



**HAL**  
open science

## Multimodal localization in the context of a medical telemonitoring system

Toufik Guettari, Paulo Armando Cavalcante Aguilar, Jérôme Boudy, Hamid Medjahed, Dan Istrate, Jean-Louis Baldinger, Imad Belfeki, Martin Opitz, Matthias Maly-Persy

► **To cite this version:**

Toufik Guettari, Paulo Armando Cavalcante Aguilar, Jérôme Boudy, Hamid Medjahed, Dan Istrate, et al.. Multimodal localization in the context of a medical telemonitoring system. EMBC 2010: 32nd Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Aug 2010, Buenos Aires, Argentina. pp.3835 - 3838, 10.1109/IEMBS.2010.5627688 . hal-01305112

**HAL Id: hal-01305112**

**<https://hal.science/hal-01305112v1>**

Submitted on 20 Apr 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Multimodal Localization in the Context of a Medical Telemonitoring System

Toufik GUETTARI, Paulo A. C. AGUILAR, Jerome BOUDY, Hamid MEDJAHED, Dan ISTRATE, Jean-Louis BALDINGER, Imad BELFEKI, Martin OPITZ, Matthias MALY-PERSY

**Abstract** - This paper addresses a localization system which is based on a combination of information from two modalities: a Smart Home Person Tracking (SHPT) composed of infrared sensors and an Audio Person Tracking (APT) which uses microphones able to estimate azimuth of acoustic sources. This combination improves precision of localization compared to a standalone or separated module. The localization software facilitates the integration of both SHPT and APT systems, to display the position in real time, to record data and detect some distress situations (some kind of fall). Results on implementation show good adaptation for Smart Home environments and a robust detection.

## I. INTRODUCTION

As time passes by, elderly persons become more or less autonomous and are exposed to the risk of domestic accidents, especially in huge residential places. In the context of medical telemonitoring [1-4], it is not sufficient if we only detect the distress situation of the Care Receiver person (CR). Another important parameter to monitor is the location of a person. This information can be used for ADL (Activity of Daily Life) measurements [5-6], especially to track the CR's activities in the course of one or several days. This gives the possibility to measure of the frequency of the usage of rooms and to inform the medical service in case of emergencies such as falls and abrupt changes of vital parameters of the CR. Therefore we decided to implement a localization platform based on several modalities.

Research on localization systems have been widely developed, especially in the field of robotics, based on several approaches, such as multi-sensor signals (sound, video and laser) [7-8], sound [9] or/and image [10] signal processing, Bayesian [11] and Markov [12] methods.

In this paper, we present a platform for localization of a person in his dwelling. Therefore the proposed platform is based on a combination of existing commercially available movement sensors (infra-red) and sound sensors (microphones) installed in this environment. This multimodal combination is based on an expert system which allows for efficient usage of these sensors and palliation of possible failing of the single modalities.

This work is supported by the European Commission in the frame of the Seventh Framework Program (FP7/2007-2011) within the CompanionAble Project (grant agreement n. 216487) [17].

T. Guettari, P. A. C. Aguilar, J. Boudy, J. L. Baldinger, Imad BELFEKI, Télécom & Management SudParis, Évry, France. E-mail: {toufik.guettari, paulo.cavalcante, jerome.boudy, imad.belfeki, jean-louis.baldinger@it-sudparis.eu}.

H. Medjahed, D. Istrate; Esigetel, Fontainebleau, France; E-mail: {dan.istrate, hamid.medjahed@esigetel.fr}.

Martin OPITZ, Matthias MALY-PERSY; AKG Acoustics Wien, Austria; E-mail: {martin.opitz, matthias.maly-persy@harman.com}.

In order to improve the precision of localization we have developed a scalable platform where we can add any other modality to help us to take a better decision, for instance additional modalities based on vision and accelerometers processing.

Section II describes each localization modality used in the platform, section III presents the proposed methodology and algorithm description and section IV provides experimental results obtained so far. Section V concludes and gives perspectives on these research works.

## II. LOCALIZATION MODALITIES

Our system of localization is based on the combination of two modes: the SHPT module (Smart Home Person Tracking) and the APT module (Audio Person Tracking).

The SHPT system is composed of a network of infrared sensors (IR) installed in the elderly person environment. The APT system is composed of a specific adaptive directivity-based microphone designed by AKG that permits to estimate the direction of arrival (the azimuth) of a given resonant source.

### A. SHPT Module : Infrared sensors network for presence detection

The SHPT system uses a set of infrared sensors [Legrand®] distributed and fixed in different rooms and areas of the CR's house [14]. As depicted in Fig. 1 this system allows to determine in which room (Room ID) the person is and precisely in which zone or area (Area ID) of the given room.

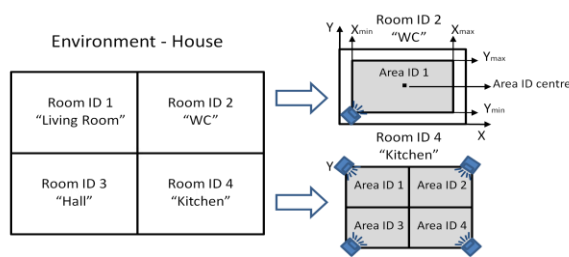


Fig. 1. SHPT system: Room ID and Area ID.

The CR's life place regroups several Room IDs (lounge, room, kitchen...) that are composed of one or several Area IDs. Each IR sensor determines an Area ID formed by a surface defined by coordinates  $X_{min}$ ,  $X_{max}$  and  $Y_{min}$ ,  $Y_{max}$ . Moreover if the zone of detection covered by an IR sensor is large, there will be less precision to localize the person in his/her lifeplace Therefore the use of large number of IR sensors increases the number of Area IDs and thereby

increases the precision of the system. For each detected Area ID, the person's position is estimated by the coordinates of the Area centre (Area ID centre, see Fig. 1).

The SHPT system is characterized by its reliability of movement detection. On the other hand, it presents a weak precision in the case where only one IR sensor is used in a room, which is often the case in most habitation (dwellings). In order to refine the localization process we propose its combination with the acoustic person tracking (APT) system.

### B. APT Module : Intelligent sound tracking sensor

The APT system is composed of microphones capable to estimate the direction of arrival (DOA or azimuth) of a given sound source. This technology is based on a new concept called CMT (coincident microphone technology) designed and developed by the AKG<sup>®</sup> Company [13]. It can give, with a reliable precision, the DOA of a person in a given area when he/she is speaking. This is very interesting for localizing the person in his/her environment.

The precision of localization can be influenced by echoes and reflections generated by the geometry and materials of the house. Another limit is that, for the detection of the azimuth, the person must speak or must produce a sound. However, the combination with the SHPT system previously described can increase the performance of localization. This will be shown in the sections below.

1) *Description of the Algorithm:* The CMT system is based on the simultaneous use of four coupled-microphones, one with an omni directional directivity and three with a cardioid directivity (hence directional).

As shown in Fig. 2 right side, each of the 3 first order gradient microphone transducers (MIC1, MIC2 and MIC3) is oriented in such a way that the direction of maximum sensitivity is 120 degrees apart from its neighbor transducer. The acoustic centers of these 3 transducers lie within a sphere of a few millimeters. The reference direction of 0 degrees is showing upside opposite to the cable plug.

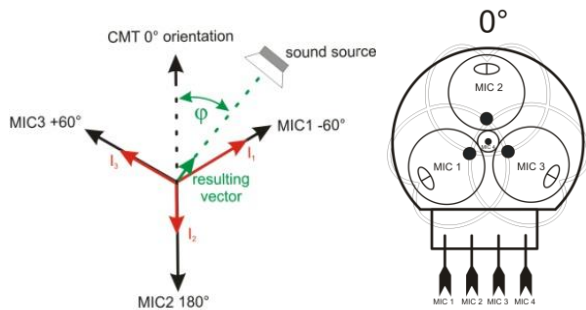


Fig. 2. Schema of azimuth localization based on momentary sound vector. Left side: Azimuthal direction. Right side: The CMT microphone.

As the position of the microphone is known, a fine position hypothesis can be generated taking into account the geometry of the apartment. That means, the origin of the sound has to be somewhere between the microphone and the next walls.

Echoes and reflections from the walls may influence the accuracy of localization. Sound absorbing walls and sound absorbing floor by using curtains and carpets will have a favorable influence on this accuracy. Such materials will minimize the reverberation time RT60. The value of RT60 should be smaller than 150 milliseconds.

The strategy selected for acoustic localization (Fig. 2, left side) is focusing on the azimuthal direction of that part of the sound field that causes the largest RMS (root mean square) value in the microphone signals (corresponding to the sound energy received). The digitized microphone signals are split in non overlapping blocks of a predefined length. For each block the unweighted RMS values are calculated for the 3 gradient microphones. These 3 RMS values ( $I_1, I_2$  and  $I_3$ ) are attributed to the respective direction of the 3 gradient microphones spaced by one another by 120 degrees. In this way 3 vectors are computed with azimuthal distances of 120 degrees. The vector lengths are proportional to the measured RMS values of the respective microphone signals. Ideally the argument of the summation vector  $\vec{I}(\varphi)_{sum}$  is showing in the direction of the loudest sound event, as in (1),

$$\begin{aligned} \vec{I}(\varphi)_{sum} &= \vec{I}(\varphi)_1 + \vec{I}(\varphi)_2 + \vec{I}(\varphi)_3 \\ (\varphi) &= \arg[\vec{I}(\varphi)_{sum}] \end{aligned} \quad (1)$$

The size of I is a measure for the reliability if the result. In the case of reflections, for example on walls, floor, ceiling, furniture etc. a deflection of the summation vector away from the sound source direction is detected. Additional background noise further decreases the accuracy of azimuth estimation.

### III. METHODOLOGY

The combination of the SHPT and APT systems permits to localize the person in a more precise manner while preserving the robustness of the positions estimated by the SHPT system when it is performing alone.

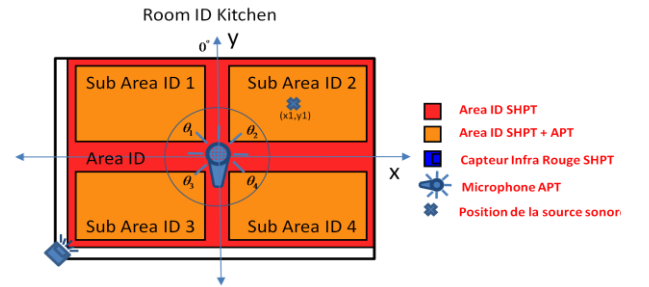


Fig. 3. Combination of the SHPT and APT systems.

Indeed, we consider the SHPT system as the reference modality because of its detection reliability (namely its detection continuity) and the APT system as secondary modality because its working depends on the CR's speech or sound event occurrence for the evaluation of the azimuth. In the ideal case the SHPT system determines the Area ID in which the person is present and the APT system refines its localization as shown in Fig. 3.

In the diagram of the Fig. 3, the microphone is placed in the middle of the room and allows, thanks to its ability to detect azimuth direction, to divide the Area ID detected by the unique IR sensor for the considered room (red zone) in several Sub Area IDs (yellow or clearer zones). In this case the person is localized in the Sub Area ID 2 for  $[0, 90^\circ]$  where the APT system gives an estimated azimuth between  $0$  and  $90^\circ$ . For each detected zone (Area ID or Sub Area ID), the person's position is estimated by coordinates  $(x, y)$  of the centre.

1) *Description of the Algorithm* : The algorithm is therefore based on expert rules allowing to make a “combined” decision targeting at a better precision than the one given by the SHPT system alone, which is considered as reference system under absence of APT measurements.

If only one sensor is in the room, which is the most frequent situation, each room is characterized by its room ID and is defined by its  $(x_r, y_r)$  centre coordinates. As the room is sub-divided in sub areas with Sub-Area IDs, we have the following main steps of the implemented rules based algorithm applied to each detected azimuth direction  $\varphi$  provided by the APT module :

**Step 1:** for  $1 \leq j \leq N$ , where  $N$  is the number of Sub AreaID, compute the angle  $(x, y)_j$ , where  $(x, y)_j$  is the centre of Sub Area ID  $j$ ;

**Step 2:** obtain the CR's position by  $(x_{cr}, y_{cr}) = ArgMin\{\varphi - \text{angle}(x, y)_j\}$  for all  $j$ , if  $\varphi - \text{angle}(x, y)_j < \gamma$ , where  $\gamma$  is a threshold related to the APT precision;

**Step 3:** if for all  $j$ ,  $\{\varphi - \text{angle}(x, y)_j\} > \gamma$ , then  $(x_{cr}, y_{cr}) = (x_r, y_r)$ .

In the case of noises or resonant non-speech events we consider the SHPT system only to localize the person (Area ID considered by default): this non-event index detection is provided by a separate module doing sound recognition (Anason [16]).

In the case of sensitivity problems and non detection of the person's movements by the IR sensors, the APT module will permit to estimate the CR's position at least based on his DOA.

#### IV. EXPERIMENTAL RESULTS

In the setting of the CompanionAble project our system was implemented in a Smart Home environment located in Eindhoven where different modalities for medical telemonitoring were installed. The installed platform is composed of 12 IR sensors for movement detection [Legrand®] and 2 CMT microphones [AKG®]. The whole configuration is shown in Fig. 4.

In the diagram of the Fig. 4, we have 12 Area IDs (red or dark zones) defined by the IR sensors and 8 Sub-Area IDs (yellow or clearer zones) determined from the evaluation of the resonant source direction by the microphone. Each IR sensor covers a maximum distance of 12 meters with an opening angle of  $90^\circ$ .

Experiments were performed in the living room of the Smart Home environment in Eindhoven as shown in Fig. 5.

This scenario consists of a person walking along the path (dashed line) in the living room. We tagged 7 positions

along the path (points A, ..., G) and computed the Euclidean distance between the tagged positions (black circles in Fig. 5) and the positions estimated by the systems SHPT and SHPT-APT combination (blue crosses in Fig. 5), as shown in Table 1.

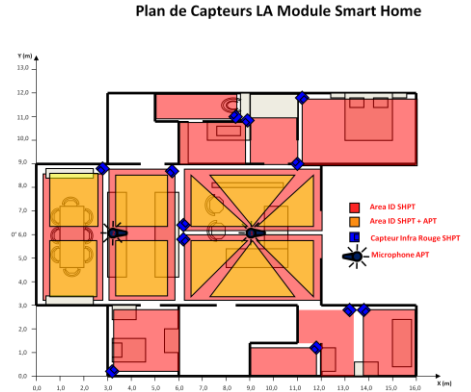


Fig. 4. Installation map of the localization system in Smart Home Eindhoven.

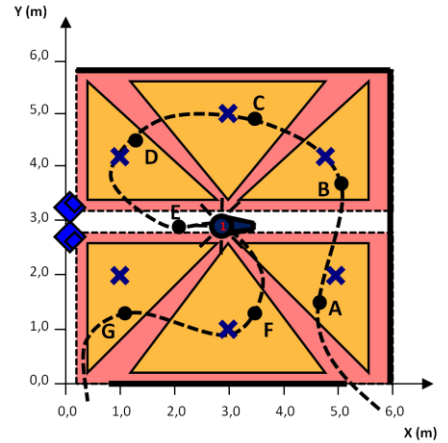


Fig. 5. Scenario for localization in the living room Smart Home Eindhoven.

TABLE I  
RESULTS OF EXPERIMENTATIONS

Position (x,y) real	Position (x,y) Estimated SHPT	Distance measured with SHPT	Position (x,y) Estimated SHPT+APT	Distance measured with SHPT+APT
A(5.0, 1.5)	(3.0, 1.0)	2.0616	(5.0, 2.0)	0.5000
B(5.0, 3.5)	(3.0, 5.0)	2.5000	(5.0, 4.0)	0.5000
C(3.5, 5)	(3.0, 5.0)	0.5000	(3.0, 5.0)	0.5000
D(1.5, 4.5)	(3.0, 5.0)	1.5811	(1.0, 4.0)	0.7071
E(2.0, 3.0)	(3.0, 3.0)	1	(3.0, 3.0)	1
F(3.5, 1.0)	(3.0, 1.0)	0.5000	(3.0, 1.0)	0.5000
G(1.0, 1.0)	(3.0, 1.0)	2	(1.0, 2.0)	1
Average error		1.4490		0.6724

We can observe along the whole CR's path that the combination of the two modalities increases the position estimation accuracy. On the average the measured distance between the estimated and the real position of the CR is reduced from 1.45 m to 0.67 m corresponding to an improvement of 54%.

The above-presented localization algorithm is

implemented in a LA (Localization Abstraction) module software whose graphical user interface is presented in Fig.6; this LA software allows us to visualize in real time the person's position in his/her living place. Positions are recorded in a database and are possibly associated to the person's activity measured through a sensor excitations ratio computed per minute; this permits us to define an average profile of the person's activity to be exploited for ADL measurements. This displacement profile is computed to detect potential pre-distress situations (profile ruptures).



Fig. 6. Interface of the localization software (LA Module).

## V. CONCLUSION AND OUTLOOK

The combination of two localization modules such as the SHPT and APT modules, has been implemented in a Smart Home type environment. This permits us to localize the person in a more efficient and reliable manner taking into account the respective limits of each separate modality or module. However, the well proven localization performance of the APT system can be influenced by echoes and reflections in the room. Another inconvenience of this system is the dependence of its working on the occurrence of sound to estimate the azimuth.

The integration of the two systems is possible thanks to the localization software that permits us to record data, to visualize positions in real time, to measure the activity and to detect some situations of distress of the person. Experimental results show a good adaptation in environments of Smart Home type and robustness in the detection of displacements.

The current LA software module has been programmed in an open and scalable way. For future works we propose to extend the current combination to other localization modalities provided by the CompanionAble application such as the Video Person Tracking [18] and the person's Behavior Analysis through the use of a wearable device. The application of multimodal fusion techniques as described in [15] will permit further increase of the precision and the robustness of the system.

## ACKNOWLEDGMENT

We thank Smart Home in Eindhoven and LEGRAND Company, partners of the CompanionAble [17] project for the disposal of the intelligent apartment and the infrared

sensors that allowed us to perform all these experimentations.

## REFERENCES

- [1] D. Istrate, E. Castelli, M. Vacher, L. Besacier, J. F. Serignat, "Information Extraction from Sound for Medical Telemonitoring", *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, no. 2, april 2006.
- [2] J. Boudy, F. Delavault, M. Muller, I. Farin, R. V. Andreão, S. Torres-Müller, A. Serra, D. Gaiti, F. Rocaries, C. Dietrich, A. Lacombe, F. Steenkeste, M. Schaff, M. Baer, A. Ozguler, and S. Vaysse, "Telemedicine for elderly patient at home: the TelePat project". In *International Conference on Smart Homes and Health Telematics*. Belfast, Northern Ireland, 2006.
- [3] W. Soudiene, D. Istrate, H. Medjahed, J. Boudy, J. L. Baldinger, I. Belfeki, and J. F. Delavault, "Multi-modal platform for in-home healthcare monitoring (emutem)", in *International Conference on Health Informatics (HEALTHINF 2009)*, Porto, Portugal, January 2009, pp. 381–386.
- [4] Rialle V, Lamy J. B, Noury N, Bajolle L., "Telemonitoring of patients at home: a software agent approach", *Comput Meth Prog Bio*, (2003), 257-268.
- [5] G. LeBellego, N. Noury, G. Virone, M. Mousseau, and J. Demongeot, "A Model for the Measurement of Patient Activity in a Hospital Suite", *IEEE Trans. Information Technology in Biomedicine*, vol. 10, n°1, January 2006.
- [6] H. Medjahed, D. Istrate, J. Boudy and B. Dorizzi, "Human Activities of Daily Living Recognition Using Fuzzy Logic For Elderly Home Monitoring", *FUZZ-IEEE 2009*, ISBN 978-1-4244-3597-5, pp. 2001-2006, 20-24 August 2009, Jeju Island, Korea.
- [7] J. Fritsch, M. Kleinhagenbrock, S. Lang, G. A. Fink, and G. Sagerer. "Audiovisual person tracking with a mobile robot". In *Proc. Int. Conf. on Intelligent Autonomous Systems*, pages 898–906. IOS Press, 2004.
- [8] Böhme, H.-J., et al., "An approach to multi-modal human-machine interaction for intelligent service robots". *Robotics and Autonomous Systems*, 2003. 44(1): p. 83-96.
- [9] J.-M. Valin, F. Michaud, J. Rouat, and D. Létourneau, "Robust sound source localization using a microphone array on a mobile robot," in *Proceedings International Conference on Intelligent Robots and Systems*, 2003.
- [10] S. Shiry, Y. Nakata, T. Takamori, M. Hattori, "Human Detection and Localization at Indoor Environment by Home Robot", *IEEE SMC Conference*, Nashville, USA, Oct. 2000.
- [11] W. Burgard, D. Fox, D. Hennig, and T. Schmidt. "Estimating the absolute position of a mobile robot using position probability grids". In *Proceedings of the Thirteenth National Conference on Artificial Intelligence*, Menlo Park, August 1996. AAAI, AAAI Press/MIT Press.
- [12] D. Fox, W. Burgard, and S. Thrun. "Markov localization for mobile robots in dynamic environments". *Journal of Artificial Intelligence Research*, 11:391–427, 1999.
- [13] K. Freiberger "Development and Evaluation of Source Localization Algorithms for Coincident Microphone Arrays" Diploma Thesis, Institute of Electronic Music and Acoustics (IEM), Univeristy of Music and Performing Arts Graz, Austria.
- [14] F. Steenkeste, H. Bocquet, M. Chan, B. Vellas. "Remote monitoring system for elders in a geriatric hospital". *Promoting Independence & Quality of Life for older persons: an international conference on aging Arlington 2 – 4 December 1999 (USA)*.
- [15] T. Guettari, H. Medjahed, J. Boudy, D. Istrate, I. Belfeki, J.L. Baldinger, "Etude de différentes Approches de Fusion. Une Application à la Télévigilance Médicale", *Congrès SFTAG '09*, 18-20 Novembre 2009, Troyes, France.
- [16] D. Istrate, E. Castelli, "Information extraction from sound for medical telemonitoring", *IEEE Transactions on Information Technology in Biomedicine*, vol. 10, no. 2, pp. 264-274, April 2006.
- [17] <http://www.companionable.net/>
- [18] Z. Zhang, "A flexible new technique for camera calibration". *Pattern Analysis and Machine Intelligence, IEEE Transactions*, 22(11):1330–1334, 2000.