Information Processing and Management 51 (2015) 204–214

Contents lists available at ScienceDirect

Information Processing and Management

journal homepage: www.elsevier.com/locate/infoproman

Human activity data discovery from triaxial accelerometer sensor: Non-supervised learning sensitivity to feature extraction parametrization

Inês P. Machado [a,c,*], A. Luísa Gomes [a], Hugo Gamboa [a,b], Vítor Paixão [c], Rui M. Costa [c]

[a] Faculty of Sciences and Technology, New University of Lisbon, 2829-516 Caparica, Portugal
[b] PLUX, Wireless Biosignals, Avenida 5 de Outubro, 70, 1050-059 Lisbon, Portugal
[c] Champalimaud Neuroscience Programme, Champalimaud Institute for the Unknown, Avenida de Brasília, 1400-038 Lisbon, Portugal

ARTICLE INFO

Article history:
Received 10 December 2013
Received in revised form 30 June 2014
Accepted 24 July 2014
Available online 15 November 2014

Keywords:
Human activity recognition
Interactive knowledge discovery
Feature extraction
Dimensionality reduction
Clustering algorithms

ABSTRACT

Background: Our methodology describes a human activity recognition framework based on feature extraction and feature selection techniques where a set of time, statistical and frequency domain features taken from 3-dimensional accelerometer sensors are extracted. This framework specifically focuses on activity recognition using on-body accelerometer sensors. We present a novel interactive knowledge discovery tool for accelerometry in human activity recognition and study the sensitivity to the feature extraction parametrization. *Results*: The implemented framework achieved encouraging results in human activity recognition. We have implemented a new set of features extracted from wearable sensors that are ambitious from a computational point of view and able to ensure high classification results comparable with the state of the art wearable systems (Mannini et al. 2013). A feature selection framework is developed in order to improve the clustering accuracy and reduce computational complexity.[1] Several clustering methods such as K-Means, Affinity Propagation, Mean Shift and Spectral Clustering were applied. The K-means methodology presented promising accuracy results for person-dependent and independent cases, with 99.29% and 88.57%, respectively. *Conclusions*: The presented study performs two different tests in intra and inter subject context and a set of 180 features is implemented which are easily selected to classify different activities. The implemented algorithm does not stipulate, *a priori*, any value for time window or its overlap percentage of the signal but performs a search to find the best parameters that define the specific data. A clustering metric based on the construction of the data confusion matrix is also proposed. The main contribution of this work is the design of a novel gesture recognition system based solely on data from a single 3-dimensional accelerometer.

1. Background

The past two decades have seen an increase in research activity in the field of human activity recognition. With activity recognition having considerably growing so did the number of challenges in developing, implementing and evaluating

* Corresponding author at: Faculty of Sciences and Technology, New University of Lisbon, 2829-516 Caparica, Portugal.
E-mail addresses: ip.machado@campus.fct.unl.p (I.P. Machado), alg.gomes@campus.fct.unl.pt (A. Luísa Gomes), hgamboa@plux.info (H. Gamboa), vitor.paixao@neuro.fachampalimaud.org (V. Paixão), ruicosta@fchampalimaud.orgt (R.M. Costa).

[1] The software OpenSignals (Gomes, Nunes, Sousa, & Gamboa, 2012) was used for signal acquisition and signal processing algorithms were developed in Python Programming Language (Rossum & de Boer, 1991) and Orange Software (Curk et al., 2005).

http://dx.doi.org/10.1016/j.ipm.2014.07.008
0306-4573/© 2015 Published by Elsevier Ltd.

solutions that may help the medical diagnosis of chronic motor diseases. A major challenge in our networked world is the increasing amount of data, which require efficient and user-friendly solutions with intuitive approaches of data visualization and extraction of relevant information.

The constant concern with the human physical and psychological well being has been the drive for research studies that have led to promising advances in medicine and engineering. In Biomedical Engineering, the demand for objectivity in clinical diagnosis of pathologies so subjective and hard to trace such as Obsessive Compulsive Disorder (National Institute of Mental Health, 2013) or Autism Spectrum Disorder (Simon, Errico, & Raffle, 2013), has been one of the greatest challenges. The physical activity disturbance is one of the essential signs of psychiatric disorders and many of these, including Depression (Song, Lee, Baek, & Miller, 2012) and Parkinson's Disease (Palmerini, Mellone, Avanzolini, Valzania, & Chiari, 2013), exhibit diagnostic criteria that require an assessment of the motor activity changes of the patient. The behavioural classification usually relies on observation and therefore a highly experienced analyst is always needed. Analysing human action is particularly challenging due to its complexity in the non rigid and self occluding nature of articulated human motion. Human body movement shows up to 244 degrees of freedom (C.H., 2006), which leads the modelling of structural and dynamic features for human activity recognition to a complex task.

The current research arises the hypothesis of accelerometry (ACC) as suitable technique for monitoring movement patterns in freely moving subjects during long periods of time. This technique can be used to measure quantitative parameters which may provide clinical insight into the subject health.

This study aims to understand the signals produced by a triaxial accelerometer through the human movement interpretation and clinical relevant parameters identification from the data. Signal processing techniques are implemented with the purpose of examining accelerometer data and finding new information that is difficult to identify from the raw data. We present a method for convenient monitoring of ambulatory movements in daily life, using a portable measurement device employing a single triaxial accelerometer.

This tool is based on an architecture of signal sensor processing, feature extraction, feature selection and clustering algorithms. To achieve the proposed goals, the following steps were implemented:

- Acquire and analyse the data produced by a triaxial accelerometer, placed on different parts of the body, during human movement, creating a database and proposing an annotation structure for motion data;
- Develop a framework for the interpretation of the data provided by an accelerometry monitoring system: design a set of time, statistical and frequency domain features from several research areas such as speech recognition and activity recognition;
- Develop algorithms to extract relevant information from the data;
- Apply clustering techniques based on feature representation of accelerometer data;
- Explore the choice of features and signals window size on the performance of different clustering algorithms;
- Develop a framework to differentiate human activities, such as sitting, walking, standing, running and lying;
- Evaluate the use of the system in daily life settings and the data interpretation framework: intra and inter subject context discovery;
- Implement a new metric based on classification performance evaluation: classification-based evaluation.

We present a method for convenient monitoring of detailed ambulatory movements in dailylife through a portable measurement device which employs a single triaxial accelerometer and machine learning techniques.

2. Related work

Most approaches to activity recognition, using body-worn accelerometers, involve a multi-stage process (Ghahramani, 2004). Firstly, the sensor signal is divided into a number of small time segments, referred to as windows, each of which is considered sequentially. For each window, one or more features are derived to characterize the signal. These features are then used as input to a clustering algorithm which associates each window with a cluster.

2.1. Composition of triaxial accelerometer signal

Before using the accelerometer system in any monitoring context, and before the algorithms development to interpret data recorded by the system, it is mandatory to understand the nature of the signals produced by the triaxial accelerometer unit. The signal measured by each fixed-body accelerometer is a linear sum of, approximately, three components (Mathie, 2003):

- Acceleration resulting from body movement – Body Acceleration Component;
- Acceleration resulting from gravity – Gravitational Acceleration Component;
- Noise intrinsic to the measurement system.

The first two components provide different information about the wearer of the device: the gravitational acceleration (GA) provides information about the space orientation of the device, and the body acceleration (BA) provides information