

Proceedings

# Real-Time Audio Event Detection over a Low-Cost GPU Platform for Surveillance in Remote Elderly Monitoring <sup>†</sup>

Rosa Ma Alsina-Pagès \*  and Marcos Hervás 

GTM-Grup de recerca en Tecnologies Mèdia, c/Quatre Camins, 30, 08022 Barcelona, Spain;  
mhervas@salleurl.edu

\* Correspondence: ralsina@salleurl.edu; Tel.: +34-93-290-2455

† Presented at the 4th International Electronic Conference on Sensors and Applications, 15–30 November 2017;

Available online: <https://sciforum.net/conference/ecsa-4>.

Published: 14 November 2017

**Abstract:** The average of life expectancy of the population and the prioritization of authorities in active and home aging has increased recently. This has led governments and private organizations to increase efforts in caring the elder and dependant segment of the population. The latest advances in technology and communications point out new ways to monitor those people with special needs at their own home, increasing their quality of life of the elderly or the dependant in a cost-affordable way. This same proposal can improve the quality of caring in retirement homes, giving support to the caring services. The purpose of this paper is to present an Ambient Assisted Living (AAL) able to identify, analyze and detect specific events in the daily life environment—mostly, at home or in a residence—defined by medical and assistant staff that can be considered as an emergency situation. It is designed to be deployed in controlled environments, where social services or medical staff are thought to be nearby. This hybrid network service is intended activate several alarms in the central services when certain situations occur in the monitored place. This tele-care proposal for certain predefined risk situations is validated through a proof of concept that takes benefit of the high performance computing capabilities of a NVIDIA Graphical Processing Unit on an embedded system named Jetson TK1 to be able to process and detect the events locally, even the situations that last in time. This platform holds the basic implementation of the acoustic event detection system, for both in-home or residence-based caring service. The system is nowadays designed to identify eight different situations along time, and set the correspondent alarm when one of the situations is detected.

**Keywords:** ambient assisted living; acoustic sensor network; graphics processor unit; home monitoring; residence assistance; surveillance

---

## 1. Introduction

Human life expectancy is increasing in the modern society, and it will keep growing during the next century [1]. There is a strong economic reason from governments to empower both elderly and partially dependent people to live independently, or at least, with the minimum caring services required. This would minimize costs and improve elderly independent life. Acoustic smart ambient assisted living technologies (AAL) present suitable solutions to provide a minimum-intrusive emergency detection in both home and residential environments. There is a huge market for companies capable of offering privacy-preserving AAL solutions to minimize the cost of care [2], most of them focused on monitoring the events in the home or residence.

In this paper, we present a proposal for the development of an acoustic event detection for surveillance purposes to support home and independent aging. This proposal includes the hardware solution, a low-cost GPU platform, to be used in both home and residences, and it operates with the acoustic event detection algorithms but also with the communications with a predefined central services by means of a hybrid network. This proposal is the continuation of the project named homeSound [3], focused on programming a low-cost GPU platform [4] for the audio event detection of fifteen in-home common sounds (e.g., water, walking, glass breaking, dog barking, etc.). The GPU platform is capable of computing the feature extraction and the machine learning methods to classify the environmental sounds real-time, and send the results to the cloud to be registered via Ethernet or activate any kind of alarm.

Section 2 reviews the state of the art of ambient assisted living and the automatic acoustic event detection projects. In Section 3, the definition of the requirements of the proposal is detailed. In Section 4, the system requirements are defined throughout the definition of the building typologies to propose a network architecture, a list of platforms involved and a signal processing solution. Finally, in Section 5, a discussion of the main constraints of the problem as well as the relevant aspects of the proposed solution is conducted.

## 2. Related Work

In recent years, there has been a rapid evolution of AAL related technologies due to the gradual aging of society, aiming to provide care and support to older adults in their daily life environment or to support the surveillance in residences. This section reviews the state-of-the-art of AAL applications, including some representative platforms used to process the data at home, giving special attention to those WASNs designed for this purpose.

### 2.1. AAL Research Projects

Several projects have been conducted within the AAL framework, which is wide and includes several technologies. We first detail a group of projects focused on allowing people to age at home, being some of them funded by the Assisted Living Joint Programme [5]. The project Aware Home [6] uses a wide variety of sensors, covering from specifically designed smart floors to more typical video and ultrasonic sensors, together with social robots to monitor and help older adults [7]. Another topic of interest of AAL projects during the last years has been behaviour or activity monitoring. In this sense, Project House [8] presents an alternative to track the house activity using sensors like cameras or microphones, which need a signal processing computation to derive behaviour conclusions. In the Gloucester Smart House project [9], a tele-care system was designed, based on lifestyle monitoring, with the pretension of continuously gathering information about the person's activity during daily routines.

### 2.2. WASNs for Tele-Care

A WASN is a group of wireless microphone nodes spatially distributed over an indoor or outdoor environment. Its design has to take into account at least the scalability of the network, the delay of the acoustic signal, the synchronization of the nodes and the decision of where the computing is performed [10]. One of the applications of sound source localization is the positioning of the person living alone [11] by means of a central system that aligns and analyzes the data coming from all the sensors. Another typical application of WASNs deployed in AAL environments is Acoustic Activity Detection (AAD) [10]. The primary purpose of AAD is to discriminate the overall acoustic events from the background noise [12], overcoming those approaches only based on energy threshold detector. Among the AAD, Voice Activity Detection plays a significant role for AAL solutions including acoustic interfaces [10].

Acoustic sensors at home can also be used -as in our proposal - for surveillance applications when taking care of the elderly or the disabled [13]. In [14], an acoustic fall detection system oriented to

the elderly age group living at home is described. The CIRDO project [15] was a multimodal effort to build a healthcare system to ensure the safety of seniors and people with decreasing independence at home. To that effect, CIRDO implements an audiovisual system that runs standard audio processing and video analysis tasks on a GPU. Despite the project's effort of privacy and private data, the patients are still not comfortable living with a system that processes real-time the video of the home activity.

### 3. Problem Description

The objective of this work is to present a solution of acoustic event detection at home or in a retirement home environments. In order to elaborate the proposal in the appropriate way, we give details of the acoustic nature of the events to be detected and the type of housing to which coverage should be provided.

#### 3.1. Acoustic Nature of Events

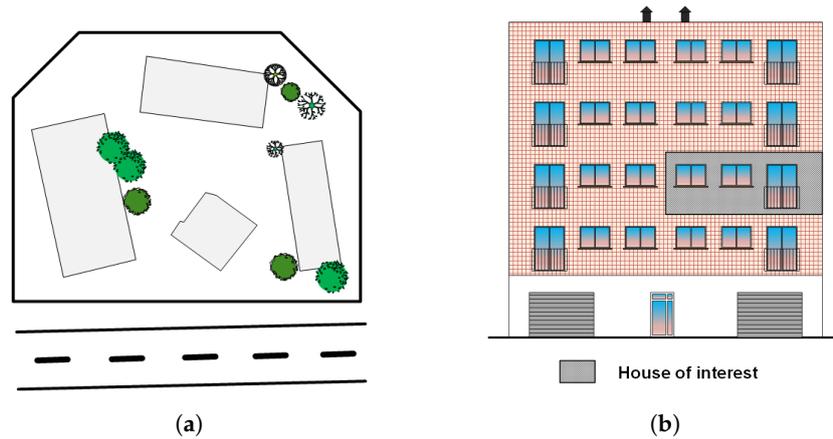
There are many acoustic events that may be likely to give an alarm in an apartment or room where an elderly or pseudo-dependent person lives. But for the implementation of this proposal four lines of action have been specified that contemplate both different types of sound (or voices) and its temporal analysis to determine if it is an event classified to give a warning.

1. *Door bell or phone ring*—a doorbell ringing for a long time, or a phone ringing constantly while nobody answers can be considered an important element to give an alarm. It means that there is nobody to answer at home, or that the person who is in the home is not in conditions of answering.
2. *Presence of more people at home*—the presence of many people at home or in a certain room is a possible risky behavior. They may have entered without consent, or it may be a situation of coercion for the elderly or pseudo-dependent inhabitant. It may also be that it is not a risk situation, but to prevent possible problems, the alarm will be raised.
3. *Patient shouting*—the patients' screaming are always a sign of alarm. They can come caused by not being well, by suffering some anxiety or panic attack, or by any other possible emergency situation (fire, theft, etc.).
4. *Activity at home after hours*—voices, television, music or any other sign of activity after hours are also cause for alarm. Being awake and active during the night can indicate disorientation or any other type of emergency at home home.

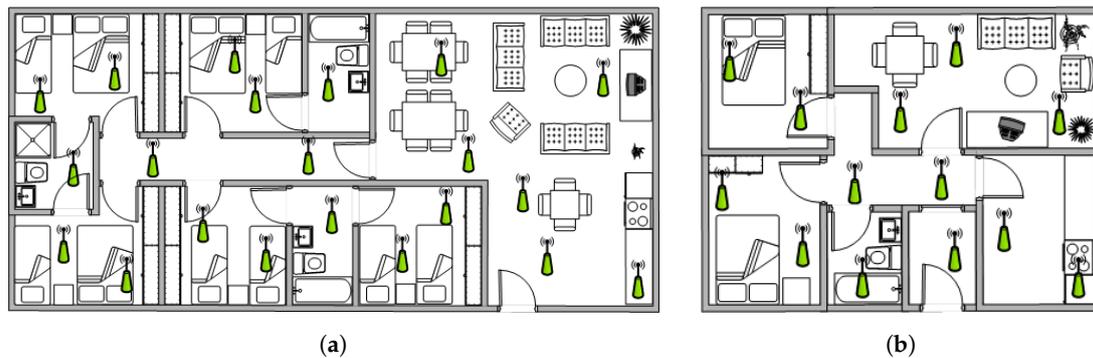
#### 3.2. Housing for the Elderly or the Pseudo-Dependent

There are two complementary proposals for this centralized service of acoustic alarms. The first and mainly supervised is the deployment in a retirement home (see Figures 1a and 2a). In this case, the sensory will be distributed in the common zones and overall, in the private areas. The goal of the service in this type of elderly housing is to minimize the supervision of the patients by caring personnel, especially in private areas, so that they can devote most of their attention to take care of them and not to surveillance tasks.

The second proposal is implemented in a private house or flat (see Figures 1b and 2b). The acoustic sensors and the processing and communication platform can be installed in any private home, in order to monitor its inhabitants from the centralized system, and set the alarm to the caring services. In this case, both the day and night zones will be taken into account for the deployment of the sensory, since the acoustic alarms predefined require surveillance in both areas.



**Figure 1.** Alternatives of flat or residence to be supported by an acoustic alarm service. (a) Residence site, with several buildings/houses with common spaces (dining room, etc.) and with private facilities for elderly or pseudo-dependent; (b) Isolated flats where elderly or pseudo-dependent people need variate surveillance support.



**Figure 2.** Detail of the alternative implementation of an acoustic alarm service, taking into account the sensing of all the available spaces in each type of living place. (a) Residence site, with several buildings/houses with common spaces (dining room, etc.) and with private facilities for elderly or pseudo-dependent; (b) Isolated flats where elderly or pseudo-dependent people need basic surveillance support.

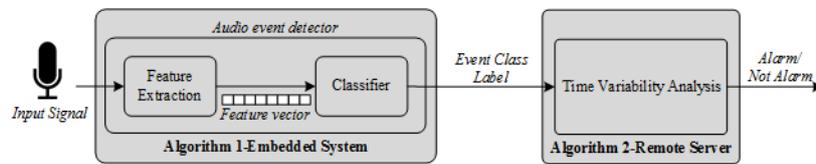
#### 4. System Proposal

The aforementioned topology of the buildings offered by this kind of service presents constraints concerning the sizing of the network, because the number of acoustic sensor deployed can be large enough to collect and process them in a couple of wireless data concentrator and a digital signal processing devices. It is worth noting that the raw acoustic data will be processed twice for different reasons: (a) in the frequency domain to classify the acoustic event, and (b) these events will be processed to decide whether there is an alarm; this second part of the study will face the time evolution of the labels.

##### 4.1. Signal Processing Solution

An audio event detection algorithm consists of two different stages, as detailed in Figure 3: the feature extraction and the classification. Feature extraction is a signal processing procedure to parametrize the key characteristics of the audio events by means of a set of representative coefficients with a lower dimensionality than the original samples [3]. When those coefficients are available, they are fed to the analysis and classification module, to obtain the identification of what type of acoustic

event has occurred in the predefined universe. In this application, also the time of the day is a key issue to classify whether the event detected has to set an alarm or it is a false alarm situation.



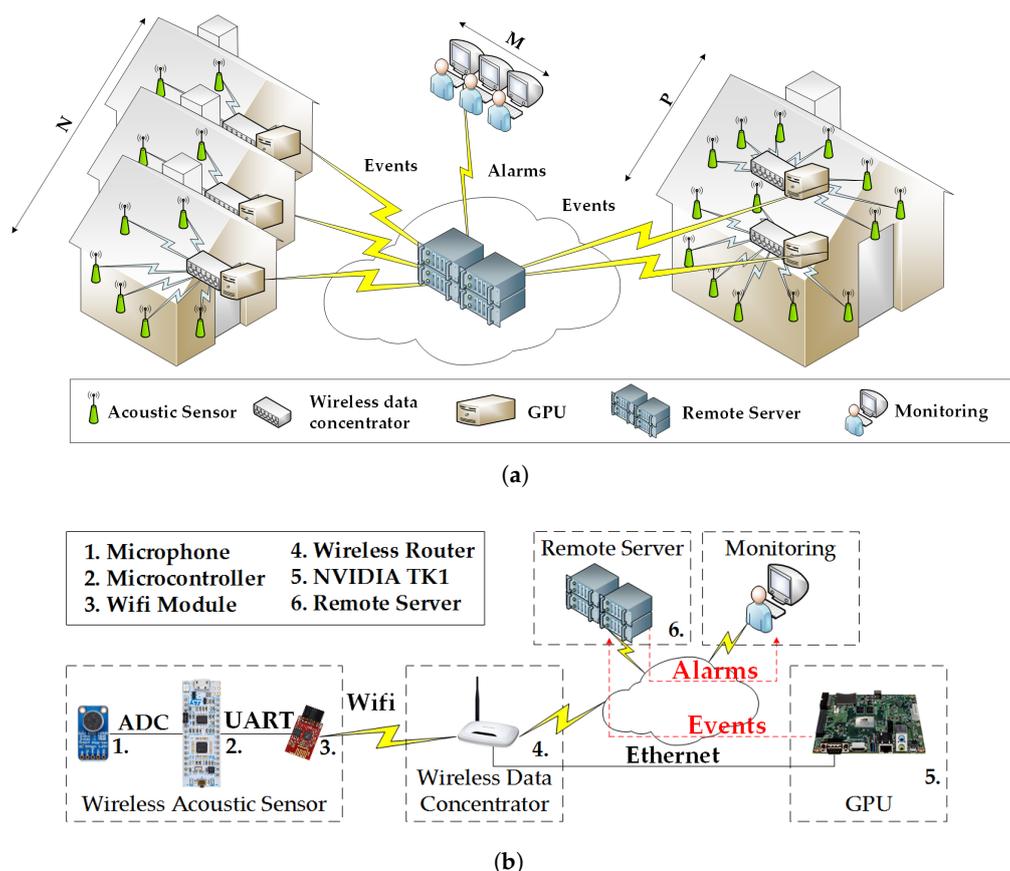
**Figure 3.** Block diagram of the algorithm proposed and where those algorithm will be implemented.

#### 4.2. Network Topology

To tackle the sizing the network, we propose dividing the large amount of wireless acoustic sensors into smaller segments of network, where a single wireless data concentrator combined can solve the signal processing and the communication. Actually, there are 5 well-defined subsystems in the proposed network topology (see Figure 4a). The subsystems features are listed hereafter.

1. The acoustic sensors in charge of sampling the raw audio at 44.1 ksps, and they send these data to the wireless concentrator.
2. The wireless data concentrator collecting all the information provided by the acoustic sensor and sending it to the GPU to be processed. It performs as a router.
3. The GPU will process the data collected from the acoustic sensor in frequency domain to classify th event. This device will be able to produce a label for every mote. Finally, the GPU send these labels to the remote server throughout the wireless concentrator.
4. The remote server carries out a time-evolution study of the predefined labels to inform to the caring service about the possible alarms.
5. Finally, the monitoring and caring service is in charge of visualizing and representing the alarms observed in the places of interest to inform the emergency services.

The hardware of the platforms involved in the AAL system presented is described in Figure 4b. The wireless acoustic sensors composed by three modules: (a) an electric microphone with the amplifier MAX9814 breakout board, (b) a Nucleo 32 development platform with the STM32L432KC ARM cortex-M  $\mu$ controller from ST and (c) a Wifi module based on the ESP8266. The output of the microphone breakout, with a cost of around 7 €, is connected to an ADC input of the Nucleo 32, with a cost of around 10 €, and the Nucleo 32 transmits the raw samples to the GPU throughout a UART port to the ESP8266, with a cost of 6 €. So, the acoustic sensor has a total cost of 23 €. The wireless concentrator can be a router; the aim of this device is to link the sensors with the GPU, and then the GPU with a remote server. The embedded system in charge of the signal processing of the data coming from the acoustic sensor of each network segment is the embedded GPU JETSON TK1 of NVIDIA [4] with a price around of 160 €. The remote server receives the information about the classification of the events sent by the GPUs and processes them to study the time evolution of the events, in order to determine whether an alarm should be set. Finally, these alarms will be sent to the monitoring and caring team.



**Figure 4.** Network topology description of the proposed AAL system proposed and the technological proposal of the devices which integrate that system. (a) Network topology of the system and (b) subsystems that integrate that full system and the elements used in each subsystem.

### 5. Discussion

The proposal that we present in this paper combines the network structure for a hybrid system, where the service is applied both in private homes and in retirement homes, with a similar type of alarm service in the caring system.

The proposal presents two major challenges. The first one is the low-level design of the network structure, from the single acoustic sensor installed, to the database that should decide whether an alarm is activated due to an emergency or maybe discard it, because it is not necessary. One of the key tests to be performed is the data throughput real-time implementation; we have to guarantee that the signal processing and event transmission is conducted real-time.

The second is the processing of the raw acoustic signal, the frequency study for the acoustic event detection. In a retirement home or in a private home there are several sensors installed; therefore, the final result of the acoustic event detection must take into account the network acoustic sensors processing. Although this consideration makes the algorithm more complex to design and to implement, and increases the computational cost, it is a more robust detection system than the AED using a single sensor.

**Acknowledgments:** The research presented in this work has been partially supported by the the Secretaria d'Universitats i Recerca del Departament d'Economia i Coneixement (Generalitat de Catalunya) under grant ref. 2014-SGR-0590.

**Author Contributions:** Rosa Ma Alsina-Pagès has written the introduction, the related work and the problem description. Marcos Hervás has written the technical proposal. Both authors have reviewed the paper and have worked together in the technical proposal detail.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

AAL	Ambient Assisted Living
AED	Acoustic Event Detection
AAD	Acoustic Activity Detection
WASN	Wireless Acoustic Sensor Network
UART	Universal Asynchronous Receiver-Transmitter

## References

1. Suzman, R.; Beard, J. *Global Health and Aging—Living Longer*; National Institute on Aging: Bethesda, MD, USA, 2015.
2. Vacher, M.; Portet, F.; Fleury, A.; Noury, N. Challenges in the processing of audio channels for ambient assisted living. In Proceedings of the 2010 12th IEEE International Conference on e-Health Networking Applications and Services (Healthcom), Lyon, France, 1–3 July 2010; pp. 330–337.
3. Alsina-Pagès, R.M.; Navarro, J.; Alías, F.; Hervás, M. HomeSound: Real-Time Audio Event Detection Based on High Performance Computing for Behaviour and Surveillance Remote Monitoring. *Sensors* **2017**, *17*, 854.
4. NVIDIA. JETSON TK1. Available online: <http://www.nvidia.com/object/jetson-tk1-embedded-dev-kit.html> (accessed on 15 May 2016).
5. Comission, E. Active and Assisted Living Programme. ICT for Ageing Well. Available online: <http://www.aal-europe.eu/> (accessed on 21 February 2017).
6. Abowd, G.D.; Mynatt, E.D. Designing for the human experience in smart environments. In *Smart Environments: Technologies, Protocols, and Applications*; John Wiley & Sons: Hoboken, NJ, USA, 2005; pp. 151–174.
7. Chen, T.L.; King, C.H.; Thomaz, A.L.; Kemp, C.C. Touched by a robot: An investigation of subjective responses to robot-initiated touch. In Proceedings of the 2011 6th ACM/IEEE International Conference on Human-Robot Interaction (HRI), Lausanne, Switzerland, 6–9 March 2011; pp. 457–464.
8. Tapia, E.M.; Intille, S.S.; Larson, K. Activity recognition in the home using simple and ubiquitous sensors. In *Pervasive Computing*; Springer: Berlin/Heidelberg, Germany, 2004; Volume 4, pp. 158–175.
9. Barnes, N.; Edwards, N.; Rose, D.; Garner, P. Lifestyle monitoring-technology for supported independence. *Comput. Control Eng. J.* **1998**, *9*, 169–174.
10. Cobos, M.; Perez-Solano, J.; Berger, L. Acoustic-based technologies for ambient assisted living. In *Introduction to Smart EHealth and ECare Technologies*; Taylor & Francis Group: Boca Raton, FL, USA, 2016; pp. 159–180.
11. Sertatlı, C.; Altinkaya, M.A.; Raoof, K. A novel acoustic indoor localization system employing CDMA. *Digit. Signal Process.* **2012**, *22*, 506–517.
12. Temko, A. Acoustic Event Detection and Classification. Ph.D Thesis, Universitat Politècnica de Catalunya, Barcelona, Spain, 2007.
13. Ellis, D. Detecting alarm sounds. In Proceedings of the Workshop on Consistent and Reliable Acoustic Cues CRAC-2000, Aalborg, Denmark, 2 September 2001.
14. Popescu, M.; Mahnot, A. Acoustic fall detection using one-class classifiers. In Proceedings of the 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Minneapolis, MN, USA, 3–6 September 2009; pp. 3505–3508.
15. Bouakaz, S.; Vacher, M.; Chaumon, M.E.B.; Aman, F.; Bekkadjja, S.; Portet, F.; Guillou, E.; Rossato, S.; Desserée, E.; Traineau, P.; et al. CIRDO: Smart companion for helping elderly to live at home for longer. *IRBM* **2014**, *35*, 100–108.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).