of human binaural hearing to as much as is possible to know the complex and not fully understand the processes in the human brain on auditory perception and in particular human perception of the direction from which sound waves arrive.

II. DEVELOPMENT OF GEOMETRICAL MODEL FOR SOUND SOURCE LOCALIZATION

The geometrical model of microphone array based on the human binaural hearing is proposed in Fig.2. It is based on the human hearing aid and binaural sound source localization. It consists of two microphones \( M_1 \) and \( M_2 \), with the identical parameters. They are situated on the short distance \( d \) between them, with the value similar to the distance of the ears in the human head. The positions of the two microphones are shown in Fig. 2 as coordinates \( x_{m1}, y_{m1} \) and \( x_{m2}, y_{m2} \), respectively (assuming \( x \) as horizontal, \( y \) as vertical coordinate and the origin of coordinate system in the left upper corner in Fig.2). Sound source, producing corresponding sound source signal \( S \), (usually a speaking person or speaker producing sound waves from an audio system) is positioned on the other end of the room. It is be described (for the proposed in Fig. 2 planar case of geometric model) with the coordinates \( (x_{src}, y_{src}) \). In Fig. 2 are shown also the model of the imaginary or artificial head (as the circle in dark) as model of the human head and also two circles presenting the microphone beamforming diagrams.

Usually the number of sound waves arriving to the human ears or to each of the microphones in the microphone array is also infinite in real situations, but it is very difficult or impossible to describe the summary sound wave arriving to each of microphones in the microphone array. Therefore, in this article based on human hearing model, a simplified geometric model (Fig. 2) is proposed for sound waves propagation to each of two microphones in a microphone array, assuming the existence only of one direct and two reflected sound waves arriving to each the left and to right microphones, respectively. In Fig. 2 are shown for simplicity only the trajectories of direct and two reflected sound waves arriving from the sound source to the left standing microphone \( M_1 \) in microphone array. The corresponding trajectories lengths of the direct, two reflected sound waves arriving to each the left microphone \( M_1 \) and reflection angles are indicated in Fig. 2 in the following way:

- \( l_d \) direct sound wave trajectory length;
- \( l_L = l^L_1 + l^L_2 \) first reflected sound wave trajectory length with two parts \( l^L_1 \) and \( l^L_2 \);
- \( \theta_1 \) reflection angle for first reflected sound wave;
- \( l_2^R = s^R_1 + s^R_2 + s^R_3 \) second reflected sound wave trajectory length with three parts \( s^R_1, s^R_2 \) and \( s^R_3 \);
- \( \theta_2 \) angle of the for second reflected sound wave.

The direct and two reflected sound waves with the lengths \( l_d, l^L \) and \( l^R_2 \) produce the corresponding sound signals \( S_d, S_1 \) and \( S_2 \) at the position of the left microphone \( M_1 \), arriving from the sound source \( S \). The explained above simplified geometric model (Fig. 2) can be considered also for the right standing microphone \( M_2 \), assuming, that it is necessary only to change the index \( L \) with the index \( R \) for the corresponding indication of trajectories lengths of the direct, two reflected sound waves arriving to each the right microphone \( M_2 \) and reflection angles.

Neded geometric parameters, shown in Fig. 2, are room dimensions (length – \( R_l \) and width \( R_w \)), distance between microphones – \( d \), sound source coordinates \( (x_{src} , y_{src}) \) and positions of the two microphones \( (x_{m1}, y_{m1}, x_{m2} \) and \( y_{m2}) \). Some of these geometric parameters, for example \( x_{m2} \) and \( y_{m2} \), can be calculated in the following way from \( x_{m1} \) and \( y_{m1} \), if they are defined:

\[
x_{m2} = x_{m1} + d, \quad (2)
\]

\[
y_{m2} = y_{m1}, \quad (3)
\]

where

\( x_{m2} \) and \( x_{m1} \) are \( x \)-coordinates of two microphones \( M_1 \) and \( M_2 \);
\( y_{m2} \) and \( y_{m1} \) – \( y \)-coordinates of two microphones \( M_1 \) and \( M_2 \);
- distance between them.

\[
\theta_n = \arcsin\left(\frac{L_{\text{per}}}{L_{\text{di}}}\right) \tag{7}
\]

In the similar way, using the ordinary geometric relations is possible to calculate also each of the angles of arrival \(\theta_n\) to the microphones \(M_1\) or \(M_2\) of reflected sound waves after \(n=1,2,3,...\) number of reflections from the walls of the room (Fig. 2). Therefore, the resultant signals \(S_{M1}\) and \(S_{M2}\), received from microphones \(M_1\) and \(M_2\), respectively are the sum of corresponding signals produced by the arrived direct and reflected sound waves to each of them, after \(n=1,2,3,...\) number of reflections:

\[
S_{M1} = k^1_0S^1_d + k^1_1S^1_l + k^1_2S^1_2 + ... + k^1_nS^1_n
\]

\[
S_{M2} = k^2_0S^2_d + k^2_1S^2_l + k^2_2S^2_2 + ... + k^2_nS^2_n, \tag{8}
\]

where \(k^1_0, k^1_1, k^1_2, ..., k^1_n\) and \(k^2_0, k^2_1, k^2_2, ..., k^2_n\) are the attenuation indexes of the corresponding parts (direct and reflected) sound signals produced by direct and reflected sound waves arrived to each of the microphones \(M_1\) and \(M_2\).

The attenuation indexes are important for description of the interaural intensity difference (IID) in algorithms of sound localization. Indirectly in the resultant signals \(S_{M1}\) and \(S_{M2}\) in equation (8) exist as the lengths of the reflected sound wave trajectories and corresponding time difference of arrival, as estimation of interaural time difference (ITD).

The values of the defined in equation (8) attenuation indexes \(k^1_0, k^1_1, k^1_2, ..., k^1_n\) and \(k^2_0, k^2_1, k^2_2, ..., k^2_n\) can be determined (in decibels) from microphone beamforming diagram (Fig. 3) of chosen type of microphones \(M_1\) and \(M_2\), for the corresponding values of angles of arrival \(\theta_n\) to the microphones \(M_1\) or \(M_2\) of direct and reflected sound waves after \(n=0,1,2,3,...\) number of reflections from the walls of the room.

For simplification, in the determination of the values of the attenuation indexes \(k^1_0, k^1_1, k^1_2, ..., k^1_n\) and \(k^2_0, k^2_1, k^2_2, ..., k^2_n\), is assumed to ignore the losses in reflection and absorption in the walls of the room.

For the calculations of the values of the lengths \(L_{\text{di}}\), \(L^L\) and \(L^R\) of direct and two reflected sound waves, is necessary to known the following additional geometric parameters:

\[
L_{\text{per}} = |y_{\text{si}} - y_{\text{sv}}| \tag{4}
\]

\[
d_i = |x_{\text{si}} - x_{\text{sv}}| \tag{5}
\]

where

\(L_{\text{per}}\) is the vertical component of the distance between sound source \(S\) and microphones \(M_1\) and \(M_2\) if it is assumed the placement of the two microphones \(M_1\) and \(M_2\) in a horizontal line, i.e. the equation (3) is satisfied;

\(d_i\) - the horizontal component of the distance between sound source \(S\) and microphone \(M_i\).

Using equations (4) and (5) is possible to calculate the distance \(L_{di}\) (for \(i=1,2\)) between sound source \(S\) and each of microphones \(M_1\) and \(M_2\) as follow:

\[
L_{di} = \frac{1}{2} \sqrt{d_i^2 + 4L_{\text{per}}^2} \quad \text{for } i=1,2 \tag{6}
\]

The angle of arrival \(\theta_n\) of the direct sound wave, for example to the microphone \(M_1\), can be calculated using equations (4) and (6):

Fig.2. The proposed geometrical model of microphone array based on the human binaural hearing.
The resultant signals $S_{M1}$ and $S_{M2}$ of microphones $M_1$ and $M_2$, derived by equation (8) can be used as input audio information in development and testing the algorithms of sound source localization. Simulating the appropriate situations and scenarios of positions for different cases for sound waves reflections from the room walls and in room constructions and dimensions is possible.

The proposed simplified planar geometrical model of sound waves propagation from sound source to the corresponding microphones in microphone array

To evaluate the characteristics of the proposed simplified planar geometrical model of sound waves propagation from sound source to the corresponding microphones $M_1$ and $M_2$ in microphone array based on the human binaural hearing model, in this article, the necessary simulations are carried out.

### III. SIMULATION OF THE GEOMETRICAL MODEL FOR SOUND SOURCE LOCALIZATION

The simulations proposed to test the proposed simplified planar geometrical model of sound waves propagation from sound source to the corresponding microphones $M_1$ and $M_2$ in microphone array presented in Fig.4.

It is based on the human binaural hearing model, arranged following the algorithm and then realized as the corresponding Matlab program.

In the first part of the algorithm and the corresponding Matlab program are defined some necessary input parameters for room dimensions - length $r_l$ and width $r_w$, also the distance between microphones $d$, coordinates of sound source ($x_{src}$ and $y_{src}$) and positions of the two microphones ($x_{m1}$, $y_{m1}$, $x_{m2}$ and $y_{m2}$).

With these defined parameters the different angles of arrival are calculated depending on the situations if there is no reflection, there is one reflection or there are two reflections.

The generation of microphone beamforming diagram (Fig. 3) of microphones $M_1$ and $M_2$ is done using the existing in Matlab class of microphones as elements of basic class Microphone array. In this case is chosen to use the class “Custom Microphone Element” as the type of microphones $M_1$ and $M_2$, which allow to define their beamforming diagram of the type “cardioid”, similar to the human ear hearing sensitivity and with appropriate orientation like in the human ears on human head. These conditions are involved in the following module of Matlab program and the definition of beamforming diagram of the left (microphone $M_1$) and the right (microphone $M_2$) orientation is prepared with chosen “+” or “-” in the lines 4 and 5 in the module of Matlab program:

```matlab
h = phased.CustomMicrophoneElement;
h.PolarPatternFrequencies = [500 1000];
h.PolarPattern = mag2db([...
0.5±0.5*cosd(h.PolarPatternAngles);...
0.6±0.4*cosd(h.PolarPatternAngles)]);
resp = step(h,[500 1500 2000],[0 0;40 50]);
plotResponse(h,500,'RespCut','Az','Format','Polar');
```

Fig.4. Simulation model is arranged as an appropriate algorithm.
The corresponding beamforming diagrams of the left microphone $M_1$ and the right microphone $M_2$, shown in Fig. 5, is possible to obtain the information about the values of attenuation indexes $k_0^1, k_1^1, k_2^1, \ldots, k_n^1$ and $k_0^2, k_1^2, k_2^2, \ldots, k_n^2$ in the correspondence of calculated from equation (7) angles of arrival $\theta_n$ to the microphones $M_1$ or $M_2$ of direct and reflected sound waves after $n=0,1,2,3,\ldots$ number of reflections from the walls of the room. The values of the attenuation indexes calculated by the human binaural hearing model are presented in Table I.

The initial parameters, needed for the calculations with equation (7) of the angles of arrival $\theta_n$ to the microphones $M_1$ or $M_2$ of direct and reflected sound waves after $n=0,1,2,3,\ldots$ number of reflections from the walls of the room are collected in Table I.

For example, in concrete simulation the values of the angles of arrival of direct sound wave, i.e. without reflection, is $\theta_0 = 89$ degree (marked with red arrow in Fig.6), which correspond to the index of attenuation $k_0^1 = -6.02 dB$ (Fig.6).

### IV. EXPERIMENTAL RESULTS

Experimental results after the simulations carried out with the proposed and developed simplified planar geometrical model of sound waves propagation from sound source to the corresponding microphones $M_1$ and $M_2$ in microphone array based on the human binaural hearing model are presented in the following way.

The results in Table II are for these calculated from Matlab program, based on the proposed algorithm, shown in Fig.4. The used above values of the angles of arrival $\theta_0, \theta_1, \theta_2$ and the attenuation indexes $k_0^1, k_1^1, k_2^1$ are calculated executing the developed for these simulations Matlab program angles of arrival $\theta_0, \theta_1, \theta_2$ (theta0, theta1 and theta2 in Table II) and the attenuation indexes $k_0^1, k_1^1, k_2^1$ ($k_0^1, k_1^1, k_2^1$ in Table II). The values of the attenuation indexes calculated by the Matlab program $k_0^1, k_1^1, k_2^1$ ($k_0^1, k_1^1, k_2^1$ in Table II) are substituted in the equation (8) for determination of the resultant signal $S_{M1}$ received from microphone $M_1$ as the sum of corresponding signals produced by the arrived to the microphone $M_1$ direct sound wave and reflected sound waves after $n=1,2,3,\ldots$ number of reflections:

$$S_{M1} = -6.02S_d^1 - 1.57S_1^1 - 28.01S_2^1$$  (9)

The similar equation to the equation (9) can be achieved for the right microphone $M_2$ using the similar results from Matlab program of the attenuation indexes $k_0^2, k_1^2, k_2^2$ substituted in the equation (8) for determination of the resultant signal $S_{M2}$ received from microphone $M_2$ as the sum of corresponding signals produced by the arrived to the microphone $M_2$ direct sound wave and reflected sound waves after $n=1,2,3,\ldots$ number of reflections.
TABLE II
THE CALCULATED IN MATLAB PROGRAM ANGLES OF ARRIVAL AND ATTENUATION INDEXES

<table>
<thead>
<tr>
<th>Name</th>
<th>Value</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>ens</td>
<td>-14.9621</td>
<td>-14.96...</td>
<td>-14.96...</td>
</tr>
<tr>
<td>d</td>
<td>0.2000</td>
<td>0.2000</td>
<td>0.2000</td>
</tr>
<tr>
<td>d1</td>
<td>0.1000</td>
<td>0.1000</td>
<td>0.1000</td>
</tr>
<tr>
<td>d2</td>
<td>-0.1000</td>
<td>-0.1000</td>
<td>-0.1000</td>
</tr>
<tr>
<td>h</td>
<td>&lt;1x1 phased,Custom...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>k0</td>
<td>-6.0206</td>
<td>-6.0206</td>
<td>-6.0206</td>
</tr>
<tr>
<td>k1</td>
<td>-1.5708</td>
<td>-1.5708</td>
<td>-1.5708</td>
</tr>
<tr>
<td>k2</td>
<td>-28.0138</td>
<td>-28.01...</td>
<td>-28.01...</td>
</tr>
<tr>
<td>l</td>
<td>4.5003</td>
<td>4.500...</td>
<td>4.500...</td>
</tr>
<tr>
<td>l1</td>
<td>4.5003</td>
<td>4.500...</td>
<td>4.500...</td>
</tr>
<tr>
<td>l2</td>
<td>4.5003</td>
<td>4.500...</td>
<td>4.500...</td>
</tr>
<tr>
<td>resp</td>
<td>(0.0,2.000,0,2.000:0.257...</td>
<td>0</td>
<td>0.4061</td>
</tr>
<tr>
<td>df</td>
<td>6</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>nw</td>
<td>4</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>thetah1</td>
<td>1.1538</td>
<td>1.1538</td>
<td>1.1538</td>
</tr>
<tr>
<td>thetah2</td>
<td>89</td>
<td>89</td>
<td>89</td>
</tr>
<tr>
<td>thetah3</td>
<td>49</td>
<td>49</td>
<td>49</td>
</tr>
<tr>
<td>thetah4</td>
<td>22</td>
<td>22</td>
<td>22</td>
</tr>
<tr>
<td>nxm</td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
<tr>
<td>nxm2</td>
<td>2.1000</td>
<td>2.100...</td>
<td>2.100...</td>
</tr>
<tr>
<td>nxr</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>nym</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>nym2</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>nyr</td>
<td>0.5000</td>
<td>0.500...</td>
<td>0.500...</td>
</tr>
</tbody>
</table>

V. CONCLUSION

The aim of the proposed in this article geometrical model of sound waves propagation from sound source with combination of a simple two microphone array based on the human binaural hearing model is fully achieved.

The equations derived from the proposed geometrical model lead to achieve the important descriptions of direct and reflected sound waves propagation after the defined simple cases of two numbers of reflections from the walls of the room, especially the resultant signals received from two microphones in the in the microphone array. These equations hold the information of interaural intensity difference (IID) as defined attenuation indexes and information of interaural time difference (ITD) as the lengths of the reflected sound wave trajectories. Both, the interaural intensity difference (IID) and interaural time difference (ITD) are used in the popular algorithms to achieve the sound source localization with the appropriate precision. The derived from the proposed geometrical model equations are applied in the developed algorithm for simulating the microphone array as the model of binaural human perception and sound source localization. Used microphones are with similar receiving characteristics like the human ears and their type is chosen to have the beam forming diagram of sensitivity similar to human ears beam forming diagram of sensitivity. The results from simulations are presented and show the ability to use in tasks of sound source localization the proposed geometric model as combination from descriptions of direct and reflected sound waves propagation with a two microphone array on the base of human binaural hearing model. The results, derived as equations and simulations, will be used for further deep analyses and will be extended and complicated for the different positions of sound source, increasing the number of sound waves reflections.

ACKNOWLEDGEMENT

This paper was supported by Technical University – Sofia inner program to support PhD research projects under Contract 142 PD 0018-07: “Development of methods and tools to locate audio sources in information and communication networks”.

REFERENCES