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

Viktor Hristov[1], Snejana Pleshkova[2]

Abstract – **Investigation of the human hearing system and attempts for its simulation are one the most popular themes nowadays. If the sound signals are received from a person it is easy to determinate the proper direction only with the help of the human ears and human brain. But if the human is replaced with 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 appropriates 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.

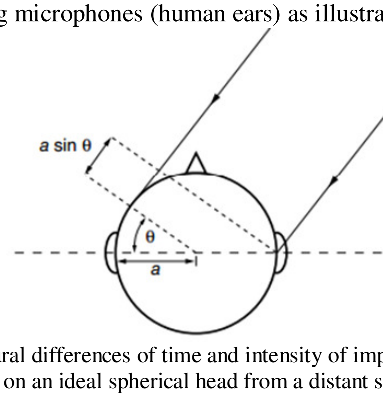

Fig.1. Interaural differences of time and intensity of impinging sound wave on an ideal spherical head from a distant source

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

[1]Viktor Hristov is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: hristov.viktor@gmail.com

[2]Snejana Pleshkova is with the Faculty of Telecommunications at Technical University of Sofia, 8 Kl. Ohridski Blvd, Sofia 1000, Bulgaria, E-mail: snegpl@tu-sofia.bg

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 θ radians can be approximately described with the following equation using diffraction theory [11, 12 and 13]:

$$\tau = \left(\frac{a}{c}\right) 2 \sin\theta, \tag{1}$$

where

τ is the interaural time difference (ITD);
a - the radius of the imaginary or artificial model of human head;
c = 343 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 perception. 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 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.

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} and 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;
- $l^L = l_1^L + l_2^L$ first reflected sound wave trajectory length with two parts l_1^L and l_2^L;
- θ_1 reflection angle for first reflected sound wave;
- $l_2^R = l_{21}^R + l_{22}^R + l_{23}^R$ second reflected sound wave trajectory length with three parts l_{21}^R, l_{22}^R and l_{23}^R;
- θ_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_2^R 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.

Needed 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} and 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:

$$x_{m2} = x_{m1} + d, \tag{2}$$

$$y_{m2} = y_{m1} = y_m, \tag{3}$$

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;

d - distance between them.

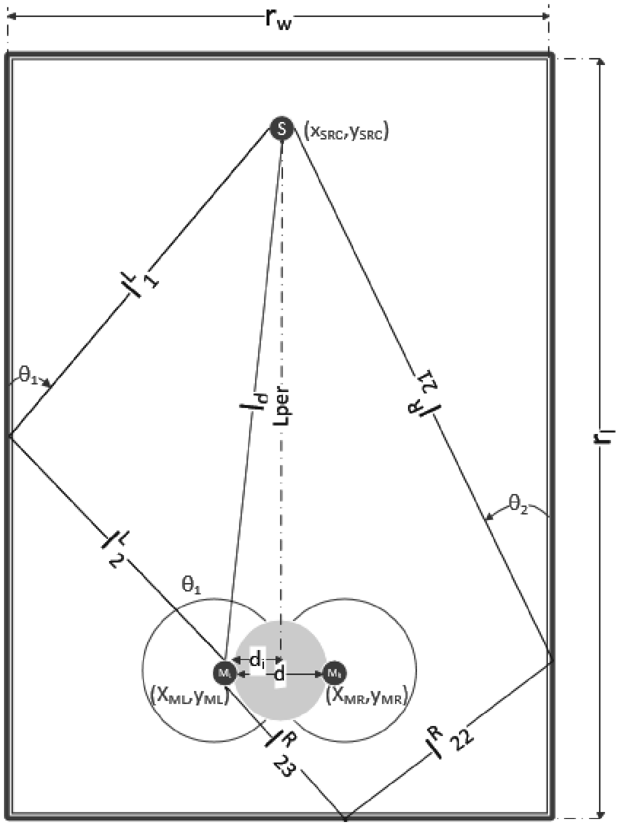

Fig.2. The proposed geometrical model of microphone array based on the human binaural hearing

For the calculations of the values of the lengths l_d, l^L and l_2^R of direct and two reflected sound waves, is necessary to known the following additional geometric parameters:

$$L_{per} = |y_{m1} - y_{src}| \tag{4}$$

$$d_i = |x_{m1} - x_{src}| \tag{5}$$

where

L_{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_1.

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:

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

The angle of arrival θ_0 of the direct sound wave, for example to the microphone M_1, can be calculated using equations (4) and (6):

$$\theta_0 = \arcsin(\frac{L_{per}}{L_d}) \tag{7}$$

In the similar way, using the ordinary geometric relations is possible to calculate also each of the angles of arrival θ_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:

$$S_{M1} = k_0^1 S_d^1 + k_1^1 S_1^1 + k_2^1 S_2^1 + ... + k_n^1 S_n^1$$
$$S_{M2} = k_0^2 S_d^2 + k_1^2 S_1^2 + k_2^1 S_2^2 + ... + k_n^2 S_n^2, \tag{8}$$

where

$k_0^1, k_1^1, k_2^1, ..., k_n^1$ and $k_0^2, k_1^2, k_2^2, ..., k_n^2$ 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 $k_0^1, k_1^1, k_2^1, ..., k_n^1$ and $k_0^2, k_1^2, k_2^2, ..., k_n^2$ 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 θ_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 $k_0^1, k_1^1, k_2^1, ..., k_n^1$ and $k_0^2, k_1^2, k_2^2, ..., k_n^2$, is assumed to ignore the losses in reflection and absorption in the walls of the room.

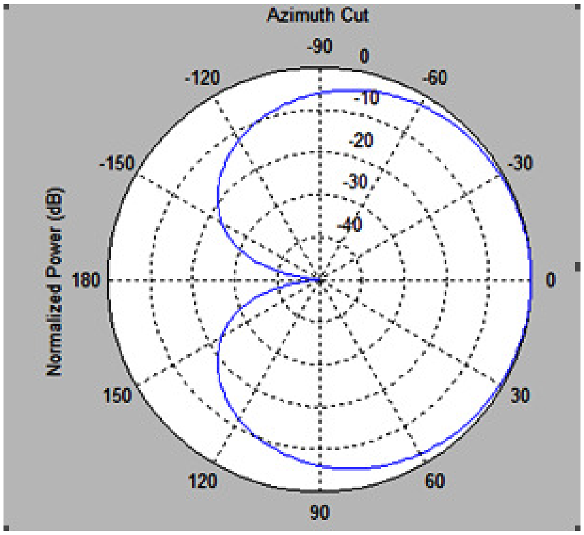

Fig.3. Microphone beamforming diagram

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:

```
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');
```

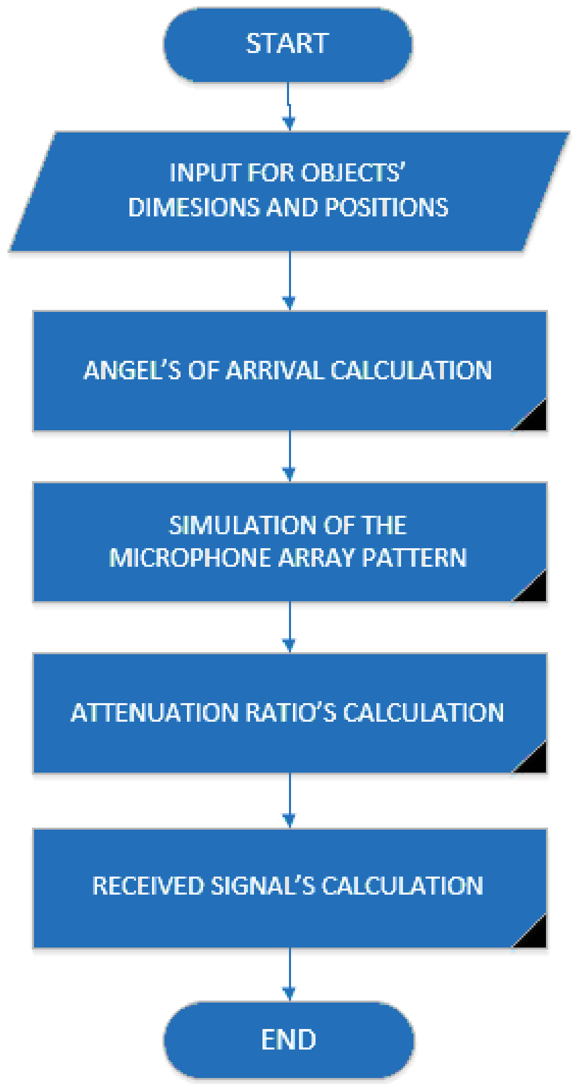

Fig.4. Simulation model is arranged as an appropriate algorithm.

The corresponding beamforming diagrams of the left microphone M_1 and the right microphone M_2 with their achieved necessary orientation are presented in Fig.5.

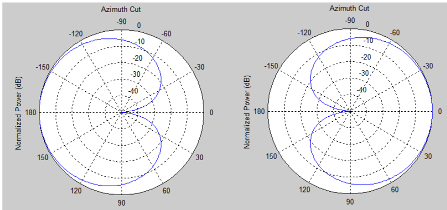

Fig.5. Beamforming diagram of the used microphones

From the generated 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 θ_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. The values of attenuation indexes $k_0^1, k_1^1, k_2^1, \ldots, k_n^1$ and $k_0^2, k_1^2, k_2^2, \ldots, k_n^2$ are necessary in the equation (8) for definition of the resultant signals S_{M1} and S_{M2} received from microphones M_1 and M_2 as the sum of corresponding signals produced from the arrived to each of two microphones direct and reflected sound waves.

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 initial parameters, needed for the calculations with equation (7) of the angles of arrival θ_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 are collected in Table I.

For example, in concrete simulation the values of the calculated angle 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 $k_0^1 = -6.02dB$ (Fig.6).

TABLE I
INITIAL ENVIRONMENT'S PARAMETERS

Room length – rl	4000 mm
Room width - rw	6000 mm
Source X coordinates - xsrc	2000 mm
Source Y coordinates - ysrc	500 mm
Mic1 X coordinate – xm1	1900 mm
Mic1 Y coordinate – ym1	5000 mm
Mic2 X coordinate – Xm2	2100 mm
Mic2 Y coordinate – Ym2	5000 mm

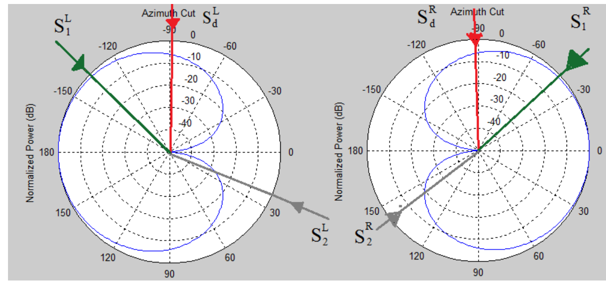

Fig.6. Experimental results

In the similar way, the angle of arrival of sound wave after one reflection is $\theta_1 = 49$ degree (corresponding to 131 degree in Fig. 6 and marked with green arrow in Fig. 6) with attenuation index $k_1^1 = -1.70dB$ (Fig.6). Also the angle of arrival of sound wave after two reflections is $\theta_2 = 22$ degree (not marked in Fig.6), which correspond to the attenuation index $k_2^1 = -28.01dB$ (Fig.6).

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, based on the proposed algorithm, shown in Fig.4. The results in Table II are for these calculated from 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, k_1, k_2 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, k_1, k_2 in Table II) are substituted in the equation (8) for determine 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…. number of reflections:

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

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 determine 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…. number of reflections.

TABLE II
THE CALCULATED IN MATLAB PROGRAM ANGLES OF ARRIVAL AND ATTENUATION INDEXES

Name	Value	Min	Max
ans	-14.9621	-14.96...	-14.96...
d	0.2000	0.2000	0.2000
d1	0.1000	0.1000	0.1000
d2	-0.1000	-0.1000	-0.1000
h	<1x1 phased.Custom...		
k0	-6.0206	-6.0206	-6.0206
k1	-1.5708	-1.5708	-1.5708
k2	-28.0138	-28.01...	-28.01...
l	4.5000	4.5000	4.5000
l1	4.5003	4.5003	4.5003
l2	4.5003	4.5003	4.5003
resp	[0,0.2000,0.2000;0.257...	0	0.4061
rl	6	6	6
rw	4	4	4
tgtheta1	1.1538	1.1538	1.1538
theta0	89	89	89
theta1	49	49	49
theta2	22	22	22
xm1	1.9000	1.9000	1.9000
xm2	2.1000	2.1000	2.1000
xsrc	2	2	2
ym1	5	5	5
ym2	5	5	5
ysrc	0.5000	0.5000	0.5000

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

[1] Yong Rui and Dinei Florencio, “New direct approaches to robust sound source localization”, Multimedia and Expo, 2003. ICME '03. Proceedings. 2003 International Conference,Volume:1, 2003, pp. I-737-I-740.

[2] Iain McCowan, “Microphone Array”, April 2001, pp. 61-85.

[3] J. Benesty, and J. Chen, “Study and Design of Different Microphone Arrays”, 2013, pp. 15-31.

[4] Muller, R.S., Howe, R.T., Senturia, S.D., Smith, R.L., and White, “Microsensors”, R.M, IEEE Press, New York, 1991, pp. 350-351.

[5] J. W. Strutt, “On our perception of sound direction”, Volume 13, pp. 214–232, 1907.

[6] C. Trahiotis, L. R. Bernstein, R. M. Stern, and T. N. Buell, “Interaural correlation as the basis of a working model of binaural processing”, Sound Source Localization, Springer Handbook of Auditory Research, Chapter 7, Springer-Verlag, Heidelberg, 2005, pp. 238–271.

[7] DeLiang Wang, Guy J. Brown, “Binaural sound localization”, Chapter 5, Book “Computational Auditory Scene Analysis: Principles, Algorithms, and Applications”, John Wiley & Sons, September 29, 2006, 147-187.

[8] M. Bodden, “Modelling human sound-source localization and the cocktail party effect.” Acta Acustica, Volume1, 1993, pp. 43-55.

[9] M. Wu, D. L. Wang, and G. J. Brown, “A multipitch tracking algorithm for noisy speech”, IEEE Transactions on Speech and Audio Processing, volume 11(3), 2003, pp. 229–241.

[10] C. Faller and J. Merimaa, “Sound localization in complex listening situations: Selection of binaural cues based on interaural coherence.” Journal of the Acoustical Society of America, Volume116 (5), 2004, pp. 3075–3089.

[11] G. F. Kuhn. “Model for the interaural time differences in the azimuthal plane.” Journal of the Acoustical Society of America, Volume 62, 1977, pp. 157–167.

[12] G. F. Kuhn, “Physical Acoustics and Measurements Pertaining to Directional Hearing". In W. A. Yost and G. Gourevitch, editors, Directional Hearing, Chapter 1, Springer-Verlag, New York, 1987, pages pp. 3-25.

[13] E. A. G. Shaw, “Acoustical features of the human external ear.” In R. H. Gilkey and T. R. Anderson, editors, Binaural and Spatial Hearing in Real and Virtual Environments, Chapter 2, Lawrence Erlbaum Associates (Mahwah, New Jersey), 1997, pp. 25–47.