A Sound Database For Health Smart Home

Leila Abdoune
Computer science department, University of Badji Mokhtar
Annaba, Algeria
lilouchetoday@yahoo.fr

Mohamed Fezari
Electronic department, University of Badji Mokhtar
Annaba, Algeria
Mohamed.fezari@uwe.ac.uk

Abstract—**This paper describes the construction phase of a corpus of everyday life sounds for a system of separation and classification of audio sources in a habitat; an application of telemonitoring of the elderly or disabled. We first presented the key concepts of the research area, and then we presented works and projects that address detection of distress situations and recognition of activities. Finally, we studied the different sounds used in an application of telemonitoring from which we are inspired to create our database.**

Keywords- health smart home; distress situation; audio channel; everyday life sounds.

I. INTRODUCTION

Activity recognition is an active research area [2, 3, 4, 5, 6] but despite this, it has not yet reached a satisfactory performance, or resulted in a standard methodology [1]. Among the objectives of this area we quote: recognition of distress situations of the elderly or disabled people in their habitats with the aim of their surveillance. The help of elderly and disabled people from a remote location with the use of information and communication technologies and artificial intelligence techniques and tools is part of what we call "Health Smart Homes" (HSH). Several tools were used to meet this objective such as infrared sensors, cameras and microphones, but given the high cost of these tools, the large number of these sensors and the need for interaction research tends towards the use of the audio channel. We cite as an example the system AUDITHIS [1] which allows the analysis of sound and speech in the HIS from 8 microphones.

The whole idea of our work is the realization of a system of classification and separation of audio sources in a habitat for an application of telemonitoring of the elderly or disabled people. However, several problems have arisen [4], such as the choice of parameters, choice of classification methods, sound quality (presence of noise), volume of information (several signals from different channels are acquired simultaneously), presence of noise, and finally the problem of defining a database of everyday life sounds which is the goal of this paper.

In this paper, we first presented the key concepts of the research area, and then we presented works and projects that address recognition of activities and detection of distress situations in a HSH. After that, we have described the overall system architecture and its main modules. Finally, we studied the different sounds used in a telemonitoring application from which we are inspired to create our database.

II. KEY CONCEPTS

In this section we present the definition of telemedicine and health smart homes which are the key concepts of our research area.

A. Telemedicine

The term "telemedicine" defines systems using video, audio, digital information and other communication tools coming from the technology to transmit information and data relating to medical diagnosis and medical treatments, and providing care and health services to patients located in remote physical environments [19]. Originally, the definition applied to advisory services delivered primarily through interactive video, but since the advent of the Internet and the multimedia, "telemedicine" has evolved into the "telehealth " which has a wider scope than telemedicine networks, and considers not only patient education, prevention of disease and therapeutic decision making, but also the administrative resources, the patient's physiological data and medical databases [20].

B. Health Smart Homes

The smart home is a scope of *Information and communications technology* (*ICT*), which has seen a great interest in the scientific community in recent years. It is a habitat equipped with information and communication facilities designed to work together to anticipate and meet the needs of occupants, working to promote their comfort, safety and entertainment while preserving their natural interaction with environment [23].

An important application of smart home is Health Smart Home. In the medical field, the main challenges of HSH are: home healthcare, telemedicine and sociomedical teleassistance. Priority targeted populations are the elderly or isolated people, chronic patients (physically disabled, cardiopathes) and persons under temporary medical supervision (pregnancies at risk, etc.).

III. WORKS AND PROJECTS RELATED TO RECOGNITION OF ACTIVITIES AND DETECTION OF DISTRESS SITUATIONS

Recognition of activities and distress situations are made through the information provided by the sensors installed in the HSH. A popular trend is to use the maximum of sensors to acquire more information. An opposite trend is to use the minimum amount of sensors [1] in such a way to design the most powerful system possible and at lower prices. Since we are interested in the second trend, we chose the audio channel,

although a small number of systems have the capacity to recognize the audio [26,27], we present here the systems that process the detection of activities in the habitat basing on the treatment of information of sound type.

In this section, we present a series of works that deal with the recognition of the activities and the detection of distress situations and their application in health smart homes.

Maunder et al. in [7] who built a database of sounds of daily life acquired by two microphones in a kitchen. They tried to differentiate between sounds such as telephone, a cup falling, drop a spoon, etc. Another group in [8] collected sounds in an office environment and tried unsupervised algorithms to classify the sounds of everyday life at work. The work presented in [9] is to recognize the distress situations at home in embedded situations using hardware at affordable prices (with standard sound cards and audio microphones). Another trend [10] is to perform speech synthesis, speech recognition, and the construction of a coherent dialogue with the person. Such research has applications in robotics, the aim is then to accompany the person and reduce loneliness.

The work presented in [1] is a complete sound recognition system to identify the different sounds in the apartment to recognize the activities of daily living currently performed, combined with a voice recognition system in French to find distress keywords within the measured signal. Thus, the CARE project [15], with the use of many sensors (localization, temperature ...), allows the recognition of activities: "going the toilet" and "Exit the apartment." In Britain, [16] created a model that distinguishes the work of preparing a hot or cold beverage with health activities based on the theory of evidence. [17] Presents the selection and arrangement of a set of sensors (infra-red detectors, door switches, microphones...) in an apartment, to classify a set of seven activities of daily living: rest, dress, casual, dining, communication, hygiene and disposal. [18] Uses microphones to detect a special distress: the fall. The solution is based on floor vibration and acoustic sensing, and uses a pattern recognition algorithm to discriminate between human or inanimate object fall events.

IV. OVERALL SYSTEM ARCHITECTURE

An application of sound recognition is composed of two main modules: the extraction module of the most relevant acoustic parameters of the processed signal, and the recognition module that associates the sound with the corresponding class (most likely); it is the identification phase of the input sound.

Our system is composed of two subsystems:

- The first subsystem deal with the detection of sounds in noise and their separation (mixture of sounds); input for the second subsystem.

- The second one is a classifier, which may be SVM, HMM based, or other type of classifier, used to classify the sounds resulted from the first subsystem. The figure Fig. 1 shows the overall architecture of the system.

The implementation of such a system requires first the construction of an everyday life sound database. For pattern recognition systems, there are always standard databases for the assessment of proposed systems. However, the construction of everyday life sound database in real conditions (environmental noise) is largely less explored.
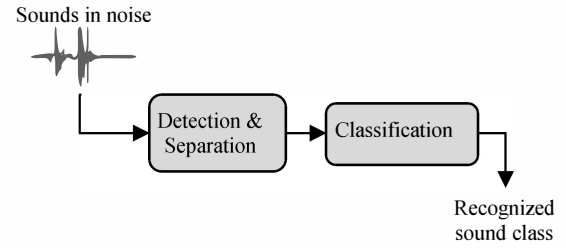

Figure 1.   : Main system modules

In effect, sounds acquired by the acquisition system and from different sources are acquired together. If we consider the number of microphones installed in the apartment is n, then we have n sound sources in a time t and therefore the information gained may be the same but with a variation in the signal power (e.g. the object falling sound in the kitchen can be acquired simultaneously by all microphones of the apartment: bathroom, living room, etc. but the sound of the fall in the kitchen is the strongest) and this if we consider that there is no noise in the apartment, or several sounds can be acquired simultaneously, e.g. TV sound in the sitting-room, the flow of water in the bathroom with the cough sound of the habitant. It is therefore important to identify the sound to take into account i.e. choose the sound the most provider of information.

The detection and separation phase is very important, it enables:

- Avoid duplication of information to process and thus, save time and storage space.

- A good determination of a distress situation if it is done with caution.

When the signals to be processed are selected, it is time to classify them. The classification is done at two levels:

- Classification of sounds like sounds of everyday life or speech.

- Recognition of the sound class or speech.

Depending on the resulting sound class or recognized speech, the system will send or not an alarm.

V. EVERYDAY LIFE SOUND DATABASES

In this section we present some existing databases of everyday life sounds and how these sounds are divided into categories.

In [12,21,11], Fleury, Vacher and Al. have defined six categories of sounds except the speech: human sounds which are related to the person, objects and supplies manipulation linked to the activity of the person, outside sounds, devices sounds, sounds of running water; this particular category

provides interesting information on activities such as disposal, hygiene, meal preparation. The last category of sounds is other sounds. Examples of each category are described in the table below.

TABLE I. SOUNDS CATEGORIES IN A HABITAT

Sound category	Examples
Human sounds	Cough, gargle, sigh, singing, whistling, wiping
Objects and supplies manipulation	Search a bag, Manipulation of the chair, handling a tray, sounds of footsteps, falling objects, sound of paper
Devices sounds	phone ringing, beep, TV
Sounds of running water	Handwashing, sink drain, flush, water flow
Outside sounds	thunder, rain
Other sounds	

In the Framework of the RESIDE-HIS project sounds are divided into two categories [13]: *Useful Sounds* (impulsive and short) as: Falling objects, broken glass, door slamming, etc., and *environmental noise* (long and stationary) as: the flow of water, hair dryer, electric shaver, etc.

Istrate in [14], for example, achieved a database of everyday life sounds for an application of telemonitoring of elderly and disabled people in a habitat. Its methodology for achieving this database is based on the one used in the speech database with regard to labeling and description files. The sound database is organized as follows: 15% of sounds recorded in the studio, 15% of the sounds recovered from a CD of effects for films [24] and 70% of sounds coming from the CD "Sound scene database in Real Acoustical Environments "(RWCP) laboratory ATR [25] . In [22], he has defined seven classes of sound: door slamming, broken glass, ringing phone, footstep sounds, screams, sounds of dishes, door lock. The corpus contains both the sounds associated with distress situations such as broken glass, falling objects, screaming, but also the usual sounds like the sounds of footsteps, door slamming, and sounds of dishes. The database must also contain the environmental sounds, such as TV, radio, hair dryer, water flow, considered as noise [14].

Everyday life sounds database of [14] consists of the following sounds: stapler sounds, clapping, moving the chair, human sounds (sneezing, yawning, laughing, snoring, coughing), opening a pressure vessel, crease or decryption of paper, footstep sounds, punch sounds, slamming different doors (entrance door, cabinet, refrigerator), electric shaver, hair dryer, door lock, dishes, chair and book fall, screams, water flow in sink, shower and glass, ringtones, health smart home background noise.

Information source of sound sensors can be [14]:

- Direct source for distress situation:
  - Recognition of distress call
  - Detection of suspicious sounds (objects falling, screaming)
  - Long absence of sound activity during the day
- An indirect source: Detection of a sequence of everyday life sounds by analyzing over a long period can be a symptom of pathology as nocturnal urinary disorders. The lack of activity during the day can be an indicator of a grave distress situation.

VI. HOW TO CREATE OUR DATABASE?

A. *Description of the Sound Database*

Sounds may be speech or everyday life sounds. In order to create our database, we have divided everyday life sounds into two categories:

- **Normal Sounds** (related to usual activities): ex. door slam, door lock, door opening, phone ringing, footstep sounds, and sounds of dishes.

- **Critical sounds** (possibility of existence of a of distress situation): ex. broken glass, falling objects, screaming.

Normal sounds are divided into two categories too:

- **Useful sounds:** which may help us to detect distress situation when combined with other information.

- **Disturbing sounds:** considered as noise, like television and radio sounds.

As it is described in [14], a part of our DB (database) is obtained by recording different types of sounds; a second part will be obtained from commercial CDs, the third part by extracting sound from mp3 or videos files. The last part will be recovered from existing DBs like the ones created in [1,14,13].

Our principle of choice of sounds is to expand the class of sounds as much as possible to ensure a high rate of recognition and thus reduce the error rate on one side, the other side, exploit this database to achieve other objectives such as the recognition of activities. Consequently, it should be noted that even if the sound does not appear significant, it is involved in the recognition of activities when combined with other information.

The list of sounds we need for our application is summarized in the table below:

B. *The Recording Parameters*

We chose a sampling frequency of 44.1 KHz to faithfully reproduce the signal after digitization. We chose the ". Wav" format for sound files because it is a standard format, and can be read by various software, in addition to that its conversion to other formats is easy [14].

The signal to noise ratio SNR of the recordings will take several values that vary between 10 and 40 to 70 dB. The exact

values are not yet defined. The length of the sound files is 20s. This length was chosen taking into account the maximum length of sounds to be treated and seeing that the initialization time for some algorithms is ≈ 5s [14]. The figures below, mention examples of some recorded sounds using Matlab environment.

TABLE II.   SOUNDS PRODUCED IN THE HABITAT

| Sound			Speech	
Critical sounds	Normal sounds		Daily speech	Distress words
	Useful sounds	Disturbing sounds		
screaming, falls of objects, glass breaking, a long silence	Sounds of dishes, doors closure, doors opening, door slamming, sounds of footsteps, Water flow, coughing, yawning.	TV, Radio, phone ringing, Electrical devices sounds, external noise.		I need help, For help, Aïe!, It hurts, …
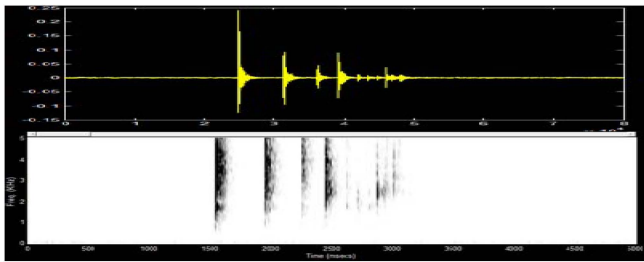
Figure 2.   : Temporal evolution and spectrogram of falling object
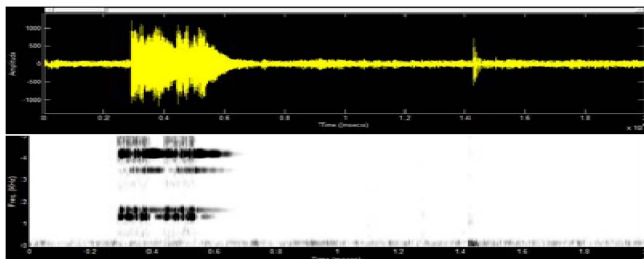

Figure 3.   Temporal evolution and spectrogram of a phone ringing

VII.   CONCLUSION

In this paper, we have made an overview on Health Smart homes and telemedicine, and then we presented various works related to recognition of activities and detection of distress situations. Subsequently, we described the existing databases created to validate systems of telemonitoring of elderly or disabled people via the detection of distress situations. Finally, we presented our gait to creating sound database which is in progress, knowing that the detailed steps of database creation were not covered here because it is not yet completed.

REFERENCES

[1] M. Vacher., A. Fleury., F. Portet, J.-F. Serignat, N. Noury .Complete Sound and Speech Recognition System for Health Smart Homes: Application to the Recognition of Activities of Daily Living. Dans New Developments in Biomedical Engineering, Domenico Campolo (Ed.) (2010) pp. 645 -- 673 - http://hal.archives-ouvertes.fr/hal-00422576.

[2] F. Albinali, , N. Davies, & A. Friday .Structural learning of activities from sparse datasets, 5th IEEE Int. Conference on Pervasive Computing and Communications. 2007.

[3] S. Dalal, M. Alwan, R. Seifrafi, S. Kell & D. Brown. A rule-based approach to the analysis of elders' activity data: detection of health and possible emergency conditions, AAAI 2005 fall symposium, workshop on caring machines: AI in eldercare.

[4] F. Duchêne , C. Garbay, & V. Rialle. Learning recurrent behaviors from heterogeneous multivariate time-series. Artificial Intelligence in Medicine 39: 25–47. 2007.

[5] T. Duong , D. Phung, H. Bui, & S. Venkatesh. Efficient duration and hierarchical modeling for human activity recognition, Artificial Intelligence 173(7–8): 830–856. 2009.

[6] A. Fleury. Détection de motifs temporels dans les environements multi-perceptifs – Application à la classification des Activités de la Vie Quotidienne d'une Personne Suivie à Domicile par Télémédecine., PhD thesis, University Joseph Fourier, Grenoble. 2008.

[7] D. Maunder, E. Ambikairajah, J. Epps, & B. Celler. Dual-microphone sounds of daily life classification for telemonitoring in a noisy environment, Proc. 30th Annual International Conference of the IEEE-EMBS 2008, pp. 4636–4639.

[8] A. Harma, M. McKinney & J. Skowronek. Automatic surveillance of the acoustic activity in our living environment, Proc. IEEE Int. Conference on Multimedia and Expo ICME 2005, pp. 634–637.

[9] D. Istrate, M. Vacher, & J.-F. Serignat, Embedded implementation of distress situation identification through sound analysis, The Journal on Information Technology in Healthcare 6(3): 204–211. (2008).

[10] S.-y. Takahashi, T. Morimoto, S. Maeda, & N. Tsuruta, Dialogue experiment for elderly people in home health care system, Text Speech and Dialogue (TSD) 2003.

[11] M. Vacher. "Analyse sonore et multimodale dans le domaine de l'assistance à domicile''. Mémoire HDR, Octobre 2011.

[12] A. FLEURY, M. VACHER, F. PORTET, P. CHAHUARA, N. NOURY. "A Multimodal Corpus Recorded in a Health Smart Home" Multimodal Corpora : Advances in Capturing, Coding and Analyzing Multimodality, LREC, 18-21 May 2010, Malta, pp. 99-105. 105. 99-105.

[13] E. Castelli, M. Vacher, D. Istrate, L. Besacier, J. F. Sérignat, "Habitat Telemonitoring System Based on the Sound Surveillance", "1st International Conference on Information Communication Technologies in Health, Greece , 141-146, July 2003.

[14] D. M. Istrate. ''Détection et reconnaissance des sons pour la surveillance médicale''. Thèse INPG, spécialité SIPT de l'Ecole Doctorale EEATS, Grenoble, 16 décembre 2003.

[15] B. KRÖSE, T. VAN KASTEREN, C. GIBSON, T.VAN DEN DOOL, CARE : Context Awareness in Residences for Elderly, International Conference of the International Society for Gerontechnology, Pisa, Tuscany, Italy, June 4-7 2008.

[16] X.HONG, C. NUGENT, M. MULVENNA, S. MCCLEAN, B. SCOTNEY, Evidential fusion of sensor data for activity recognition in smart homes, Pervasive and Mobile Computing, vol. 5, no 3, 2008, p. 1-17.

[17] A. Fleury, N. Noury, M. Vacher. Application des SVM à la classification des Activités de la Vie Quotidienne d'une personne à partir des capteurs d'un Habitat Intelligent pour la Santé, hal-00422566, version 1 - 7 Oct 2009.

[18] D. Litvak, , Y. Zigel, & I. Gannot. Fall detection of elderly through floor vibrations and sound, Proc. 30th Annual Int. Conference of the IEEE-EMBS 2008, pp. 4632–4635. 2008.

[19] M. M. Maheu, A. Allen,, 20 Aou. 2007, http://telehealth.net/glossary.html

[20] S. Bonhomme, Méthodologie et outils pour la conception d'un habitat intelligent. (Methodology and tools for the design of an intelligent home.), these, pp 19.

[21] M. Vacher, F. Portet , A. Fleury, , and N. Noury. "Development of audio sensing technology for ambient assisted living : Applications and challenges". International Journal of E-Health and Medical Communications, 2(1) :35–54. 2011.

[22] D. ISTRATE, M. VACHER, E. CASTELLI, C.-P. NGUYEN. "Sound Processing for Health Smart Home". 2nd International Conference on Smart homes and health Telematics, Singapour, 15-17 septembre 2004, 8 p.

[23] R. Harper. Inside the smart home. Springer, London, 2003.

[24] Sciascia, S. (1992). CD – bruitages, Vol.3.

[25] Partnership, R. W. C. (1998-2001). CD – Sound scene database in real acoustical environments. http//:tosa.mri.co.jp/sounddb/indexe.htm.

[26] Stäger, M., Lukowicz, P., Tröster, G.: Power and accuracy tradeoffs in sound-based context recognition systems. Pervasive and Mobile Computing, vol. 3 (3), pp. 300–327. (2007).

[27] Wang, J. C., Lee, H. P., Wang, J. F., Lin, C. B.: Robust Environmental Sound Recognition for Home Automation. Automation Science and Engineering, IEEE Transactions on, vol. 5 (1), pp. 25–31, Jan. 2008.