BBE 198 1–13
ARTICLE IN PRESS
BIOCYBERNETICS AND BIOMEDICAL ENGINEERING XXX (2017) XXX–XXX

Available online at www.sciencedirect.com
ScienceDirect
journal homepage: www.elsevier.com/locate/bbe

Review Article

Physical activity recognition by smartphones, a survey

Q1 Jafet Morales [a,*], David Akopian [b]

Q2 [a] University of Texas at San Antonio, BSE 1.500, One UTSA Circle, San Antonio, TX 78249, United States
[b] University of Texas at San Antonio, AET 2.354, One UTSA Circle, San Antonio, TX 78249, United States

ARTICLE INFO

Article history:
Received 5 December 2016
Received in revised form 8 April 2017
Accepted 20 April 2017
Available online xxx

Keywords:
Accelerometer
Gyroscope
Activity recognition
Smartphone

ABSTRACT

Human activity recognition (HAR) from wearable motion sensor data is a promising research field due to its applications in healthcare, athletics, lifestyle monitoring, and computer–human interaction. Smartphones are an obvious platform for the deployment of HAR algorithms. This paper provides an overview of the state-of-the-art when it comes to the following aspects: relevant signals, data capture and preprocessing, ways to deal with unknown on-body locations and orientations, selecting the right features, activity models and classifiers, metrics for quantifying activity execution, and ways to evaluate usability of a HAR system. The survey covers detection of repetitive activities, postures, falls, and inactivity.

1. Introduction

Researchers have attempted to recognize activities of varying levels of complexity via smartphones. This includes activities of low complexity, such as postures (i.e., remaining seated or standing), transitions (i.e., going from standing to sitting, or vice versa), or repetitive activities (i.e., walking, doing push-ups). The focus of this research is on classification of these activities, which have a low level of complexity and can be inferred from motion sensor data.

The most commonly studied activities in this type of research are walking, running, biking, jogging, remaining still, walking upstairs, and walking downstairs [1]. But the set of activities differs from one study to another. Activities with a higher level of complexity, such as *working*, may require the use additional signals such as GPS signals, WiFi signal strengths for indoor positioning, and audio [2]. But they also require more complex models due to intersubject and intrasubject variability, and the fact that in many cases these activities are composed of several low level activities performed one after the other.

Wearable motion sensor-based human activity recognition (HAR) has been explored with two types of sensor arrangements. Early research focused on recognizing activities with signals coming from one or more standalone motion sensors that were attached to the human body at locations chosen by the researcher [3–10]. Up to 5 bi-axial accelerometers attached at locations in ankle, wrist, hip, arm, and thigh were used in [3] but it was determined that there was not much recognition

* *Corresponding author at*: University of Texas at San Antonio, BSE 1.500, One UTSA Circle, San Antonio, TX 78249, United States.
E-mail addresses: jafet.aaron.morales@gmail.com (J. Morales), david.akopian@utsa.edu (D. Akopian).
http://dx.doi.org/10.1016/j.bbe.2017.04.004

Please cite this article in press as: Morales J, Akopian D. Physical activity recognition by smartphones, a survey. Biocybern Biomed Eng (2017), http://dx.doi.org/10.1016/j.bbe.2017.04.004

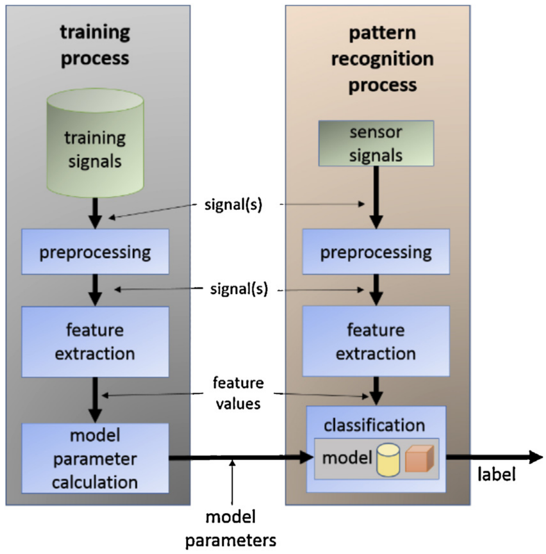

Fig. 1 – Steps for training and using a pattern recognition system. The classifier is generally trained or configured beforehand in order to recognize specific patterns.

improvement in accuracy to using only hip and wrist or thigh and wrist locations. Nowadays, many researchers are focusing on recognizing activities from the signals captured by a smartphone.

In general, the steps for pattern recognition include signal preprocessing, feature extraction, and classification. The classifier needs to be trained to recognize specific types of signals. Fig. 1 shows the general steps for pattern recognition.

This paper is organized as follows. Section 2 discusses which signals are relevant to the problem. Section 3 investigates the parameters for data acquisition that are popular with researchers (i.e., sampling rates, windowing), and the possibility of duty cycling. Section 4 discusses preprocessing algorithms (i.e., elimination of spikes and gravity). Section 5 discusses ways of dealing with signals when the mobile phones is mounted on a random on-body location. Section 6 examines ways of dealing with signals when the device has a random orientation with respect to the human body. Section 7 talks about the feature extraction stage, feature selection, and popular features used by researchers. Section 8 explores some of the machine learning aspects, including popular classifiers, activity models, and the use of contextual information (i.e., location). Section 9 discusses some of the metrics that may be useful in quantifying activity execution. Section 10 examines ways of evaluating the usability of these types of systems. Finally, Section 11 presents a conclusion.

2. Relevant signals

In most cases, accelerometer signals are used for activity estimation. The study in [11] used accelerometer, gyroscope, and magnetometer data. After applying a correlation feature selection (CFS) algorithm, 60% of the features left were accelerometer time-domain features, meaning that most of the relevant information is concentrated there. Inclusion of gyroscope and magnetometer features only increased accuracy by 5%. The study in [12] obtained an upper limit (training and testing on the same user) of 97.3% accuracy for a set of ten activities using only accelerometer signals.

The relevance of signals other than acceleration depends on the set of activities to be classified. For a simple activity set, that includes only the activities of *moving* and *not moving*, thresholding the standard deviation (STDV) of 3D acceleration magnitude can be enough to reach an accuracy of 99.4%, as shown in [2]. For more complex activity sets, context signals may be needed.

Context signals related to position include WiFi signal strengths for indoor positioning, GPS position, and barometer [2,13]. Fusing barometer data with accelerometer data can increase recognition performance for walking up or down the stairs by 20% [1,13]. Context signals related to the environment include humidity and audio signals. Audio signals have been used for HAR in [2,14]. Other context signals that are useful for on-body location estimation include proximity sensor signal, light sensor signal, and signals that provide orientation with respect to a global coordinate system [11].

Most recent studies for HAR by smartphones use Android phones [1]. Android APIs provide both, physical and virtual sensors. The signals for virtual sensors are calculated by the device already by processing and in some cases fusing signals from different physical sensors. The gravity component and the motion component of acceleration, for example, are already provided by the API so that the developer does not need to calculate them. Other signals provided by the API include orientation with respect to Earth in terms of a 3D rotation vector, significant motion detection signal, and step detection signal. Additionally, the API in [15] already allows for programming detection of activities in the form of callbacks. The target application receives a list of activities that were possibly being performed at a given time. Confidence values can be extracted for detected activities. This information can be used by the developer to program pop-up dialogs when a certain activity is detected or to calculate metrics that can be used by the application.

3. Popular parameters for data acquisition

The sampling rate is usually between 20 Hz and 50 Hz [16,17,4,18–21]. According to [22], 98% of the power for the walking activity is contained below 10 Hz, and 99% is contained below 15 Hz, which requires a minimal of 30 Hz to avoid undersampling. Also, no amplitudes higher than 5% of the fundamental exist after 10 Hz. It was also found in [23] that the main frequency components for running at the ankle on-body location are contained between 1 and 18 Hz [23]. Wearing a smartphone at the hip location would mean that the main frequencies are lower than 18 Hz [24]. Low sampling rates are attractive because they save battery life. Adaptive sampling techniques were studied in [25].

The size of the window to be captured and processed is usually 1–10 s [5,18,26,19,16,2,20,27,28,11]. It was found in [29] that a window of about 1 s is optimal for distinguishing between activity and rest, in terms of sensitivity and

Please cite this article in press as: Morales J, Akopian D. Physical activity recognition by smartphones, a survey. Biocybern Biomed Eng (2017), http://dx.doi.org/10.1016/j.bbe.2017.04.004