Future Generation Computer Systems 81 (2018) 307–313

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

A robust human activity recognition system using smartphone sensors and deep learning

Mohammed Mehedi Hassan [a,*], Md. Zia Uddin [b], Amr Mohamed [c], Ahmad Almogren [a]

[a] College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia
[b] Department of Informatics, University of Oslo, Norway
[c] Department of Electrical Engineering, University of British Columbia, Canada

HIGHLIGHTS

- A smartphone inertial sensors-based approach for human activity recognition.
- Uses deep learning based solution for successful activity recognition.
- The proposed approach was compared with traditional expression recognition approaches.

ARTICLE INFO

Article history:
Received 31 July 2017
Received in revised form 30 October 2017
Accepted 14 November 2017
Available online 22 November 2017

Keywords:
Activity recognition
Sensors
Smartphones
Deep belief network

ABSTRACT

In last few decades, human activity recognition grabbed considerable research attentions from a wide range of pattern recognition and human–computer interaction researchers due to its prominent applications such as smart home health care. For instance, activity recognition systems can be adopted in a smart home health care system to improve their rehabilitation processes of patients. There are various ways of using different sensors for human activity recognition in a smartly controlled environment. Among which, physical human activity recognition through wearable sensors provides valuable information about an individual's degree of functional ability and lifestyle. In this paper, we present a smartphone inertial sensors-based approach for human activity recognition. Efficient features are first extracted from raw data. The features include mean, median, autoregressive coefficients, etc. The features are further processed by a kernel principal component analysis (KPCA) and linear discriminant analysis (LDA) to make them more robust. Finally, the features are trained with a Deep Belief Network (DBN) for successful activity recognition. The proposed approach was compared with traditional expression recognition approaches such as typical multiclass Support Vector Machine (SVM) and Artificial Neural Network (ANN) where it outperformed them.

1. Introduction

Human Activity Recognition (HAR) has become an elegant research field for its remarkable contributions in ubiquitous computing [1–3]. Researchers use these systems as a medium to get information about peoples' behavior [4]. The information is commonly gathered from the signals of sensors such as ambient and wearable sensors. The data from the signals are then processed through machine learning algorithms recognize the events lying there. Hence, such HAR systems can be applied in many practical applications in smart environments such as smart home healthcare systems. For example, a smart HAR system can continuously observe patients for health diagnosis and medication [5] or it can be applied for automated surveillance of public places to predict crimes to be happening in near future [6].

In last few decades, many HAR systems were surveyed [7–9] where the authors focused on several activities in distinguished application domains [10,11]. For instance, the activities can be including, walking, running, cooking, exercising, etc. Regarding the duration and complexity of the activities; they can be categorized into three key groups: short activities, simple activities, and complex activities. The group of short activities consist of activities with very short duration such as transition from sit to stand. The second kind of activities is basic activities walking and reading [12]. The final one is basically combinations of progressions of basic activities with the interaction with other objects and individuals. Such kind of activities can be partying or official meeting together [13]. In this paper, we focus on recognizing basic activities.

* Corresponding author.
E-mail addresses: mmhassan@ksu.edu.sa (M.M. Hassan), mdzu@ifi.uio.no (Md.Z. Uddin), amrm@ece.ubc.ca (A. Mohamed), ahalmogren@ksu.edu.sa (A. Almogren).

https://doi.org/10.1016/j.future.2017.11.029
0167-739X/

HAR has been actively explored based on a distinguished kind of ambient and wearable sensors [1]. Some instances of such sensors include motion, proximity, microphone, video sensors. Most of the ambient sensor-based latest HAR researchers have mainly focused on video cameras as cameras make it easy to retrieve the images of surrounding environment. Video sensors are included with some other prominent sensors in some works in novel ubiquitous applications [14,15]. Though video sensors have been very popular for basic activity recognition. However, it faces very many difficulties when privacy issue arrives. On the contrary, wearable sensors such as inertial sensors can overcome this kind of privacy issues and hence; such sensors deserve more focus for activity recognition in smart homes [16].

Many HAR systems over the past used accelerometers to recognize a big range of daily activities such as standing, walking, sitting, running, and lying [17–23]. In [20], the authors have explored the accelerometer data to find out the repeating activities such as grinding, filling, drilling, and sanding. In [21–23], the authors tried to do elderly peoples' fall detection and prevention in smart environments. Majority of the aforementioned systems adopted many accelerometers fixed in different places of the human body [17–21]. However, this approach apparently not applicable to daily life to observe long-term activities due to attachment of many sensors in the human body and cable connections. Some studies tried to explore the data of single accelerometers at sternum or waist [22,23]. These works reported substantial recognition results of basic daily activities such as running, walking, lying, etc. However, they could not show good accuracy for some complex activity situations such as transitional activities (e.g., sit to stand, lie to stand, and stand to sit).

Thus, regarding different sensors in activity recognition, the accelerometer is the most commonly utilized sensor for focusing on human body motion [8]. The sensor can be deployed in two ways. First, one is in multi-sensor package such as triaxial accelerometers or Body Sensor Networks (BSN). The second one is in combination with other sensors such as gyroscopes, temperature, and heart rate sensors [24]. Bao and Intille [12] proposed one of the earliest HAR systems for the recognition of 20 activities of daily living using five wearable biaxial accelerometers and well-known machine learning classifiers. They achieved reasonably good classification accuracy reaching up to 84% considering the number of activities involved. One evident drawback was related with the number and location of the body sensors used, which made the system highly obtrusive. Gyroscopes have also been employed for HAR and have demonstrated to improve the recognition performance when used in combination with accelerometers [25,26].

In the case of wearable sensors in activity recognition, the smartphone is an alternative to them due to the support of the diversity of sensors in it. Handling sensors such as accelerometers and gyroscopes along with the device processing with wireless communication capabilities made smartphones a very useful tool for activity monitoring in smart homes [27,28]. Besides, smartphones are very ubiquitous and require almost no static infrastructure to operate it. This advantage makes it more practically applicable than other ambient multi-modal sensors in smart homes. As recent smart phones consist of inertial sensors (e.g., gyroscopes and accelerometers), they can be appropriate sensing resources to obtain human motion information for HAR [29,30].

Recently, smartphones have attracted many activity recognition researchers as they have fast processing capability, and they are easily deployable [31–34]. For instance, in [31], the authors used wirelessly connected smartphones to collect a user's data from a chest unit composed of the accelerometer and vital sign sensors. The data was later processed and analyzed using different machine learning algorithms. In [32], the authors developed a HAR system to recognize five transportation activities where data from smartphone inertial sensors were used with a mixture-of-expert model for classification. In [33], the authors proposed an offline HAR system where a smartphone with built-in triaxial accelerometer sensor was used. The phone was kept in the pocket during experiments. In [34], the authors used a smartphone mounted in the waist to collect inertial sensors' data for activity recognition. They used Support Vector Machine (SVM) for activity modeling. In [35], a smartphone was used to recognize six different activities in real-time. In [36], the authors proposed a real-time motion recognition system with the help of a smartphone with accelerometer sensors. Similarly, the authors in [37] used a smartphone with an embedded accelerometer to recognize four different activities in real-time.

As the dimension of the features from different sensors becomes very high in activity recognition, Principal Component Analysis (PCA) can be applied in this regard [37]. PCA applies a linear approach to find out the directions with maximum variations. Thus, PCA is adopted in this work to reduce the dimensions of high-dimensional features. Recently, deep learning techniques have been getting a lot of attentions by pattern recognition and artificial intelligence researchers [38–40]. Though deep learning is more efficient than typical neural networks, it consists of two major disadvantages: it has overfitting problem, and it is often much time-consuming. Deep Belief Network (DBN) is one of the robust deep learning tools that use Restricted Boltzmann Machines (RBMs) during training [39]. Hence, DBN is a good candidate to the model activity recognition system.

In this work, a smartphone-based novel approach is proposed for HAR using efficient features and DBN. The rest of the paper is organized as follows. Section 2 explains the feature extraction process from depth images. Then, Section 3 illustrates the modeling of different expression through deep learning. Furthermore, Section 4 shows the experimental results using different approaches. Finally, Section 6 concludes the work with some remarks.

2. Proposed method

Fig. 1 shows the basic flowchart of the proposed system. The proposed system basically consists of three main parts: Sensing, Feature extraction, and recognition. The part is sensing. It collects sensor's data as input to the HAR system. For this study, two prominent sensors in smartphones have been selected for data collection: triaxial accelerometers and gyroscopes. The sensors provide measurements at frequencies within 0 Hz and 15 Hz. The second major part is the feature extraction. This part starts with removing noise to isolate relevant signals such as gravity from triaxial acceleration. After removing noise, it does statistical analysis on fixed-size sliding windows over the time-sequential inertial sensor signals to generate robust features. The third key part of the system is modeling activities from the features via deep learning where DBN is adopted.

2.1. Signal processing

Triaxial angular velocity and linear acceleration signals are considered from the smartphone gyroscope and accelerometer sensors. The sampling rate of the raw signals is 50 Hz for both the sensors. These signals are then preprocessed to reduce noise. Two filters are used in this regard: median and low-pass Butterworth filter. Twenty Hz is considered as cutoff frequency for the Butter-worth filter. Another low-pass Butter-worth filter is applied to the acceleration signal with gravitational and body motion components to filter out body acceleration and gravity information. It gravitational forces are assumed to have low-frequency components and 0.3 Hz is considered as optimal corner frequency