Automatic audio-based recognition of activities of daily living

The general aim of this thesis is to validate a system that might automatically detect the activities of daily living of older persons living alone at home. The automatic detection is solely based on microphone measurements.

Promotors: Hugo Van hamme, Bart Vanrumste
Daily leaders: Peter Karsmakers (Peter.Karsmakers@esat.kuleuven.be), Tom Croonenborghs (Tom.Croonenborghs@cs.kuleuven.be)
Number of students: 1 or 2

The ongoing aging in our modern society leads to the tendency to allow older persons to stay as long as possible in their home environment. Although they are mainly able to independently organize themselves, it is for a certain group nevertheless necessary to observe their activities of daily living (ADL). Examples of such activities include sleeping, cooking, making a phone call, visiting the toilet, washing etc. Based on this automatic detection of ADL, we want to be able to detect changes in the behavior, i.e. changes in these ADL patterns. These changes include both acute and gradual changes. Acute changes are abnormal events that are critical and require an immediate alarm (e.g. a fall incident). On the other hand, we also want to detect gradual changes. These changes are important for an early detection of problems such as (early stage) dementia (e.g. ADL decline). The information about these activities and changes in behavior can then be presented to the caregivers (including family members) to adapt older people’s care plans, and as a consequence, increase their quality of care and quality of life. Hence allowing them to stay longer at their homes.

This thesis will investigate the recognition of ADLs using data captured via a network of audio nodes. An audio node is an audio capturing device possibly acquiring multiple audio channels. Recent technologies allow to build contactless, cheap, wireless, low power consuming, and compact audio nodes. Therefore a “large” number of audio nodes can be discretely spread across the living environment of the elder.

State-of-the-art methods learn in a fully supervised manner i.e. each individual sound (e.g. a microwave oven, water streaming from a tap, flushing a toilet, slamming a door, …) is first labeled then learned and afterwards combined to define an ADL. In this work an audio-based ADL recognizer will be learned using a non-unique labeling on the level of ADLs. For example given a large stream of audio with synchronously recorded video (only needed during the model training phase), segments of say 30min are labeled with corresponding ADLs, (note that a single segment may contain multiple ADLs). Such labeling requires less work than the case were each individual sound needs to be annotated. The labeled data set is then used to train the recognizer model. To accomplish this Probabilistic Latent Semantic Analysis (PLSA) such as used in e.g. [1] will be studied. Note that simulated data will be available for the thesis candidate.