Real Time Acoustic Localization of Elderly Persons in a Smart Home

Adel Thaljaoui, Damien Brulin, Thierry Val, Nejah Nasri

Real Time Acoustic Localization of Elderly Persons in a Smart Home

Adel THALJAOUI
University of Toulouse
UT2J, CNRS-IRIT-IRT
Toulouse, France
Adel.Thaljaoui@irit.fr

Damien BRULIN
University of Toulouse
UT2J, LASS-CNRS-N2IS
Toulouse, France
damien.brulin@laas.fr

Thierry VAL
University of Toulouse
UT2J, CNRS-IRIT-IRT
Toulouse, France
val@irit.fr

Nejah NASRI
University of Sfax
ENIS, ,LETI
Sfax, Tunisia
nejah.nasri@isecs.rnu.tn

***Abstract*—Due to the growing number of elderly people during the last decade, new needs came out especially in term of healthcare, resulting in more demand for self-care and home-based services. The emergence of smart technologies and home-based care has made this possible. In fact, smart home technology is designed to provide a place where old people can live independently as they wish by maintaining a control of every aspect of their daily life. Locating an elderly person in his home at any moment is one of the most concerns of smart home designers. The use of acoustic information captured in the smart home is one of the axes that can bring more possibilities in term of localization. This paper focuses on developing a sound source localization system using microphone arrays. Time difference of arrival (TDoA) using correlation technique is used for estimating the delay between two signals captured by two different microphones. The direction of arrival of the sound source can be obtained using this delay. Finally the sound source is positioned by adopting the geometric location method. This system has been tested by a set of experiments and the results are satisfactory**

Keywords— Home-care, localization, sound, microphone, TDoA

I. INTRODUCTION

Auditory perception in animals is an innate process that enables them to locate the source of any sound with a precision that can reach ±1° for some species [1]. This process is based directly on the operation of animal ears. In the case of Humans, the localization of a coming sound is carried out by combining the slightly deferred acoustic signals that arrive to each ear.

Scientific community has focused during the last decade on designing localization systems that could achieve high degrees of precision and reliability as offered by human auditory systems. Therefore, many studies have been performed especially on the field of robotics [2] [3] [4] [5] and [6].

One other concern of Scientifics was to build sound localization systems that could operate in indoor environments. Recent studies can be found in [7] [8] [9] [10] [11] [12] and [13]. Many of these studies consider a set of two spaced microphones hooked to a computer and calculate the spatial location of an emitted sound using a set of mathematic methods. One of the most common methods is the estimation of the Direction of Arrival of a sound through the computation of Time Difference of Arrival (TDoA) [14] [15] [16].

The aim of our works is to contribute on indoor localization by designing an hybrid localization system that combines acoustic, radio and light information in order to truck the position of an elderly person in a smart home at any moment. To do so, we had to focus on every one of these three localization approaches as a standalone system (acoustic, radio and light indoor localization systems) in order to figure out the potentially useful signal parameters that we can exploit when designing our hybrid system.

Our first studies deal with the acoustic localization in indoor environments. They consist on the construction and the design of hardware as well as the implementation of an application that returns the location of an acoustic sound at any moment. As a first step, we had used a set of two microphones [17] for the estimation of the angle of arrival of a sound source. The present paper introduces the second phase of our work on acoustic localization. We use a 4-microphone array for the determination of the geographic coordinates of an acoustic source based on the estimation of the Time Difference of Arrival (TDoA).

II. TIME DIFFERENCE OF ARRIVAL (TDOA)

Perception experiments conducted on humans and animals have been largely discussed and used by many researchers in the field of sound localization. As presented by Blauert [18] and Yost and Dye [19], the fundamental principles of sound localization combine both physics and geometry. Indeed, as the sound spreads out from its source, its power decreases. So, if the distances from the source to each receiver (ear) are unequal, an interaural level difference (ILD) will be noticed. Similarly, knowing the speed of sound in the air [20], and considering two spatially separated receivers, sound will travel different distances to each receiver and, therefore, will arrive at different times. Time difference of Arrival (TDoA) is one of the indicators that has been largely used in many localization systems such as radar and sonar for positioning radiating sources.

Let's consider a source that is emitting a sound $s(t)$ and an array of two spatially separated microphones array. Each one of the two microphones is receiving a signal ($s_1(t)$ for the first and $s_2(t)$ for the second). Due the distance between the two microphones (Fig. 1), a difference of time between the observations of the sound signal will be noted at each microphone, referred to as Time difference of Arrival. TDoA is computed using the spatial positions of the source and microphones.

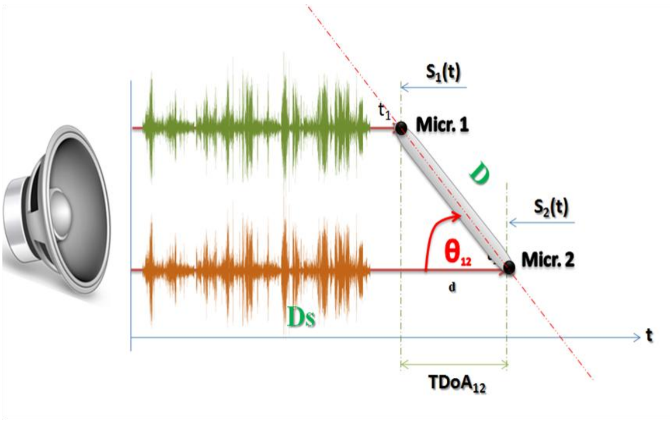

Fig. 1. Time Difference of Arrival (TDoA)

III. PROPOSED METHOD AND SPECIFIC ASSUMPTIONS

The present paper introduces a method of localization of an acoustic source within a frequency spectrum of 22Hz to 22KHz using an array of two microphone pairs. By localization, we mean the determination of the spatial coordinates of the source in a 2D space. The first step consists on estimating the time difference of arrival (TDoA) of the signals captured by two separated microphones. Then, the direction of arrival of the sound with respect to the microphones array is computed using trigonometry specifications. This first step was presented in [17]. The second step consists on localizing the acoustic source using a 4-microphone array. The process consists on merging the results obtained with each microphones couple in term of direction of arrival and then use a specific geometric positioning method in order to compute the geometric coordinates of the acoustic source in a 2D space.

Our method is designed under a set of assumptions: (1) only one acoustic source is located (2) We suppose that no Doppler shift will be detected meaning that there will be no obstacles that may change sound characteristics (3) Noise is White Gaussian (4) The set of microphones that we are using are Electret (5) we already know the spatial positions of microphones and (6) the distance between sound source and receivers is so big that we can consider that sound waves are planar.

IV. PHASE 1: ESTIMATION OF THE DIRECTION OF ARRIVAL OF AN ACOUSTIC SOURCE

This first step was published in [17]. We considered a system of two microphones and a sound source placed outside the perpendicular plane to the line passing through these two microphones and crossing through their middle (**Erreur ! Source du renvoi introuvable.**). Therefore, the wavefront of the sound source reach the two microphones with a slight time difference or delay, depending on the arrival angle of the source. The maximum value that this delay can tell about a sound source located on the same line through the two sensors (0° or 180°).

Our aim was to find the angle θ corresponding to the direction of arrival of the wave emitted by the sound source. We assume that $\theta \in]0, \pi[$.

We consider :

- D_s = distance between S and Mic 2
- $D_s - d$ = distance between S and Mic 1
- v : speed of sound in air that can be approximated as (Laplace formula):
- $v = (331 + 0.6 \times T)\ ms^{-1}$; where v is the speed of sound, in meters per second and T is the ambient temperature, in degrees Celsius. This approximation holds true for conditions near room temperature and pressure.
- TDoA also called τ..

In order to determinate the angle θ, we used trigonometry formulas for right triangle with d is the adjacent leg and D (distance between the two microphones) is the hypotenuse:

$$\cos\theta = \frac{d}{D} \qquad (1)$$

We know that for a distance d (Fig.1), le speed of the acoustic wave is calculated using this formula:

$$v = \frac{d}{\tau} \Rightarrow d = \tau \times v \qquad (1)$$

(2) becomes: $cos\ \theta = \frac{\tau \times v}{D}$

Therefore:

$$\theta = \cos^{-1}\frac{\tau \times v}{D} \qquad (2)$$

In order to find the direction of arrival (angle θ) of a sound wave, we proceed as follows:

- We first record the sound s(t), then we sample this signal according to a sampling frequency Fs (s(n)). The two sampled signals corresponding respectively to each microphone ($s_1(t)$ and $s_2(t)$) will be discerned.
- We then compute the cross-correlation between the two signals $s_1(t)$ and $s_2(t)$
- We search the temporal position of the maximum of the cross correlation and therefore we identify the delay τ
- We determine the angle θ

A. Acoustic sound recording

In this first step, we record the sound observed by the array of two microphones that we designed (Fig. 2). This signal is then sampled according to a sampling frequency Fe.

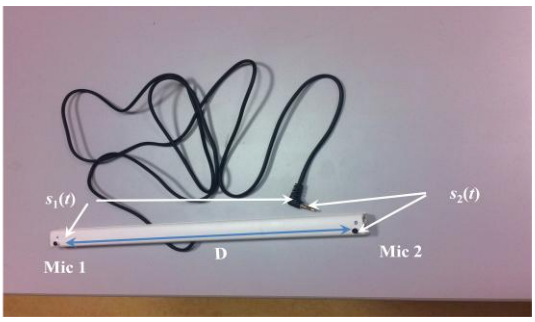

Fig. 2. Our designed 2-microphone array

The two microphones are connected to the computer through a stereo jack cable where left and right lane correspond respectively to the microphones 1 and 2. The two signals related to the two microphones are then returned.

B. *Cross-corrélation & TDOA (τ) Estimation*

There are many algorithms that can be used to calculate approximately the TDoA [21]. The cross-correlation (CC) method is one of the basic keys of this problem.

The estimated delay is obtained by finding the time-lag that maximizes the cross-correlation between the two observed signals [22]. Indeed, the signals observed by the two microphones are time shifted versions compared to each other. Therefore, in order to determine the direction of arrival of the sound source we need to detect the temporal point at which the two captured signals are at their maximum correlation. In this point, the two signals are at their closest match when they are superimposed on each other.

When the signals are continuous, the cross-correlation function can be expressed by:

$$C_{s1s2}(\tau) = \int_{-T}^{+T} s_1(t)s_2(t-\tau)dt \tag{3}$$

In our case, the observed signals are sampled, the expression that we use is as follows:

$$\mathrm{C}_{s1s2}(\tau) = \sum_{n=0}^{N-1} s_1(n)s_2(n-\tau) \tag{4}$$

where N is number of samples.

The idea is to scan the first captured signal according to an increasing time axis (i) while the second signal travels the time axis in the opposite direction $(-i)$. That's why the two compared signals should have the same length. In MATLAB, the cross-correlation of two signals, x and y ($\mathrm{length}(x) = \mathrm{length}(y)$), is a vector of length= $(2 \times \mathrm{length}(x) - 1)$ [23].

Once we have found the cross-correlation peak (delay τ between the two captured signals), we can deduce the temporal position of this maximum by applying the argument function on the maximized cross-correlation.

$$\tau = Argmax(C_{s1s2}) \tag{5}$$

V. Phase 2: Determination of the Source Spatial Coordinates

The aim of our acoustic source localization method is to be able to determine its 2D coordinates.

To do so, we consider the following layout:

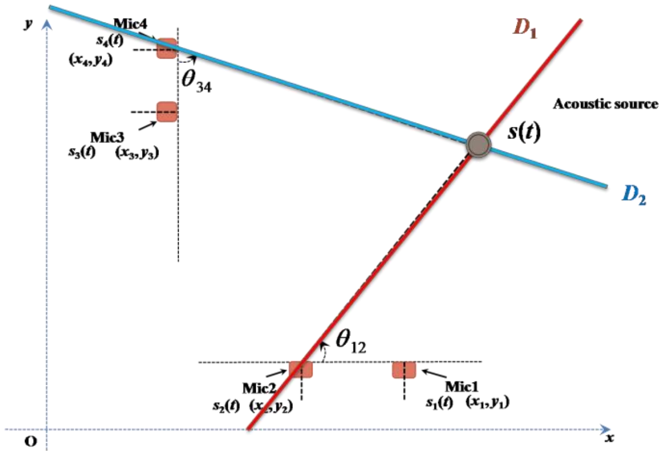

Fig. 3. Two 2-microphone pairs array

According to the assumptions we had already made, the spatial coordinates of the 4 microphones are known. Therefore, since we are able to estimate the angles θ_{12} and θ_{34} using the first step of our proposed method, we can calculate the spatial coordinates of the point of intersection of the two lines crossing the sound source with respectively microphone 2 and microphone 4 (D1 and D2).

Let's consider the two equations (7) and (8) as the respective equations of D1 and D2:

$$y = a_{12}x + b_{12} \tag{6}$$

$$y = a_{34}x + b_{34} \tag{7}$$

Where (a_{12}, b_{12}) and (a_{34}, b_{34}) are the respective couple of coefficients of each line.

To do so, we need to determinate the equations of these two straight lines. One of the mathematical solutions is to determine two points of each line and then compute their coefficients.

The coordinates of the acoustic source are than deduced through the resolution of the equation of intersection of the two equations above.

VI. Expérimentation

A. *Performances of the proposed method*

In order to evaluate the accuracy and the precision of the proposed method, we analyze the theoretic error E that we get while estimating the angle $\hat{\theta}$.

$$\mathrm{E} = \left|\theta - \hat{\theta}\right| \tag{8}$$

This error is calculated by varying the variance of the Additive White Gaussian Noise AWGN from 10dB to 10dB by 1dB steps. For more precision, we use the average of the error computed for 200 estimations of the angle θ.

Signal-to-noise ratio (SNR) is an indicator of the quality of information transmission by comparing the level of a desired signal to the level of background noise. SNR is defined as the ratio of signal power to the noise power (dB):

$$RSB = 10 \times \log_{10}\left(\frac{P_{signal}}{P_{noise}}\right) \qquad (9)$$

With P_{signal} = Power of the original signal ang P_{WGN} = power of the White Gaussian Noise.

Let's consider WGNinit an initial White Gaussian Noise (mean=0 and variance=1).

$$WGN_{init} = randn(1, N_{est}) \qquad (10)$$

Here Nest is the number of estimations of the angle θ (In our simulations we considered Nest=200)

The idea is to add to the original signal a White Gaussian Noise WGN with a null mean and a variance σ^2 : $WGN = \sigma \times WGN_{init}$

$$P_{WGN} = \sigma^2 \times P_{WGNinit} \ , \ RSB = 10 \times \log_{10}\left(\frac{P_{signal}}{P_{wgn}}\right)$$

$$\text{Therefore : } \sigma^2 = \frac{P_{signal}}{P_{WGNinit}} \times 10^{-\frac{RSB}{10}}$$

$$\sigma = \sqrt{\frac{P_{signal}}{P_{WGNinit}} \times 10^{-\frac{RSB}{10}}} \qquad (11)$$

The theoretical average estimation error that we obtained is about 5°, which is considered a good result for our case.

B. 2-Microphone array MATLAB Programming

In order to test our method, we required a set of hardware including an array of microphones and a recording device (typically a sound card) as well as specific software to be used for recording and processing the acoustic signals observed by the microphones in order to discern relevant information.

Therefore, we designed a block of two microphones (D = 30cm) that we fixed as shown in Fig. 4. The sound source was a Smartphone emitting a continuous sound. We developed then a MATLAB application (Fig. 5) that estimates the angle θ for many theoretical arrival directions as well as different values of Ds (distance between the sound source and the microphones array).

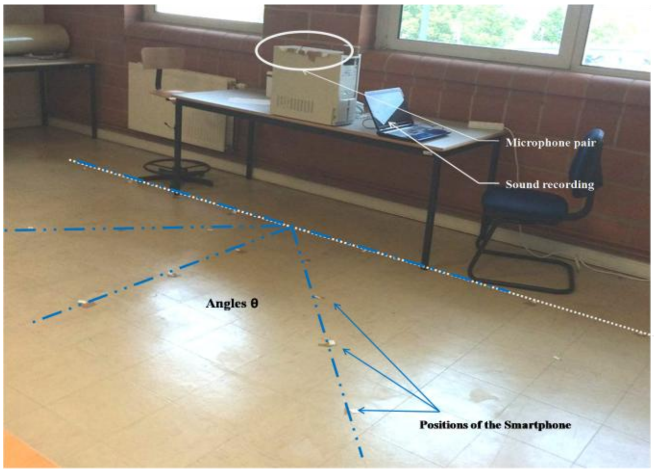

Fig. 4. First phase experimentations

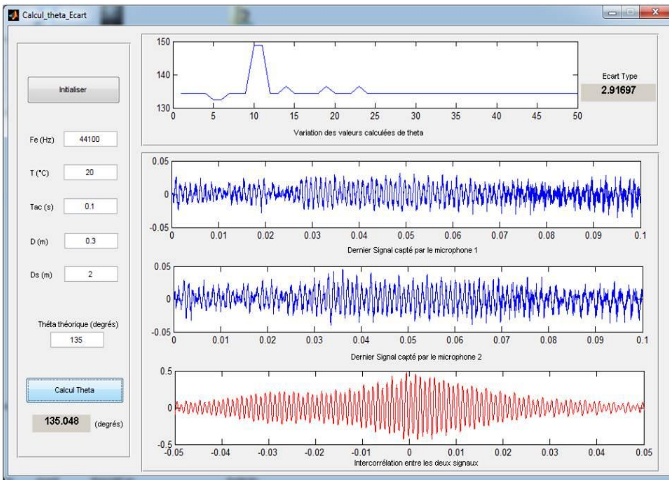

Fig. 5. Screen shot of the developed application for step 1

For more precision, we averaged the angle θ for 50 iterations(Fig. 6). For each one of the 50 estimations of the angle θ, we computed the standard deviation.

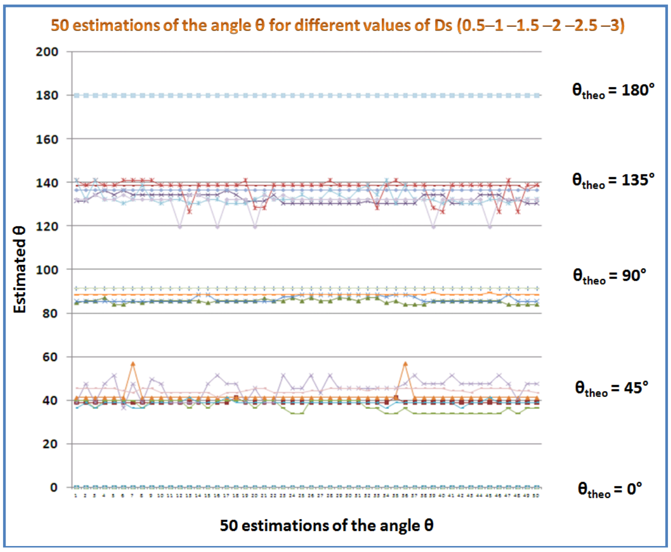

Fig. 6. 50 estimations of the angle θ for different Ds

According to the Nyquist Sampling Theorem, " If a time-varying signal is periodically sampled at a rate of at least twice the frequency of the highest-frequency sinusoidal component contained within the signal (in our case fmax=22Khz), then the original time-varying signal can be exactly recovered from the periodic samples". Thus, we chose a sampling frequency Fs = 44100Hz.

We varied Ds (0.5 - 1 - 1.5 - 2 - 2.5 - 3 m). The estimations of the direction of arrival are obtained for five theoretical positions of the angle θ (0 ° - 45 ° - 90 ° - 135 ° - 180 °) with a mean error of 3°, which is considered a good result.

C. Two pairs of Microphone array experimentation

The aim of the second part of our method is to determine the spatial location of the acoustic source by calculating its geographic coordinates in a 2D space.

As mentioned above, in order to perform this phase, we consider an array of two pairs of microphones, two computers, one Smartphone emitting a continuous sound.

Every pair of the microphones is hooked to a computer. We place the Smartphone in a already known position as shown in Fig. 7. We then run simultaneously in the two computers the software we already prepared in the first step in order to determine for each microphone pair the angle of arrival of the sound emitted by the Smartphone. The obtained values of the two angles are automatically stored in two separated excel sheets for later use.

We performed a set of tests using the designed hardware as well as the developed application (Fig. 8). This application reads the stored data from the two computers and returns the geographic position of the sound source using its estimated angles of arrival with the two pairs of microphones.

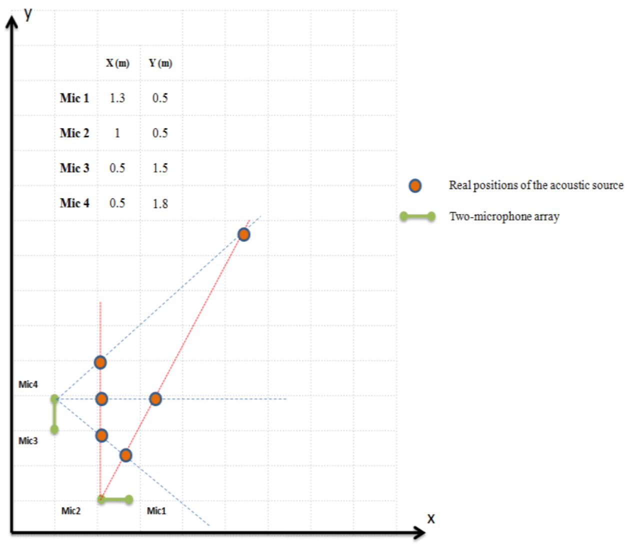

Fig. 7. Real positions of the acoustic source during the experimentation

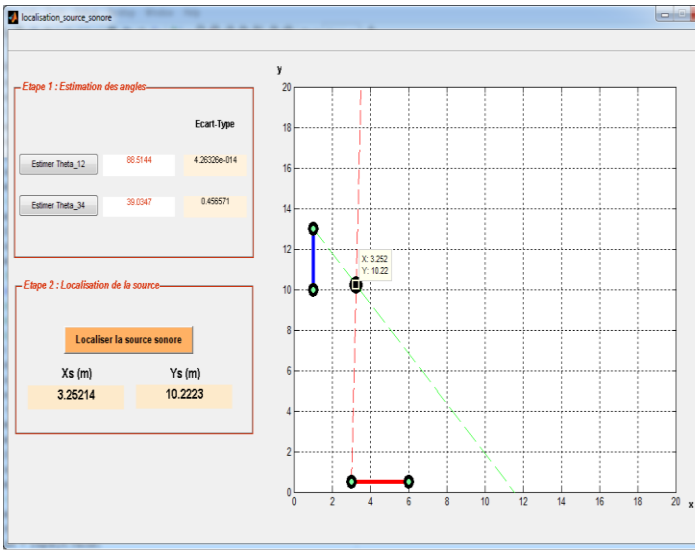

Fig. 8. Screen shot of the developed application for step 2

This application works satisfactorily (Fig. 9). In fact, the error between the estimated positions and the real ones can be explained by the assumptions we had made (see Table 1).

TABLE I. EXPERIMENTAL RESULTS

Real angles (°)		**Real positions (m)**		**Estimated angles (°)**		**Estimated positions (m)**	
θ12 theo	***θ34 theo***	***X theo***	***Y theo***	***θ 12 esti***	***θ 34 esti***	***X esti***	***Y esti***
90	45	1	1.3	94.15	39.03	0.946	1.25
90	135	1	2.3	95.9	135.048	0.831	2.132
90	90	1	1.8	84.078	89.554	1.13	1.79
45	45	1.4	0.9	44.61	41.77	1.35	0.84
45	90	2.3	1.8	53.28	87.38	1.92	1.735

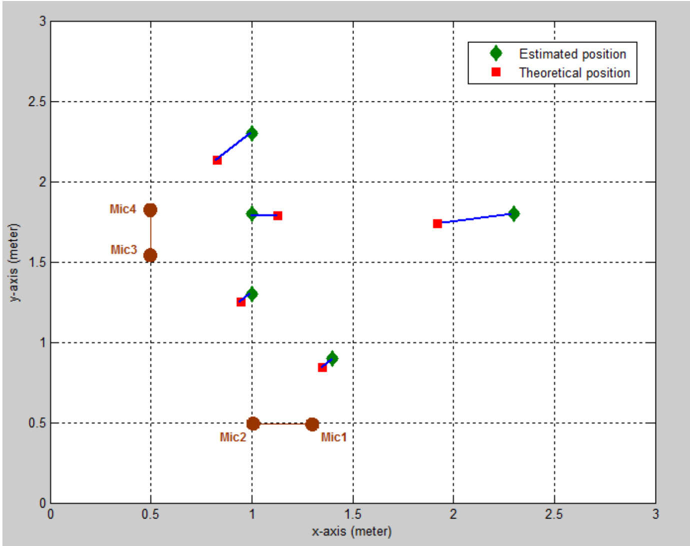

Fig. 9. Acoustic source estimated positions vs theoretical positions

VII. CONCLUSION AND FUTURE WORK

We have introduced in this paper, a method for positioning an acoustic source in a 2D space using an array of two pairs of microphones. We first computed for each pair of microphones the direction of arrival of the emitted sound (phase 1). Therefore, we estimated the time difference of arrival TDoA between the two sound signals observed by each microphone of the considered pair. In the second phase, we merged the results obtained simultaneously for each pair of microphones (in the first phase) in order to determine the spatial coordinates of the source.

In order to test our method, we carried out a set of experiments using the microphones array that we had already designed as well as a software that we developed using MATLAB. The results seem to be satisfactory.

Actually, as mentioned above, the method that we propose in this paper is based on a set of assumptions. One of these assumptions is that the environment we are working in represents no noise nor reverberations nor echoes. That is why, we used the cross correlation method to determine the delay τ which works good in such conditions. But, in real environment, this method is no longer suitable [24]. One of the common solutions is to consider the Generalized Cross Correlation (GCC) instead of the CC. Indeed, GCC transforms the signals from time domain to the frequency domain using Fourier transformation.

For a better accuracy, a weighting function or filter can be used in order to sharpen the peak of correlation (Fig. 10).

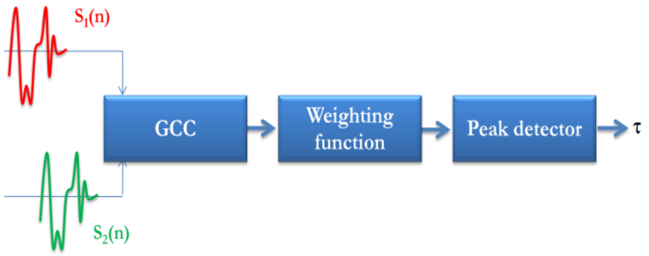

Fig. 10. Estimation of TDoA using the GCC method

The most common weighting functions are the phase transform "PHAT", the maximum likelihood "ML", the least mean square adaptive filter "LMS", SCOT, ROTH, ... As a continuation to our work, we are planning to compare these methods in order to figure out which one suits more to our case of study.

Besides, we intend to work jointly on radio and light localization methods to propose an hybrid localization system exploiting radio (Ibeacon BLE), light (LiFi) and audio (this present work) relevant information, in order to take advantage of these systems depending on the position of the person to locate.

REFERENCES

[1] HAROLD et al., Harold L. Hawkins, Teresa A McMullen, Arthur N. Popper, Richard R. Fay. Auditory Computation. Springer Handbook of Auditory Research. Springer 1996, pages 334-336., 1996.

[2] VALIN et al., Valin, J.-M.; Michaud, F.; Hadjou, B.; Rouat, J "Localization of Simultaneous Moving Sound Sources for Mobile Robot using a Frequency-Domain Steered Beamformer Approach, ICRA '04. 2004 IEEE International Conference, Volume: 1., 2004.

[3] YEOUN et al., Ji-Yeoun Lee; Su-young Chi; Jae-Yeun Lee; Minsoo Hahn; Young-Jo Cho, "Real-time sound localization using time difference for human robot interaction", World Congress, Volume # 16 | Part# 1., 2005.

[4] HÖRNTEIN et al., Jonas Hörnstein, Manuel Lopes, José Santos-Victor, "Sound Localization for Humanoid Robots -Building Audio-Motor Maps based on the HRTF", International Conference on Intelligent Robots and Systems, China., 2006.

[5] H-LIU et al., H. Liu, M. Shen Continuous Sound Source Localization Based on Microphone Array for Mobile Robots. IEEE International Conference on IROS, Taipei, Taiwan, 2010 :4332- 4339., 2010.

[6] LI et al., Xiaofei Li, Miao Shen, Wenmin Wang and Hong Liu, "Real-time Sound Source Localization for a Mobile Robot Based on the Guided Spectral-Temporal Position Method", International Journal of Advanced Robotic Systems., 2012.

[7] BIAN et al., G. D. A. Xuehai Bian, James M. Rehg, "Sound source localization in domestic environment", GVU center, Georgia Inst. of Technology., 2004.

[8] MINERO, P. Minero, "State of the art on localization and beamforming of an acoustic source", summary of Localization Techniques., 2004.

[9] KLEE et al., T. G. Ulrich Klee and J. McDonough, "Kalman filters for time delay of arrival-based source localization," EURASIP Journal on Applied Signal Processing, pp. 1–15., 2006.

[10] SECO et al., F. Seco, A. R. Jiménez, C. Prieto, J. Roa and K. Koutsou, A survey of mathematical methods for indoor localization, WISP 2009, 6th IEEE International Symposium on Intelligent Signal Processing, Budapest, Hungary, pp. 9-14., 2009.

[11] PERTILA, P. Pertila. Acoustic Source Localization in a Room Environment and at Moderate Distances. Tampereen teknillinen yliopisto. Julkaisu-Tampere University of Technology. Publication; 794., 2009.

[12] VACHER et al., M. VACHER, F. PORTET, A. FLEURY et N. NOURY : Challenges in the processing of audio channels for ambient assisted living. In 12th International Conference on E-Health Networking, Applications and Services, Lyon, France, 1-3 Jul. 2010., 2010.

[13] CANCLINI et al., A. Canclini, E. Antonacci, A. Sarti, and S. Tubaro. Acoustic source localization with distributed asynchronous microphone networks. Audio, Speech, and Language Processing, IEEE Transactions on, 21(2):439–443., 2013.

[14] MANDEL et al., M. I. Mandel, D. P. W. Ellis, and T. Jebara. An EM algorithm for localizing multiple sound sources in reverberant environments. In Proc. NIPS, pages 953–960, Cambridge, MA., 2007.

[15] LIU et al., R. Liu and Y. Wang. Azimuthal source localization using interaural coherence in a robotic dog: modeling and application. Robotica, 28(7):1013–1020., 2010.

[16] WOODRUFF et al., J. Woodruff and D. Wang. Binaural localization of multiple sources in reverberant and noisy environments. IEEE Trans. Acoust., Speech, Signal Process., 20(5):1503–1512., 2012.

[17] THALJAOUI et al., Adel THALJAOUI, Damien BRULIN, Thierry VAL et Nejah NASRI, "Localisation d'une source sonore par un réseau de microphones", Journées Nationales des Communications Terrestres, JNCT 2014 Toulouse-Blagnac - France., 2014.

[18] BLAURET, Jens Blauert, Spatial Hearing: The Psychophysics of Human Sound Localization, Cambridge, Massachusetts: The MIT Press., 1983.

[19] YOST et al., W. A. Yost and R. H. Dye, "Properties of Sound Localization by Humans," Neurobiology of Hearing: The Central Auditory System, R. A. Altschuler, et al., editors, Raven Press, Ltd, New York, pp. 389-410., 1991.

[20] LEHRMAN et al., R.L. Lehrman and C. Swartz, Foundations of Physics, New York, NY, pp. 297-299., 1965.

[21] CARTER, G. C. Carter: "Coherence and time delay estimation: an applied tutorial for research, development, test, and evaluation engineers", Piscataway, NJ: IEEE press., 1993.

[22] OMOLOGO et al., M. Omologo and P. Svaizer, "Acoustic event localization using a crosspower-spectrum phase based technique", Proceedings of IEEE International Conference on Acoustics,Speech, and Signal Processing (ICASSP'94), vol. 2, pp. 273–276., 1994.

[23] GU et al., Hung-Yan Gu, Shan-Siang Yang, A sound-source localization system using three-microphone array and crosspower spectrum phase, Machine Learning and Cybernetics (ICMLC), 2012 International Conference on (Volume:5)., 2012.

[24] VALIN et al., Jean-Marc Valin, Francois Michaud, Jean Rouat, Dominic L'etourneau, Robust Sound Source Localization Using a Microphone Array on a Mobile Robot, Intelligent Robots and Systems, vol. 2, pp 1228-1233., 2003.

[25] BALLOU et al., G. Ballou, Handbook for Sound Engineers: The New Audio Cyclopedia (Howard Sams, Indianapolis, IN), pp. 14-15., 1987.

[26] Coleman, Coleman P. An analysis of cues to auditory depth perception in free space. Psychol. Bull. 60:302–315., 1963.