Automated Audio Data Monitoring for a Social Robot in Ambient Assisted Living Environments

Rosa Ma Alsina-Pagès[a], Joan Navarro[b] and Enric Casals[a]

[a]GTM - Grup de Recerca en Tecnologies Mèdia, La Salle - Universitat Ramon Llull, Barcelona (Spain)
[b]GRITS - Grup de Recerca en Internet Technologies & Storage, La Salle - Universitat Ramon Llull, Barcelona (Spain)

Abstract— Human life expectancy has steadily grown over the last century, which has driven governments and institutions to increase the efforts on caring about the eldest segment of the population. Although this concern was initially addressed by building larger hospitals and retirement homes, these facilities have been rapidly overfilled and their associated maintenance costs are becoming far prohibitive. Therefore, modern trends attempt to take advantage of latest advances in technology and communications to remotely monitor those people with special needs at their own home, which boosts their life quality and has very few impact on their social lives. Nonetheless, this approach still requires a considerable amount of qualified medical personnel to track every patient at any time. The purpose of this paper is to present a social robot for assisted living that tracks patients status by automatically identifying and analyzing the acoustic events happening in a house. Specifically, we have taken benefit of the amazing capabilities of a Raspberry Pi together with a Nao robot to collect data inside a house and send it in realtime to the medical center. Conducted experiments verify the feasibility of our approach and open new research directions in this domain.

Keywords— Assisted living, human robot interaction, social robots, remote therapy, audio recognition

1. Introduction

Increasing the average human life expectancy is one of the greatest achievements of modern society [1]. This opens several challenges for the public (and private) health community since the number of patients to take care of has also raised accordingly. Additionally, the way of how those people who need special care are attended is changing significantly. In fact, despite the latest efforts of governments and institutions on building bigger and more modern medical facilities, they rapidly get overfilled due to the ever-raising number of people they have to serve. Nowadays, it is intended that the elderly stay at home for two reasons: on one hand, it is better for their health, while not suffering from severe deteriorations, and on the other hand, is much cheaper to health services.

Thankfully, technology can contribute to address this problem by enabling medic staff to monitor (and attend) patients when they are at home (also referred to as ambient assisted living [2]), which reduces the personnel costs and enhances their social interactions. From a general point of view, ambient assisted living consists of mining the preferred living environment of patients with intelligent devices able to track their status to improve their life quality. Nowadays, there are several social iniciatives to put robots to elderly people homes not only to cover certain routine tasks, but also with chatbot applications, to accompany them, to help them remember to take the tablets and also to make them use interactive games that help to keep their cognitive abilities.

This work further exploits the concept of social robots in ambient assisted living and presents an application, consisting of a hardware and a software, that infers the in-home context of a house using sound data. This platform is composed of a Raspberry Pi to work with any standard robot. Specifically, the Raspberry Pi acoustically analyzes all the sounds happening inside the patient's home using machine learning and data mining techniques. When the system determines that an event corresponds to an emergency (e.g., scream, silence for long periods of time, someone falling down) the robot itself activates an alarm. The social robot is also used to obtain information about what has happened in the tests, record it, and send it together with the activated alarm to the cloud to be evaluated. It is worth mentioning that this application has been designed in such a way that it is independent form the social robot used in the tests.

The remainder of this extended abstract is organized as follows. Section 2 reviews the related work on environmental sound recognition. Section 3 discusses the basics of the human robot interaction used. Section 4 elaborates on the technical details of the proposed platform. Section 5 discusses future work directions and concludes the paper.

2. Related Work

Environmental sound recognition is a hot topic today for several applications [3]; from bird sound detection to surveillance applications, several approaches can be found in literature to extract the features from the sound, and classify the source of the noise with previous training with a corpus.

One of the most popular applications nowadays of audio event recognition is its use in the smart home [4], especially when developing the system with the disabled people needs in mind. The challenges around the design of a health smart home [2] based on audio event classification: *i)* the degree of dependency of the disabled person, *ii)* the quality of life improval by means of automatization of processes and finally *iii)* the distress situations recognition and the activation of the preasigned protocols. In this paper we only adress the third of the challenges, despite in the literature [5] several solutions are given for all the others.

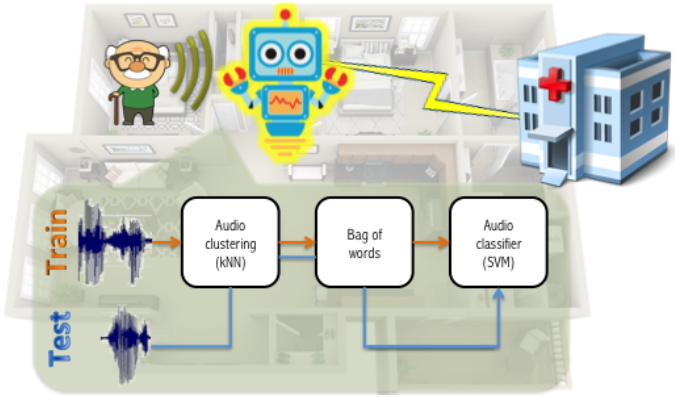

Fig. 1. Block diagram of the software modules that compose the proposed social robot for ambient assisted living

Interest in detecting in-home sounds was alive from the beginning of this technology; in 2005, Chen [6] was monitoring the bathroom activity only from the sound information. Afterwards, with research not detailed in this work, robust environment sound recognition motors were designed in 2008 [7]. Even audio scenes were classified evolving the same type of technology [8]. Finally, several works can be found about audio study in a smarthome to help doctors the early diagnose of dementia diseases for the elder [9].

3. Human Robot Interaction

Interaction between humans and robots applied is currently being used to fight loneliness in elderly people aging at home [10]. It is an instrument that can accompany them in routine daily tasks such as giving greetings in the morning, informing them of what is the weather like or reminding them to take the medication. In this regard, if there is no response in the proactive robot interaction, we get the first indicator of potential health problems. If we add the benefits of active listening of what happens at home, the patient's condition can be related with the acoustic context, and then activate the previously established protocol if necessary. In this regard, we have used Wilma, a Nao robot brought from the USA that La Salle R&D has at its disposal. It allows us to collect all the information about what happened at home and activate the alarms accordingly (e.g., when a fall is detected, the protocol is to first ask the person if everything is ok).

4. System architecture

The social robot works with a Raspberry Pi model 3 which has wi-fi connectivity and an ARM quad-core architecture very suitable to conduct audio processing. We have connected a microphone to the Raspberry and programmed a data miming architecture to process audio data (see Fig. 1). Its behavior is detailed as follows.

First, the system splits the audio sample in subsamples of 20 ms. and extracts the relevant audio features for each subsample. One of the most used feature extraction algorithms in automatic sound recognition is Mel/Frequency Cepstal Coefficients [11]. Despite in literature they are widely used in speech technologies, most of the audio recognition systems settle the use of these coefficients as baseline in terms of feature extraction [12]. When all features are extracted, a k-Nearest Neighbors (kNN) system is run [13].

Next, we apply the bag of words technique to obtain a fixed size vector for each sample. Each component of the vector corresponds to the number of audio subsamples that match a cluster of the kNN. With this information, a Support Vector Machine is trained. We have built a training dataset composed by 2850 audio samples lasting a total number of 20 hours. On the test stage, this system is able to recognize the following events with an overall accuracy of 73%: someone falling down, steps, slice, screaming, rain, printer, people talking, frying food, filling water, door knocking, dog bark, car horn, glass breaking, baby crying, water boiling.

5. Conclusions

Preliminary results of our paper encourages us to keep on working on the human robot interaction and studying the information we obtain from the tests. This will conduct us to improve the definition of the protocols and the potential feedback that the robot can give to the elderly.

ACKNOWLEDGMENT

The authors would like to thank the Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement (Generalitat de Catalunya) under grant ref. 2014-SGR-0590.

References

[1] R. Suzman, J. Beard. Global health and aging–Living longer. National Institute on Aging, 2015.
[2] M. Vacher, F. Portet, et al. Challenges in the processing of audio channels for ambient assisted living. e-Health Networking Applications and Services (Healthcom), 2010 12th IEEE International Conference on. IEEE, pp. 330337, 2010.
[3] S. Chachada, J. Kuo. Environmental sound recognition: A survey. APSIPA Transactions on Signal and Information Processing, 2014.
[4] M. Chan, D. Estève, et al. A review of Smart Homes–Present state and future challenges. Computer Methods and Programs in Biomedicine, 91(1):55-81, 2008.
[5] M. Vacher, F. Portet, et al. Development of audio sensing technology for ambient assisted living: Applications and challenges. Int. Journal of E-Health and medical communications, 2 (1), pp.35-54, 2011.
[6] J. Chen, A. H. Kam, et al. Bathroom activity monitoring based on sound. Pervasive Computing. Springer, pp. 47-61, 2005.
[7] J.C. Wang, H.P. Lee, et al. Robust environmental sound recognition for home automation. IEEE Trans. Autom. Sci. Eng., vol. 5, no. 1, pp. 25-31, 2008.
[8] X. Valero, F. Alías. Classification of audio scenes using narrow-band autocorrelation features. Proc. 20th European Signal Processing Conf., (ISSN 2076-1465), pp. 2015-2019, Bucharest (Romania), August 2012.
[9] P. Guyot, X. Valero, et al. Two-step detection of water sound events for the diagnostic and monitoring of dementia, IEEE Int. Conf. on Multimedia and Expo, July 2013, San Jose, California (EEUU).
[10] D. Portugal et al. SocialRobot: An interactive mobile robot for elderly home care. IEEE/SICE International Symposium on System Integration (SII), Dec 2015.
[11] P. Mermelstein. Distance measures for speech recognition–psychological and instrumental. Pattern Recognition and Artificial Intelligence, pp. 374-388, 1976.
[12] F. Alías, J. C. Socoró, et al. A review of physical and perceptual feature extraction techniques for speech, music and environmental sounds. Applied Sciences, 6(5):143; May 2016.
[13] M.L. Zhang, and Z.H. Zhou. "A k-nearest neighbor based algorithm for multi-label classification." 2005 IEEE international conference on granular computing. Vol. 2. IEEE, 2005.