ASSISSTING DOCTORS IN THE DIAGNOSTIC AND FOLLOW-UP OF DEMENTIA USING THE SOUND INFORMATION FROM WEARABLE SENSORS

Xavier Valero[1], Francesc Alías[1], Patrice Guyot[2], Julien Pinquier [2]

[1] Grup de Recerca en Tecnologies Mèdia, La Salle – Universitat Ramon Llull

{xvalero; falias} @salleurl.edu

[2] Institute de Recherche en l'Informatique de Toulouse

{patrice.guyot; julien.pinquier @irit.fr

ABSTRACT

A significant aging of world population is foreseen for the next decades. Thus, developing technologies to empower the independency and assist the elderly are becoming of great interest. In this framework, we investigate **tele-monitoring technologies** to support doctors in the diagnostic and follow-up of **dementia** illnesses, such as Alzheimer. Specifically, **water sounds** are very useful to track and identify abnormal behaviors form everyday activities (e.g. hygiene, household, cooking, etc.). In this work, we propose a double-stage system to detect this type of sound events. Experimental results reveal the potential of the combined system, yielding an accuracy higher than 80%.

MOTIVATION

This work is framed on the IMMED project, which proposes a video monitoring system to **help the doctors in the diagnosis of dementia** illnesses. A wearable camera, positioned on the patient's shoulder, allows recording the ADL at his/her personal place of residence. The resulting videos are added to the traditional interviews and questionnaires to assist the doctors to measure autonomy decline.

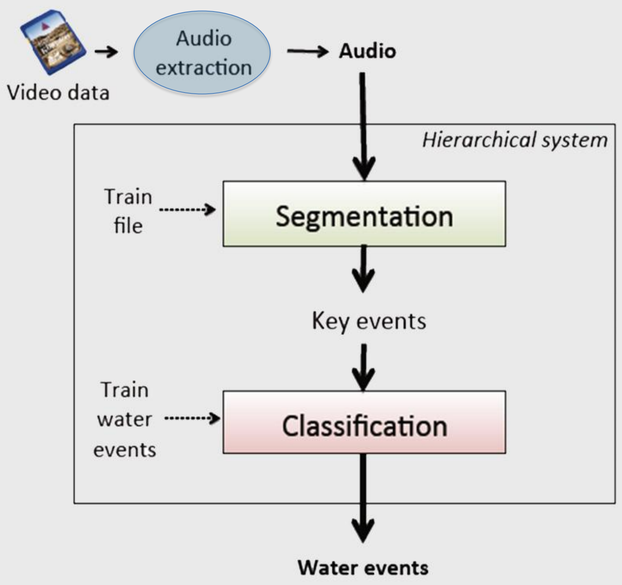
HYPOTHESIS

Water sound events are highly **representative** of several daily-life human activities, such as: hygiene activities (e.g., hand-washing, teeth brushing), diet activities (e.g., cooking, making coffee), housework (e.g., doing the dishes, moping, cleaning) and leisure (e.g., gardening).

Thus, tracking (automatically recognizing) this type of sound events may enable the indexing of daily-life activities relevant for the diagnosis of dementia.

METHOD

A classical pattern recognition approach may find difficulties to match the data due to the huge range of situations that can be found in a home environment. To overcome this problem, in this work we propose a **two-step method** for water sound event detection, using the **audio information** from the video recorded by the camera. The water event detection system is carried out in two steps:



1. SEGMENTATION:

   Without a specific training, it segments the sound stream extracting a set of key sound events which are potential candidates to belong to the water category. This part is based on **the Spectral Cover** (SC), which had been found to be more robust to speech and environmental background noises that classical low level descriptors in a water detection task.

2. CLASSIFCATION:

   Takes the isolated key events as input data and classifies them into two classes: water/no water. In this part, a detailed spectral analysis of the signal is performed by means of the **Gammatone Cepstral Coefficients**, a set of **bio-inspired** signal features that analyze the sound signals by mimicking the human auditory system. The classification of such patterns is performed by Support Vector Machines, a supervised machine learning technique.

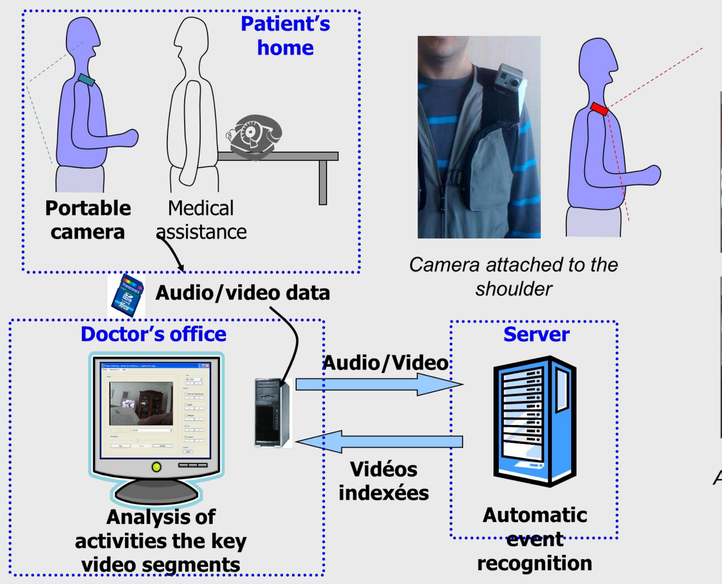
BENEFITS

For **doctors**:

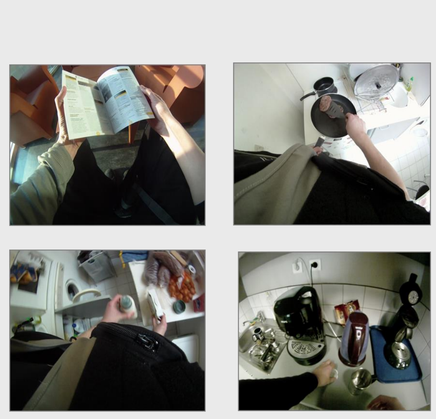
- The method **prevents** the **bias** inherent from carrying out the activities observation in a medical office.
- It ensures a more **ecological approach** to monitor the autonomy of the patient in his/her own living place.

For **patients**:

- It is far more **comfortable**, since the patient does not need to go to the hospital.



Camera attached to the shoulder



Activities recorded through the portable camera

EXPERIMENTS

A total of **20 patients** were recorded making activities of daily living in their own personal residence. Overall, **7 hours and 30 minutes** of human-daily activities were recorded. The recordings were carried out with a camera (HD Go-Pro Fisheye) placed on the shoulder of the patient. The audio part (with sampling rate of 12 KHz) was extracted from the video signal.

RESULTS

The experiments revealed that a large majority of water sounds were correctly recognized. In terms of **F-measure**, the proposed hierarchical system obtained a result of the **82%**, notably outperforming results in previous works (66% F-measure). Then, the information from the water sound events is used to improve the indexing of the key video segments.