Indoor Sound Based Localization: Research Questions and First Results

João Moutinho[1], Diamantino Freitas[2], and Rui Esteves Araújo[1]

[1] INESC TEC (formerly INESC Porto), Faculty of Engineering, University of Porto
[2] Faculty of Engineering, University of Porto
Rua Dr. Roberto Frias4200-465 Porto, Portugal
`{jnm,dfreitas,raraujo}@fe.up.pt`

Abstract. This PhD work has the goal to develop an inexpensive, easily deployable and widely compatible localization system for indoor use, suitable for pre-installed public address sound systems, avoiding costly installations or significant architectural changes in spaces. Using the audible sound range will allow the use of low cost off-the-shelf equipment suitable for keeping a low deployment cost. The state-of-the-art presented in this paper evidences a technological void in low-cost, reliable and precise localization systems and technologies. This necessity was also confirmed by the authors in a previous project (NAVMETRO®) where no suitable technological solution was found to exist to overcome the need to automatically localize people in a public space in a reliable and precise way.

   Although research work is in its first steps, it already provides a thorough view on the problem while discussing some possible approaches and predicting strategies to overcome the key difficulties. Some experiments were already conducted validating some initial premises and demonstrating how to measure the signal's time-of-flight necessary to infer on distance calculations.

Keywords: context-aware systems, sound based localization, TOA, TDOA, multilateration, trilateration, multipath, IPS, PL.

1   Introduction

One of the most popular research areas in ubiquitous or pervasive computing is the development of location-aware systems. These are systems in which electronic devices provide the users some kind of information or service depending on their location. The basilar component of a location-aware system is the location-sensing mechanism. This PhD work has the goal to develop an inexpensive, easily deployable and widely compatible localization system appropriate for indoor use, suitable for pre-installed public address sound systems, avoiding costly installations or significant architectural changes in spaces. Robustness, reliability and a within one meter precision are relevant requirements for this work and its consequent application.

   The most widely used location system in the world is the well-known GPS satellite-based navigation system. However, GPS doesn't work well without direct line-of-sight to satellites. Never the less, there are alternatives that use radio signals,

ultrasound, infrared, optical technologies, or a combination of few. They all have their own strengths and weaknesses and are very useful in their respective application domains. However these systems are not always suitable for indoor positioning since they usually imply the need of specialized hardware and infrastructure that is not typically available "in hands", turning them into prohibitively expensive solutions for wide deployment: "Good applications are those that achieve an adequate equilibrium between system requirements, technological advantages, and associated costs" [1]. Audible sound, as the signal to measure distance with its time-of-flight and consequently distance, is a possible solution for this necessary equilibrium. It's not uncommon that public indoor spaces have some kind of public address sound system with speakers uniformly distributed providing good sound coverage: train or subway stations, airports, shopping plazas, museums, etc. A subway station example is actually the starting point of this research work, as previous work has been developed by the authors in the NAVMETRO project [2] concerning acoustical localization.

2   Relationship to Internet of Things

Context-aware computing is becoming more and more necessary as technology evolves and the interaction between the user and computer is made simpler. This context-aware information can provide a specialized interface, automating tasks or may allow adjusting program settings to provide a personalized experience.

Location-aware systems address the acquisition of coordinates in a grid or at least distances to reference points. Outdoors, RADAR served for local demand and GPS satellite systems for global demand. These technologies have become ubiquitous in navigation and demonstrate the usefulness of having the user's position available. Indoor, where no satellites are visible and RADAR is not suitable due to walls and a smaller scale, no unanimous solution was yet achieved. The increasing use of the radio-frequency spectrum (Wi-Fi, Bluetooth, GSM, etc.), illumination problems, visual obstructions for cameras and the sound noise that exists in every indoor environment, complicate the problem enough so that no solution is perfect, cheap, precise or robust. Most of the existing solutions fail to be inexpensive to implement, precise, robust and costless to the user. To fill this gap, a proposed localization solution is presented based in audible sound.

The initial premise is that most indoor spaces already have public address sound systems and therefore using them would allow sparing relevant installation costs. The second premise is that everyone has a cell phone. These two assumptions together, will allow filling the low cost imposed requirement necessary for a successful wide deployment. Robustness and precision are next to follow and research and development is being conducted to overcome the several difficulties that using audible sound presents. Not many authors have approached the use of audible sound due to its intrusive characteristics in the acoustic environment, especially indoors. However, previous experience in other sound based applications [2] provided ideas and knowledge to expect overcoming the several difficulties that will present.

3   Research Questions

How to use sound to automatically localize a device indoors using a common sound installation without disturbing the acoustic environment? To successively accomplish answering this apparently simple question with complex issues, one may divide the problem into four possible sub questions.

- **How to localize a device?**

This is probably the most common subject on the research work in this area. It relies on the established framework of localization no matter the technology or the application. From the simplest example on getting the x,y coordinates (in the 2D problem) of a certain position, until the use of techniques to infer on the distance that goes from the loudspeaker to the receiver at the device's position as shown in Fig.1.
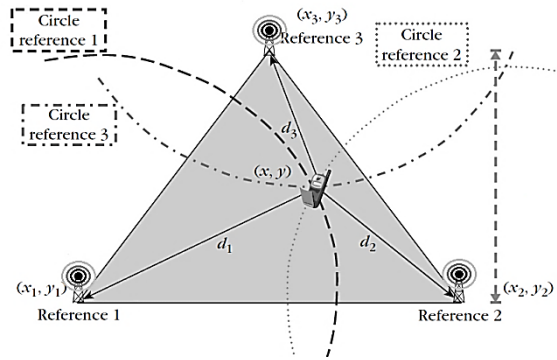

Fig. 1. Position as function of the distance from the sound sources

Localization algorithms considering this audible sound approach require measurements of range with reference fixed landmarks. Ranging can be achieved using different signal measurements such as time of arrival (TOA) or time difference of arrival (TDOA). Using sound instead of radio frequency signals can provide less numeric instability, since the speed of sound is very much smaller than the speed of light.

- **What kind of sound?**

The choice of sound to be used as the signal to obtain position of a certain device is one of the most important decisions. One of the most obvious problem constraints is the need to disguise the necessary stimuli in the "natural" environmental sound. This is the key issue in using audible sound. Increasing environmental noise may turn this position determination not desirable for the users. An approach exploring possibilities like using high frequencies (enough to pass the loudspeaker filter and still be captured by the device microphone), sound watermarking [3] or frequency or time masking, will be performed to choose the most adequate way of successfully locating the device while not disturbing the acoustic environment.

- **Which device to use?**

To explore the most convenient scenario regarding possible applications with persons and the stated low cost requirement, the possibility of using GSM cell-phones will be investigated. This would be convenient since almost all persons use one. However many problems may arise specifically due to technology and equipment limitations, starting on its microphone (the receiver transducer), its codec (to compress the audio's digital signal), network's packet delay jitter or simply the GSM communication delay from the distance to the antennas and operator.

- **Acoustic space?**

The propagation environment may cause signal shadowing and reflections that may introduce a multipath scenario in which reflected copies of a transient signal travel longer distances than the direct line of sight. Additive combinations of reflected signals cause field intensity variations or fading. These phenomena occur in acoustic environments, and their nature and characteristics are highly dependent on the operation frequency band, and even on temperature. These errors may be reduced through a wise choice on the excitation sound and with signal processing techniques like filtering, averaging practices, and multi-observation or redundancy.

4   State-of-the-Art

Indoor positioning systems (IPS) have been developed to provide location information of persons and devices. Personal networks are designed to meet the users' needs and interconnect users' devices equipped with different communications technologies in various places to form one network. Location-aware services need to be developed in personal networks to offer flexible and adaptive personal services.

4.1   Infrared Based Systems

One example is the "Active Badge Location System" where a device produces periodic signals that will be received by a spatially wide network of sensors that localize the device's signal [4]. The Active Badge system uses diffuse IR technology to realize location sensing by estimating the location of the active badges taken along by the persons. In each located place such as a room, one or more sensors are fixed and detect the IR signal sent by an active badge. The position of an active badge can be specified by the information from these sensors, which are connected by wires and forwards the location information of the tracked active badges to a central server. Although the prices of active badges and networked sensors are cheap, the cables connecting sensors raise the cost of the Active Badge system. Interference from florescent Light and sunlight raise problems.

4.2   Ultrasound Based Systems

Ultrasound technologies like the "Active Bats" [5] or the "Crickets" [6] are also used to localize a device. Active bats, has receivers placed in a square grid under the ceiling,

1.2m apart, that are interconnected by a serial link. Mobile units subject to localization are small devices called "Bats". They are equipped with an ultrasonic transmitter and a radio transceiver. A "Bat" that is to be located is addressed via the radio link. It then transmits a pulse of ultrasound at a known time. The Cricket location system and its extension, Cricket Compass, developed at the MIT, is a similar implementation [7]. But the general drawback of these approaches is that using standard transducers, ultrasound can usually propagate less than six meters. As a consequence a large number of precisely located reference units must be mounted.

4.3   Radio Based Systems

Due to the wide diffusion of Radio Frequency technologies, RSSI (Received Signal Strength Information) based localization techniques have been reported as a possible solution to low cost and easy to implement installations that allows tracking RF devices. It uses radio maps with a previously built signal strength model. The location of the mobile device is estimated by finding a best match from the signal strength model and the measurement with the use of deterministic and probabilistic techniques [8]. The attenuation rate is the rate α at which signal strength decreases over distance:

$$RSS \propto d^{-\propto} \tag{1}$$

As a practical rule, if $\alpha = 2$, then signal strength drops by 3dB every time distance doubles. This sub-linear attenuation rate means that the difference in signal strength between 1m and 2m is similar to the difference between 10m and 20m: exactly 3dB. Taking this into account, a constant level of noise can result in ever increasing error when signal strength is used to estimate distance; if RSSI noise is sufficient that we cannot tell the difference between 1 and 1.5m, one can't also tell the difference between 10m and 15m.

Other example approaches are the RFID base localization systems [9] that even though robust, have very low position resolution as it relies in having as many locations as RFID separate fields.

4.4   Optical Based Systems

Artificial vision based systems are also a possibility, and stereovision systems are pointing some new directions [10], but they are expensive due to the requirement of a high number of cameras and heavy processing.

Many false positives may be returned due to problems in the recognition of objects due to lighting conditions or simply due to occlusions and moving obstacles. Most techniques rely on the need to have previously trained conditions to correctly classify and track position. It is an approach that relies on controlled environments still not robust.

4.5   Inertial Based Systems

Inertial Navigation Systems (INS) together with RFID is also being explored to tackle the PL problem. The smartphone "advent" and their use of Micro-electromechanical

Systems (MEMS) allow integrating their information to get relative positioning. A possible example is an accelerometer based Positioning Scheme (APS) for tracking objects in indoor environments. The main idea of the APS is to compute an object's displacement by integrating acceleration and speed. Given the original coordinates of an object, the final position could be estimated according to the directional displacement. To overcome the "dead reckoning" problem of getting large accumulative errors, RFID techniques are used to calibrate position now and then [11].

4.6   Audible Sound Based Systems

Some audible sound-based techniques to infer on an object's or device's position are also found. Most of them use sound as a natural consequence of their operation, just like airplanes that produce noise that can be used to track them [12]. A 3-D IPS named Beep [13] was designed as a cheap positioning solution using audible sound technology. Beep uses a standard 3-D multi-lateration algorithm based on TOA measured by the Beep system sensors as a PDA or another device emits sound signals.

Other possibilities rely on having microphone arrays [14] to track some sound source positioning by AOA (angle of arrival techniques).

Another possible approach is a technique named "Acoustic Background Spectrum" where sound fingerprinting is employed to uniquely identify rooms or spaces in a passive way, just with the noise "fingerprint" of that space [15].

Even though using audible sound, these solutions aren't cheap, reliable and precise.

5   Research Contribution and Innovation

This research will investigate the use of watermarked coded audible sound in acoustic reverberant spaces with the objective of localizing a device that feedbacks the received sound to a central intelligence. Nothing was found in any previous work with these requirements and it will try to fill a void in the market of low cost, reliable and precise localization systems in public spaces. This research will have its contribution in the signal processing area where sound references will have to be processed to eliminate reverberations, noise and multipath disturbances. Coding and watermarking will also have to be explored and new techniques investigated so that information can travel along with the "natural" installation sound without being perceived by the listeners together with other concurrent information in the same acoustic space, i.e. multiuser. Exploring the use of the cell phone audio channel to close the loop will also require deep investigation so that no other hardware "tag" device is required therefore lowering the solution cost making more simple its wide dissemination.

6   Preliminary Results

Experiences were conducted with dozens of persons using only the human hearing and a cell phone to interact with a central intelligence [2]. It provided validation for

the use of audible sound in a public (noisy) space. Using pleasant bird sounds (chirp like in content) as cues provided the experimental proof that one may use sound in public spaces without disturbing the comfort of people in the same space while providing good navigation references [16]. Using advanced masking techniques (frequency and time) and sound watermarking [3] one can predict an even smaller impact, validating the use of sound as the signal to measure time-of-flight to infer on distance and therefore position using multi-lateration.
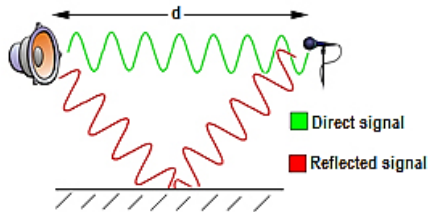

Fig. 2. Direct sound versus floor reflected sound between emitter and receiver

Results on signal processing were also already obtained in measuring time-of-flight of chirp acoustic signals in non-ideal scenarios (heavily reverberant acoustic environments). The correct wave front detection corresponding to the direct component of the emitted sound is also a compelling challenge that is necessary to overcome the precise measure time to correctly infer on distance. Using audible sound, a relatively low speed signal, is unquestionably another advantage since it requires less precision on time marking.

The front wave arrival detection instance is being performed with the use of cross correlation and Hilbert pre-envelope techniques, which has proven to be valid for measuring the time-of-flight while providing a measure for the confidence level of the measurement. In absolute distance measures, variable latency issues on the test platform are still compromising rigorous results. However, results are promising. Work-in-progress experiences are coding sound cues with embedded time code and using time difference of arrival. Results are however too preliminary.

7   Conclusions and Further Work

The presented work is still on its first developments. It is of capital importance to understand the boundaries of the problems in hands to have a good problem formulation. The research questions in hands are multidisciplinary and will require a deep and wide analysis on the subjects. A state-of-the-art was already performed and allowed to specify the requirements in hands. There was no similar approach found in the literature and therefore investigation is stepping on some new ground.

Preliminary results on a previous project showed that a similar principle with similar constraints, using a manual mechanism based in the human auditory system, provides excellent results. Turning this mechanism fully automatic do the user, while dealing with the non-ideal or controllable aspects, will be the challenge to overcome.

References

1. Munoz, D., Lara, F.B., Vargas, C., Enriquez-Caldera, R.: Position Location Techniques and Applications. Academic Press (2009)
2. Moutinho, J.N.: Sistema de Apoio Acústico à Navegação Pessoal. Master's Thesis, Faculty of Engineering, University of Porto (2009)
3. Lazic, N., Aarabi, P.: Communication over an acoustic channel using data hiding techniques. IEEE Transactions Multimedia 8, 918 (2006)
4. Want, R., Hopper, A., Falcão, V., Gibbons, J.: The Active Badge Location System. ACM Transactions on Information Systems 40(1), 91–102 (1992)
5. Hazas, M., Hopper, A.: A Novel Broadband Ultrasonic Location System for Improved Indoor Positioning. IEEE Transactions on Mobile Computing 5(5) (2006)
6. Priyanthaa, N.B.: The cricket indoor location system: PhD Thesis, 199 p. MIT (2005)
7. Linde, H.: On Aspects of Indoor Localization, University of Dortmund (2006)
8. Whitehouse, K., Karlof, C., Culler, D.: A practical evaluation of radio signal strength for ranging-based localization. ACM SIGMOBILE Mobile Computing and Communications Review 11(1) (2007)
9. Ni, L.M., Liu, Y., Lau, Y.C., Patil. A.P.: LANDMARC: Indoor Location Sensing Using Active RFID. In: IEEE International Conference on Pervasive Computing and Communications (2003)
10. Krumm, J., Harris, S., Meyers, B., Brumitt, B., Hale, M., Shafer, S.: Multi camera Multi person tracking easy living. In: Third IEEE International Workshop on Visual Surveillance, Dublin, Ireland, pp. 1–8 (2000)
11. Hsu, C., Yu, C.: An Accelerometer based approach for indoor localization. In: Symposia and Workshops on UIC 2009 and ATC 2009 Conferences, pp. 223–227 (2009)
12. Blumrich, R., Altmann, J.: Medium-range localization of aircraft via triangulation. App. Acoustics 61(1), 65–82 (2000)
13. Mandal, A., Lopes, C.V., Givargis, T., Haghighat, A., Jurdak, R., Baldi, P.: Beep: 3D Indoor Positioning Using Audible Sound. In: Proc. IEEE CCNC, Las Vegas (2005)
14. Atmoko, H., Tan, D.C., Tian, G.Y., Bruno, F.: Accurate sound source localization in a reverberant environment using multiple acoustic sensors. Meas. Sci. Technol. 19 024003 10pp (2008)
15. Tarzia, S.P., Dinda, P.A., Dick, R.P., Memik, G.: Indoor localization without infrastructure using the acoustic background spectrum. In: MobiSys 2011, pp. 155–168. ACM, NY (2011)
16. Moutinho, J.N., Freitas, D.S.: Sound cue selection methods for source localization in noisy environments. In: Proceedings Inter-Noise (2010)