2012

Comparing Features Extraction Techniques Using J48 for Activity Recognition on Mobile Phones

Blázquez Gil, Gonzalo

Springer

Omatu, S. et al. (eds.), 2012. Distributed Computing and Artificial Intelligence: 9th International Conference (Advances in Intelligent and Soft Computing, 151), Springer, pp. 141-150.
http://hdl.handle.net/10016/18242

This document is published in:

Omatu, S. et al. (eds.), 2012. *Distributed Computing and Artificial Intelligence*: 9th International Conference (Advances in Intelligent and Soft Computing, 151), Springer, pp. 141-150.
Doi: 10.1007/978-3-642-28765-7_18

Comparing Features Extraction Techniques using J48 for Activity recognition on Mobile Phones

Gonzalo Blázquez Gil, Antonio Berlanga de Jesús, and José M. Molina Lopéz

Applied Artificial Intelligence Group, Universidad Carlos III de Madrid,
Avd. de la Universidad Carlos III, 22, 28270, Colmenarejo, Madrid, Spain
{gonzalo.blazquez,antonio.berlanga,josemanuel.molina}@uc3m.es
http://www.giaa.inf.uc3m.es

Abstract. Nowadays, mobile phones are not only used for mere communication such as calling or sending text messages. Mobile phones are becoming the main computer device in people's lives. Besides, thanks to the embedded sensors (Accelerometer, digital compass, gyroscope, GPS, and so on) is possible to improve the user experience. Activity recognition aims to recognize actions and goals of individual from a series of observations of themselves, in this case is used an accelerometer.
Keywords: Mobile device, Activity Recognition, Ambient Assisted Living, J48, features extraction.

1 Introduction

Ambient intelligence (AmI) could be viewed as human interface metaphor. In AmI people are empowered through a context aware environment that is sensitive, adaptive and responsive to their needs, habits, gestures and even emotions [?]. AmI sees a world where a huge mesh of sensors are integrated into daily objects, clothing, people and so on. Using this information, the environment itself can provide context-aware services to support its inhabitant.
Taking into account that people is the main actor in AmI enviroments, it is mandatory to provide a tool for user to comunicate with the enviroment. At this point, mobile phones present several advantages: they are considered essential in people life's, so they could be considered as a non-intrusive sensor; they experience the same physical forces, temperature and noise that the person who carries them out [?]. On the contrary, obtaining physical actions from smartphones presents several problems [?] [?], basically, they are not built to collect information and infer activities.
Considering these advantages, it may be possible to consider a smartphone like a non-intrusive device to obtain people activities. The ability to understand human life patterns by analyzing user mobile phone behaviour is becoming a new challenge for researchers [?].
The inference of user activities implies a large number of sensors distributed over the body and/or the environment, depending on the activities to detect [?]. Smartphones are especially well-suited to this task

because they have integrated Microelectromechanical systems (MEMS) which make easier to obtain user information. They may obtain and process physical phenomena from embedded sensors (MEMS) and send this information to remote locations without any human intervention [?]. Smartphones should take advantage of mobile contextual information, such as position, user profile or device features; to offer amazing services.
GPS, Wi-Fi, Bluetooth and microphone are the most known sensors in mobile phones, however, recently, new kind sensors have been added: accelerometer, gyroscope, compass (magnetometer), proximity sensor, light sensor, etc. [?]. As a result of this, not only phone numbers and addresses are collected in the mobile phone but also location, temperature, noise, physical forces may be collected to offer user new kind of amazing applications.
Finally, a performance studio about three different ways to realize activity recognition using smartphones (Spectrogram, Continuous Wavelet Transform and mean, standard deviation and other features) is presented in this paper. Besides, a dataset is created using an HTC Magic mobile phone with Android Operating System. The quality of the given solution is measured using a J48 tree.
The paper deals with the topic of recognizing user's activities by analyzing the data produced by motion sensors embedded in mobile phones. Sensory data is collected by a mobile application made in Android and it sends to a server where pre-learnt activities are recognized in real-time. Besides this study rely on the power of the GPS in order to tag every action that the mobile phone takes using speed value.

2 Related work

There are many different methods to retrieve user activity information from raw sensor data in the literature. However, the principal steps can be categorized as preprocessing, segmentation, feature extraction, dimensionality reduction and classification [?].
Normally, raw sensor data is collected using ad-hoc accelerometers over the body. However, placing sensors in multiple body locations could be annoyed for the user. New researches try to make this more comfortable for users using a smartphone. In this section, some architectures are briefly described where a smartphone is used for that purpose.

2.1 Cenceme architecture

Probably, Cenceme is the most known system to recognize physical activities from mobile devices (Figure 1a). This architecture aims to infer user physical activities from mobile devices and to share them on social network. The proposed architecture is split in three layer: Sense, learn and share.
Sense layer aims to collect raw sensor data from sensors embedded in the phone. In learn layer, they propose to use a variety of data mining techniques to infer user rules. These techniques are used to interpret mobile

data extracted from sensor layer. Their approach is to share information in a web portal where sensor data and inferences are easily displayed.

2.2 lifeMap architecture

Yohan Chon et all. [?] present LifeMap, an Smartphone-based Context Provider for Location-based Services. The presented architecture is split in four component: (i) All the sensor are placed on the low level, this level send the information (ii) to the Component Manager where information is processed and provide high-level information. Using high-level information from the Component Manager, (iii) the Context Generator generates a point of interest (POI) which contains the user context. The context map is stored in a database to match and aggregate user contexts. And finally, (iv) The Database Adapter is an interface to provide user context to other applications.

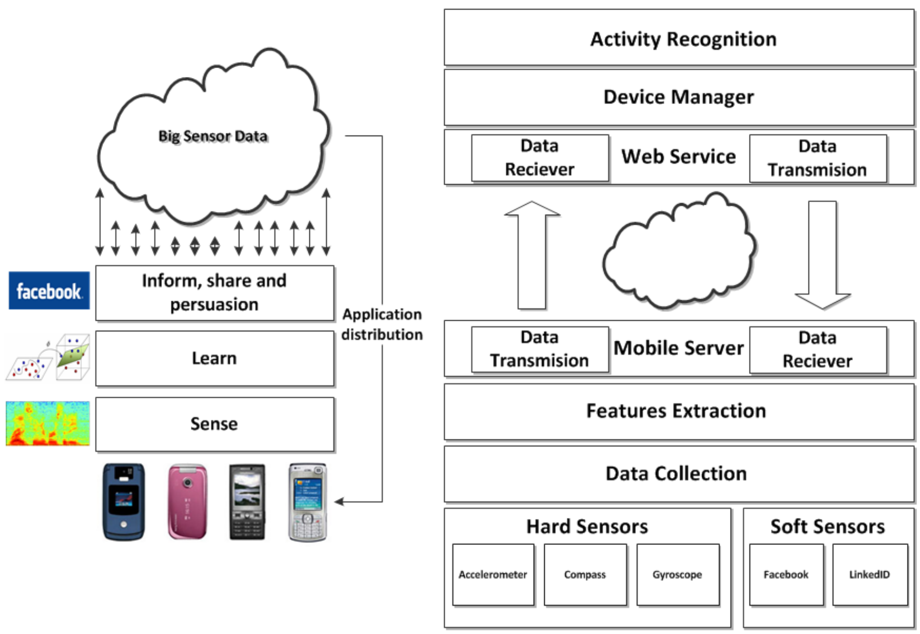

Fig. 1. Proposed Architecture by Cenceme and inContexto.

2.3 inContexto architecture

Finally, in [?] was presented inContexto (Figure 1b). It is a distributed architecture to retrieve user context information from mobile phones.

The architecture is split in five components. Some of them run on the cloud and some others run on the mobile phone. This paper rely on this architecture to obtain sensory data from mobile phones.

- Data Acquisition: A low-level sensing module continuously gathers relevant information about the user activities using sensors. The proposed architecture this component is located on the mobile device.
- Features Extraction: The features extraction level is also implemented in the mobile phone. The module processes the raw sensor data into features that help discriminate between activities. This level aims to process and select which features are better to identified an action.
- Activity recognition: The last layer is classification module that uses the features selected in the mobile phone to infer what activity an individual or group of individuals is engaged in, for example: Walking, running, sitting, standing. In this component, it will be implemented the algorithms (Supervised learning, Probabilistic classification, Model-based or instance-based learning) to figure out the taken action.

3 Experimentation

Pattern recognition answer to the description and classification of measurements taken from physical or mental processes [**?**]. In order to provide an effective and efficient description of patterns, pre-processing is often required improve performance, removing noise and redundancy in measurements. Then a set of characteristic measurements, which could be numerical or not, and relations between them, are extracted representing the patterns.
Collecting data is a hard task, hence in order to generate enough trajectories examples to make the training process, the data collection was made in a different way. This process has four steps: Data Collection, Trajectories generation, Features extraction and Training process.
Tagging process has been made using mobile phone GPS which distinguish every single action whith the speed value. Thus, it is not neccesary user involvement in tagging process, every time the GPS is enable, the application start to log data.
It was used three individuals who made five different activities (Running, Walking and Standing up). An HTC magic mobile phone equipped with Android OS was chosen to perform the measurement. Finally, a dataset was created for the research community and it is available online in [1].

3.1 Data Collection

In this study, the accelerations and azimuths of the pedestrian were collected using a HTC Magic. The created dataset has the following attributes: 3-axis accelerometer values in the mobile device Cartesian reference system, 3-axis Compass values, 3-axis accelerometer values in the

[1] GIAA Web page http://www.giaa.inf.uc3m.es/

real world reference system, GPS precision and GPS speed. Next table show the number of instances for each activity.
The sampling frequency can be adjusted according to the action studied. In this case, rely on the next study[**?**], the sampling frequency range requiring to obtain human actions is $0.6Hz$ to $2.5Hz$. Consequently, to prevent aliasing problem, the Nyquist-Shannon sampling theorem is followed:

$$F_E \geq 2 * F_{Max}$$

Finally, the sampling frequency was fixed to the maximum that the Android OS permits, in this case 50 Hz which is more than sufficient compared to 5Hz recognize the activities of the pedestrian.

	Running	Standing	Walking
Instances	5,118	7,321	24,825
Seconds	102.36	146.42	496.5

Table 1. Number of second and samples for each activity.

3.2 Trajectories generation

It is necessary a big amount of trajectories to make correctly the training process. However, it is quite costly to generate enough trajectories to make this process.
In this case, the selected trajectories are made semi-automatically. First of all, we have 3 files corresponding each activity (Running, Walking, and Standing up). Subsequently, a Java program has been created to mix all the activities generated a unique trajectory. Finally all the generated trajectories have been stored to continue the pattern recognition process. However, there are some requirement to make this trajectories as real as possible:

1. All the trajectories start with a Standing up action.
2. The next action could be the action besides (Figure 2) or the same action again.
3. The minimum duration of each action is 2 seconds and the maximum is 7 seconds.
4. Finally, each trajectory consists in 10 actions.

When the trajectories generation process is over, it is necessary to discretize the speed value due to J48 tree users nominal values. Thus, all the samples are discretized in 5 classes:

– Stop class: It is when the GPS speed measurements are less than 1 km/h.

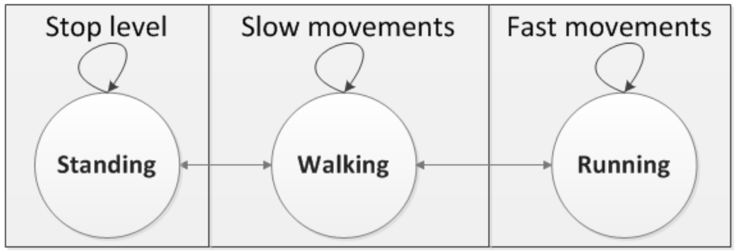

Fig. 2. Generation trajectories model.

- Walking class: Speed value from the GPS is more than 1 km/h and less than 4.
- Walking fast class: In this case, GPS speed values are among 4-6 km/h.
- Running class: It is when the GPS speed measurements are more than 6 km/h and less than 10 km/h.
- Running fast class: Finally the last class takes the GPS speed values upper than 10 km/h.

Finally, 1000 trajectories was created to infer activities. Every trajectory is different, in duration and actions, to each other. Weka[2] was used as the machine learning tool in this paper and it is necessary to transform data into arff format.

3.3 Features Extraction

Some research present different ways to obtain features in order to infer physical action. A comparison study using wavelet and frequency features (DWT, CWT, and STFT) is presented in [**?**]. These techniques provide several advantages, one particular advantage of frequency modulation is its resilience to signal level variations. On the contrary, in [**?**] is presented another features to infer activities using accelerometer values.
In this study, features were extracted from the raw accelerometer signals via a sliding window of 512 samples (Approximately 5 seconds), 256 of which overlap with consecutive ones. An sliding windows with 50% overlap has been defined in previous works [**?**]. This work uses GPS in order to obtain user's speed who is taking place the action, thus, the classifier output value is the mean of the speed in the sliding window.
This study is focused on compare three kind of features:

- The first one is based on Spectrogram function (STFT, Short-Time Fourier Transform). A spectrogram is a time-varying spectral representation that shows how the spectral density of a signal varies with time.
- The second one is Continuous Wavelet Transformation, is used to split a continuous-time signal into wavelets. Unlike Fourier transform, the CWT is able to construct a time-frequency representation

[2] Weka web page http://www.cs.waikato.ac.nz/ml/weka/

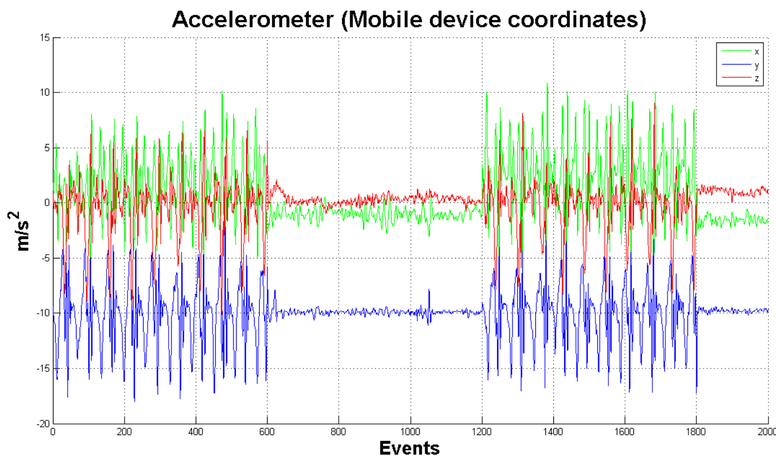

Fig. 3. Trajectory example with 4 actions (Walking, Standing up, Walking and Standing up. First graph shows the variance of the three coordinates mobile devices. The second one depicts the value of the real vertical world coordinate. The last one shows an example of spectrogram feature.)

of a signal which offers very good time and frequency localization. Both of these techniques (STFT and CWT) present several vales (higher than 150), however, they are not necessary all of them. For that reason, only the first 25 frequencies were selected such us possible feature. Besides, the signals need to be transformed from mobile device coordinate to real world coordinates following the next formula:

$$a_{realworld} = a_{mobiledevice} * R * I$$

Where I matrix is a simple rotation around the X axis and the rotation matrix R which is the identity matrix when the device is aligned with the world's coordinate system.

- The last way to make the training process is a set of different values from the raw accelerometer data, specifically, tree axes mean values, 3 axes standard deviation values, correlation between each axis and signal energy for each axis.

3.4 Results

The selected machine learning algorithm is a J48 classifier which is the Weka version from the C4.5 decision tree algorithm. J48 was chosen because gives results in tree model which can be easily transformed into real time applications.

The selected parameters for the J48 decision tree are:

- Confidence Factor = 0.25

- Minimum number of object = 2
- unpruned = false
- Test-options = 10 folds Cross-validation

After processing the training and testing sets with the J48 classifier in Weka, the results are highly accurate in vector and spectrogram features, however results are poorly accurate if CWT features extraction is used.

	Features	Leaves	Tree size	Time(s)	Accuracy	Mean absolute
error						
CWT	25	8741	17481	129.32	62.85 %	0.1631
Spectrogram	25	1007	2013	41.44	95.63 %	0.0198
Vector	12	648	1295	14.57	97.20 %	0.0131

Table 2. Features of J4 tree generated by Weka.

Table 2 shows result from each selected technique to extract features. The best implemented technique is features vector set, which is not only more accurate than the other ones, otherwise it provides the smallest tree generated and the minimum generation tree time. The size of the tree is very important because it will be implemented in a real application in a mobile phone. A bigger size of the tree causes more energy consumption according to the increase of CPU cycles. Another way to study the quality of the feature extraction techniques is using the confusion matrix (Figure 4).

CWT

```
    a     b     c     d     e   <-- classified as
 9462  3357   269    67   132 |   a = Stopping
 3115 16957  2804   997  1852 |   b = Walking
  302  3601  2025   658  1652 |   c = WalkingFast
  114  1277   775  1100  2856 |   d = Running
  169  2144  1439  1860 20266 |   e = RunningFast
```

Spectrogram

```
    a     b     c     d     e   <-- classified as
12815   472     0     0     0 |   a = Stopping
  473 24592   640    20     0 |   b = Walking
    0   667  7217   352     2 |   c = WalkingFast
    0    23   395  5501   203 |   d = Running
    0     0     4   212 25662 |   e = RunningFast
```

Vector

```
    a     b     c     d     e   <-- classified as
12979   308     0     0     0 |   a = Stopping
  286 25047   392     0     0 |   b = Walking
    0   360  7639   239     0 |   c = WalkingFast
    0     1   247  5676   198 |   d = Running
    0     0     0   187 25691 |   e = RunningFast
```

Fig. 4. Confusion matrix from CWT, Spectrogram and Vector features.

CWT technique is the worst of all the studied technique, besides it does not present any advantage over the other ones. Secondly, spectrogram achieves great results, besides, this technique uses only one signal (vertical movement in the real world) in order to obtain the spectrogram. Although confusion matrix shows that is possible to classify an instance in a class not next to the real class. Thus, the best performance (high accurate and less tree size) is presented by Vector technique. Besides, confusion matrix figure shows that Vector features extraction just fail with the class near the one which is classified (e.g. Running instead of Running fast).

4 Conclusions

In this paper, a study comparing three different techniques in order to infer activity recognition using a J48 decision tree was presented. Besides, the study rely on inContexto architecture to collect accelerometer data Overall, the presented work further demonstrates that using a mobile phone providing with accelerometers is enough to infer actions that user is taking place.

Selected features is an important field inside the Activity recognition systems. This paper aims to identify and record in real-time selected features related on user activity using a mobile device.

The best given solution obtained an overall accuracy of 97.20 % well classify instances of 79250 different actions. This solution is a vector composed by: Energy, mean, standard deviation and correlation of each axes.

The flexibility of the Android OS along with the phones hardware capability allows this system to be extended. For example, create an application which is able to send a sms or call to your relatives if you are doing strange movements. This application may be interesting in ancient people. Another application may be a indoor GPS. Based on the user movement fingerprint, this application could follow a person who is moving indoors.

5 Acknowledgments

This work was supported in part by Projects CICYT TIN2011-28620-C02-01, CICYT TEC2011-28626-C02-02, CAM CONTEXTS (S2009/TIC-1485) and DPS2008-07029-C02-02