Predicting Physical Activities from Accelerometer Readings in Spherical Coordinate System

Kittikawin Lehsan and Jakramate Bootkrajang*

Department of Computer Science, Chiang Mai University
Chiang Mai, Thailand
{kittikawin.l,jakramate.b}@cmu.ac.th

Abstract. Recent advances in mobile computing devices enable smartphone an ability to sense and collect various possibly useful data from a wide range of its sensors. Combining these data with current data mining and machine learning techniques yields interesting applications which were not conceivable in the past. One of the most interesting applications is user activities recognition accomplished by analysing information from an accelerometer. In this work, we present a novel framework for classifying physical activities namely, walking, jogging, push-up, squatting and sit-up using readings from mobile phone's accelerometer. In contrast to the existing methods, our approach first converts the readings which are originally in Cartesian coordinate system into representations in spherical coordinate system prior to a classification step. Experimental results demonstrate that the activities involving rotational movements can be better differentiated by the spherical coordinate system.

Keywords: Activity recognition, Classification, Spherical coordinate system

1 Introduction

Advances in semi-conductor and sensors technology foster the development of practical wearable devices. Many of such devices, for example a heart rate monitor or a GPS unit, can be easily spot in our daily life activities. People use them either for recreational, health or security purposes. The interesting thing is that raw data recorded by those devices can often be useful in understand the nature of the activities. With the help of data analysis techniques currently available, we can now gain more insights into regularities and patterns in the data.

The mobile phone industry also benefits from the technological advancement. We now see manufacturers packed a number of sensors into its mobile phones. Unlike wearable devices which is quite specific to its task, mobile phone is more ubiquitous. This leads to the idea of alternatively using mobile phone to sense the world instead of using specialised wearable devices.

There are increasing number of applications which make use of sensory data gathered from mobile phone sensors. For example, [10] used mobile phone data

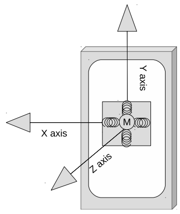

Fig. 1. An illustration of an accelerometer of a typical mobile phone. The sensor measures accelerations in x, y, z axes.

to detect person's mode of transportation. Readings from mobile phone's accelerometer can be used to signal the falling of elderly people [4]. Similarly, [6] proposed a methodology for efficiently recognising human activities as well as a method for detecting the falling. A work in [5] studied the using of mobile phone's GPS unit together with predefined points of interest for inferring user's activities based on current position of the user. Along the same line, a combination of an accelerometer and a gyroscope was used to recognise daily activities such as walking, standing and sitting [12]. Also, [3] proposed a method to detect walking as well as counting steps using smartphone's sensory data. Apart from sensory data from mobile phone, [8] studied activity classification using pressure sensors attached to five different spots on the body.

In this work, we are interested in inferring physical activities using data readings from mobile phone's accelerometer. Briefly speaking, an accelerometer is an electro-mechanical sensor used for measuring acceleration. The sensor is composed of a tiny mass attached to springs. A change in acceleration causes the springs to compress or extend, and the degree of compression/extension is translated to acceleration accordingly. In a typical mobile phone, accelerations in the x, y, z axes are usually provided. A picture of an accelerometer and its axes is illustrated in Figure 1.

A notable work that pioneered activities recognition using data from an accelerometer is probably the one in [9]. In that work, a special accelerometer unit is paired with an IPAQ personal digital assistant for data processing purpose. The work considered the recognition of eight different daily activities. The authors found that it is sufficient to use the mean, the standard deviation and the correlation of readings in the x-axis, the y-axis and the z-axis of the accelerometer to correctly recognise most of the activities. From that moment, researchers then started to investigate the topic from different perspectives and in different environments. The work in [12] compared various classification algorithms for activities recognition. They found that k-NN and boosting algorithms are among the top performer for the task. Studies that focus on the effect of smartphone's mounting positions on recognition rate can be found in [2] and [1].

The work which are most related to our work are probably the work in [7] and [11]. Both studied the recognition of physical activities using only the accelerometer. The empirical results using different classification algorithms suggested that accurate recognition can be achieved only with the data from the accelerometer.

However, and rather interestingly, all of the above work only consider the data represented in the Cartesian coordinate system. Motivated by the fact that a spherical coordinate system is often used to represent data which contains rotational movements, we postulate that the spherical coordinate system might be more suitable in capturing the dynamics of activities which involve rotational movements compared to the typical Cartesian coordinate system. Accordingly, we consider the following points to be our contributions.

- We investigate the using of the spherical coordinate system to represent the data instead of the typical Cartesian coordinate system.
- We empirically study an appropriate time frame [1] for extracting data from a stream of sensor readings.
- We extensively test the proposed representation using five different well-known classifiers.

The rest of the paper is organised as follow. Section 2 presents our approach that uses the spherical coordinates to represent the data readings. Section 3 then presents empirical evaluations while Section 4 concludes the study.

2 The Proposed framework

The movement of smartphone during activity recognition can be of both translational and rotational motions. For example, the movements for sit-up and squating are rather similar except that the movement for sit-up additionally contains rotational motions. Differentiating the two activities can be challenging. Motivated by a capability to capture rotational motions, in this work we will investigate the using of a spherical coordinates to represent our data. The spherical coordinate system is a generalisation of a polar coordinate system to three dimensional vector space. Specifically, a point (x, y, z) in the Cartesian coordinate system can be converted to a point (r, θ, ϕ) in the spherical coordinate system by

$$r = \sqrt{x^2 + y^2 + z^2} \tag{1}$$

$$\theta = \arccos \frac{z}{r} \tag{2}$$

$$\phi = \arctan \frac{y}{x} \tag{3}$$

Essentially, the quantity r in (r, θ, ϕ) tuple represents the length of a vector measured from the origin, θ represents an angle between the vector and the z-axis while ϕ represents an angle that the projection of the vector on the xy plane makes with the x-axis.

One interesting property of the spherical coordinate system is that the distance between any two points is no longer a straight line. Instead the distance is the length of the curvature which is known as a geodesic. This non-linearity

[1] Later on, we will use the term ‘window’ and ‘time frame’ interchangeably.

Table 1. A list of 9 features extracted from raw data collected from the smartphone.

Feature	Calculation	Description
μ_r	$\frac{\sum_{i=1}^{n} r_i}{n}$	The mean of r
μ_θ	$\frac{\sum_{i=1}^{n} \theta_i}{n}$	The mean of θ
μ_ϕ	$\frac{\sum_{i=1}^{n} \phi_i}{n}$	The mean of ϕ
σ_r	$\sqrt{\frac{\sum_{i=1}^{n}(r_i-\mu_r)^2}{n-1}}$	The standard deviation of r
σ_θ	$\sqrt{\frac{\sum_{i=1}^{n}(\theta_i-\mu_\theta)^2}{n-1}}$	The standard deviation of θ
σ_ϕ	$\sqrt{\frac{\sum_{i=1}^{n}(\phi_i-\mu_\phi)^2}{n-1}}$	The standard deviation of ϕ
$corr(r,\theta)$	$\frac{cov(r,\theta)}{\sigma_r\sigma_\theta}$	The correlation between r and θ
$corr(r,\phi)$	$\frac{cov(r,\phi)}{\sigma_r\sigma_\phi}$	The correlation between r and ϕ
$corr(\theta,\phi)$	$\frac{cov(\theta,\phi)}{\sigma_\theta\sigma_\phi}$	The correlation between θ and ϕ

of the distance measure is beneficial for the case of non-linearly separable data. Therefore, distance-based classifiers such as a nearest-neighbour classifier might be able to take advantage of this transformation. Further, as mentioned earlier, we believe that geodesic distance is also more natural for this task since we are interested in movements which is composed of rotational elements.

Next, we will follow the steps previously used in [9] for summarising the stream of sensor readings. The steps involve dividing the stream of data into equal window of length n seconds. Commonly in the literature including that found in [9], a 10-second window is used. In this work, we will additionally investigate windows of different lengths namely 5, 10, 15 and 20 seconds. A set of sensor readings that fit in one time frame will undergo feature extraction steps in order to produce a point representation for the readings in that time frame. To extract the features, we calculate the means, the standard deviations of each of the three variables, i.e., r, θ and ϕ, as well as the correlations between all pairs of the variables. This, in total, results in a 9-dimensional point representation for the raw sensor readings in one time frame. Although, other statistical features, i.e., max, min, inter quartile range, energy, can be extracted from the set of sensor readings, we empirically observed that these 9 features can adequately capture the dynamics of the activities under consideration. The finding is also in agreement with [9]. Table 1 summarises the 9 features extracted from one time frame of length n.

In summary, our framework involves first converting the readings into the spherical coordinates. A set of 9 features are then extracted from the converted data producing a set of 9-dimensional input vectors. The set of input vectors together with manually assigned class labels are used to train a classifier.

3 Experiments

3.1 Data collection

We manually collected data by asking volunteers to perform each of the physical activities namely, walking, jogging, push-up, squating and sit-up for 50 repeti-

(a) Walking and Jogging (b) Squatting (c) Sit-up (d) Push-up

Fig. 2. Smartphone mounting positions for each of the activities.

tions. In each repetition we asked the volunteers to exercise for 20 seconds. We mounted the mobile phone on the right shoulder of the participants as shown in Figure 2.

During data collection, we sampled data from the accelerometer sensor which are accelerations in the x-coordinate, the y-coordinate and the z-coordinate using 1Hz sampling rate. We converted the raw readings into the spherical coordinates so that the 9 features can later be extracted. In total, there are 5000 seconds of sensory data available (1000 seconds for each of the five activities). Now, labelled datasets of different size can be constructed from the sensory data depending on the length of the time frame. For example, using 10-second time frame would result in a dataset containing 100 input instances for each class, while using 20-second time frame would produce 50 input instances for each class.

3.2 Protocol

In this study, we are interested in whether transforming accelerometer readings into the spherical coordinate representation is beneficial. To evaluate this, we will compare the performances obtained from five well-known classifiers namely, k-nearest neighbour, linear discriminant analysis, naive Bayes, Support Vector Machine with linear kernel and Support Vector Machine with polynomial degree 2 operating in the Cartesian coordinate system and the newly proposed spherical coordinate system. We note that, due to the small sample size nature of the data, we will use leave-one-out cross validation technique to measure the predictive performance of the classifiers. In addition, we are also interested in the effect of the length of the time frame on the classification accuracy. For this purpose, we will additionally study the comparative performance of the classifiers learning from data sets produced by setting the time frame to 5 seconds, 10 seconds, 15 seconds and 20 seconds, respectively.

3.3 Results

Firstly, we would like to establish a window length that gives the best recognition performance. Table 2 summarises recognition accuracies for 4 different windows

Table 2. Mean accuracies (%) of the two coordinate systems as a function of window length averaged over all classifiers. Boldface entries indicates the best window length.

Window length (seconds)	Cartesian	Spherical
5	84.35 ± 12.49	74.63 ± 12.33
10	87.72 ± 10.76	85.03 ± 11.74
15	88.69 ± 9.34	88.40 ± 9.98
20	**89.59 ± 8.75**	**90.33 ± 9.08**

Table 3. Predictive accuracy of 5 classifiers. Window length is fixed to 20 seconds. Boldface highlights the coordinate system in which each of the classifier works best.

Classifier	Cartesian	Spherical
5-NN	97.20	**100.00**
Decision tree	**98.00**	95.60
SVM (Linear Kernel)	96.80	**100.00**
SVM (Polynomial Kernel degree 2)	98.80	**99.60**
Naive Bayes	97.20	**99.20**
Linear Discriminant Analysis	97.60	**100.00**

length tested. We can see, from the table, that larger time frame is generally better compared to the smaller ones in both coordinate systems. Comparing the two coordinate systems, we observe that for smaller time frame the Cartesian coordinate system has an upper hand but as window length increases to 20 seconds we see that using the spherical coordinate system is preferable. The accuracy averaged over all the classifiers is 90.33% for the spherical system while the accuracy for the Cartesian coordinate system slightly lags behind. We did not test window lengths beyond 20 seconds but we speculate that the result might not be worse than the 20 seconds mark. Moreover, as we shall subsequently see that using 20 seconds window is already very satisfying as we achieved perfect classification performance. Therefore, increasing the length of the time frame might be overkill.

Next, we take what we have learnt from the previous experiment that the 20 seconds time frame yielded the best averaged performance. In this experiment, we fix the window length to 20 seconds and compare the performances of the two coordinate systems using five classifiers. The classification accuracies of the five classifiers trained using data represented in the Cartesian and the spherical coordinate systems are summarised in Table 3. From the table, it is quite clear that the spherical coordinates representation gives relatively better predictive performance compared to the Cartesian coordinates counterpart. Interestingly, we also see that three out of five classifiers namely, 5-NN, SVM with linear kernel and Linear Discriminant Analysis, all learnt from the data represented in the spherical coordinate system achieved perfect classification results.

To gain further insights regarding why using spherical coordinates is superior to using the Cartesian coordinates, Table 4 presents the confusion matrix of the 5-NN classifier operated in the Cartesian coordinate system where it achieved 97.2% classification accuracy. We noticed that the 5-NN misclassified sit-up as

Table 4. Confusion matrix for the 5-NN under the Cartesian coordinate system.

	Walking	Jogging	Squatting	Push-up	Sit-up
Walking (predicted)	50	0	0	0	0
Jogging (predicted)	0	50	0	0	0
Squatting (predicted)	0	0	47	0	3
Push-up (predicted)	2	0	0	48	0
Sit-up (predicted)	0	0	2	0	48

squatting and vice-versa. It is worth noting that the orientation of the smartphone in these two activities are almost identical (see Figure 2). Further, the movement of the device during the action are also similar, with an exception that there exists rotational movements around the hip when performing sit-up. It seems that the Cartesian representation cannot capture this rotational movements adequately. However, this does not seem to be a problem for the spherical representation. The claim is confirmed by the perfect predictive performance of the 5-NN operates in the spherical coordinate system. The explanation also applies to the case of walking and push-up. For the sake of exposition, let us consider 2-dimensional trajectories of a smartphone in Figure 3. The blue trajectory contains rotational elements while the red one does not. If we were to extract, for example, the mean of x and y components from the readings (represented by the blue and red dots) from the two trajectories under the Cartesian coordinate system, we would end up with the similar means. As such, differentiating the two trajectories using the means would be problematic. However, converting the reading into the polar coordinates (a special case of the spherical coordinates in 2D) can clearly alleviate the problem since the means of r and θ from the two trajectories are quite distinct. We believe that this is a reason why working in the spherical coordinate system is desirable.

Overall, based on the empirical evidences we can conclude that transforming sensor readings from the Cartesian coordinates to the spherical coordinates is advantageous for physical activities recognition especially when the activities involve rotational motions.

4 Conclusion and Future work

In this work, we studied comparative performance of representing accelerometer readings in the Cartesian and the spherical coordinate system for human activities recognition. As part of the experiment we also investigated a suitable window length for summarising a stream of sensory data sampled from the accelerometer. Experimental results based on recognition rates of five well-known classifiers suggest that a 20 seconds window and the spherical coordinate system is well-suited for the task, and especially so for capturing the dynamics of physical activities that involve rotational motions. Future work includes the relaxation on the mounting position of the smartphone. For that purpose, a rotation-invariant data representation is required, so that the accelerometer readings are relatively the same regardless of the position of the mobile device.

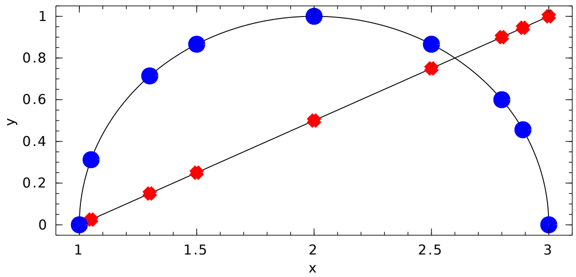

Fig. 3. Two trajectories that produce similar features under the Cartesian coordinate system but not under the polar coordinate systems as well as the proposed spherical coordinate system.

References

1. Antos, S.A., Albert, M.V., Kording, K.P.: Hand, belt, pocket or bag: Practical activity tracking with mobile phones. Journal of neuroscience methods 231, 22–30 (2014)
2. Ayu, M.A., Ismail, S.A., Matin, A.F.A., Mantoro, T.: A comparison study of classifier algorithms for mobile-phone's accelerometer based activity recognition. Procedia Engineering 41, 224–229 (2012)
3. Brajdic, A., Harle, R.: Walk detection and step counting on unconstrained smartphones. In: Proceedings of the 2013 ACM international joint conference on Pervasive and ubiquitous computing. pp. 225–234. ACM (2013)
4. Dai, J., Bai, X., Yang, Z., Shen, Z., Xuan, D.: Perfalld: A pervasive fall detection system using mobile phones. In: Pervasive Computing and Communications Workshops (PERCOM Workshops), 2010 8th IEEE International Conference on. pp. 292–297. IEEE (2010)
5. Furletti, B., Cintia, P., Renso, C., Spinsanti, L.: Inferring human activities from gps tracks. In: Proceedings of the 2nd ACM SIGKDD International Workshop on Urban Computing. p. 5. ACM (2013)
6. Kozina, S., Gjoreski, H., Gams, M., Luštrek, M.: Efficient activity recognition and fall detection using accelerometers. In: International Competition on Evaluating AAL Systems through Competitive Benchmarking. pp. 13–23. Springer (2013)
7. Kwapisz, J.R., Weiss, G.M., Moore, S.A.: Activity recognition using cell phone accelerometers. ACM SigKDD Explorations Newsletter 12(2), 74–82 (2011)
8. Moncada-Torres, A., Leuenberger, K., Gonzenbach, R., Luft, A., Gassert, R.: Activity classification based on inertial and barometric pressure sensors at different anatomical locations. Physiological measurement 35(7), 1245 (2014)
9. Ravi, N., Dandekar, N., Mysore, P., Littman, M.L.: Activity recognition from accelerometer data. In: Aaai. vol. 5, pp. 1541–1546 (2005)
10. Reddy, S., Mun, M., Burke, J., Estrin, D., Hansen, M., Srivastava, M.: Using mobile phones to determine transportation modes. ACM Transactions on Sensor Networks (TOSN) 6(2), 13 (2010)
11. Weiss, G.M., Lockhart, J.W.: The impact of personalization on smartphone-based activity recognition. In: AAAI Workshop on Activity Context Representation: Techniques and Languages. pp. 98–104 (2012)
12. Wu, W., Dasgupta, S., Ramirez, E.E., Peterson, C., Norman, G.J.: Classification accuracies of physical activities using smartphone motion sensors. Journal of medical Internet research 14(5), e130 (2012)