An unconstrained Activity Recognition Method using Smart Phones

Naciye Çelenli, Kamile Nur Seviş,
Muhammed F. Esgin, Kemal Altundağ, Umut Uludağ

BİLGEM TÜBİTAK
(The Scientific and Technological Research Council of Turkey)
Gebze, Kocaeli, Turkey

naciye.celenli@tubitak.gov.tr
nur.sevis@tubitak.gov.tr
muhammed.esgin@tubitak.gov.tr
kemal.altundag@tubitak.gov.tr
umut.uludag@tubitak.gov.tr

Abstract: In this paper, we provide human activity recognition performance rates, using accelerometer and gyroscope signals acquired using smart phones. Covering seven basic actions (walking, running, jumping, standing, ascending stairs, descending stairs, and standing up and sitting down as one action) and a complex action (getting in and out of a car), with more than 100 subjects in a database collected in different environments, we provide recognition results on the largest database in the published literature. Utilizing features (e.g. extrema, zero crossing rates...) extracted from time-windows (e.g. with a duration of 2 seconds), K-Star classifier led to the best performance among 6 classifiers tested, exceeding 98% recognition accuracy. A detailed comparison with current approaches is provided, along with possible future research directions. The associated technology could be helpful for health-related monitoring of one's activities, generating automatic status feeds for social networking sites, and calculating precise/adaptive calorie intake needs for individuals.

1 INTRODUCTION

Smart phones are becoming ubiquitous: as their abilities increase in sensing, computing and networking fronts, people are using them as an indispensable part of their lives, not just for making and receiving phone calls. One can run business, entertainment, shopping, health, communication etc applications, which add extra value to basic services provided by mobile phones of the previous decade. Motivated by this, we are trying to see if smart phones can be used for human activity recognition, and to what extent. There have been prior studies in this area, albeit with many involving very small databases. In this paper, utilizing more than 100 subjects' data, we run classifiers on features extracted from accelerometer and gyroscope signals, originating from popular smart phones. Of the 6 classifiers we experimented with, K-star led to recognition rates exceeding 98%.

The organization of the paper is as follows: Section 2 describes previous studies in this area. In Section 3, details about the subject set and activities along with data collection and feature extraction steps are explained. Section 4 describes which experiments were performed and their results. Section 5 includes concluding remarks and ideas for further research.

Table 1: Comparison with previous work

	# of actions	**# of subjects**	**frequency**	**overall best success**	**Orientation**
[BGC09]	6	Not given	30Hz	80%	Not specified
[Yan09]	6	4	36Hz	90.6%	Not specified
[SZL$^+$10]	7	7	10Hz	93.1%	Only 4 allowed
[KWM11]	6	29	20Hz	91.7%	Not specified
[HTM11]	6	5	50Hz	86.38%	Only 2 allowed
[DDK$^+$12]	8	10	80Hz	92%	Not specified
Proposed Method	*7*	***102***	*100Hz*	***98.16%***	*Not specified*

2 RELATED WORK

Activity recognition using sensors such as accelerometers and gyroscopes is not a new topic. However, until around 2009, most of the works in this area considered the case where accelerometers for non-commercial use are strapped on a human body, in which case there is very little change in the orientation of the sensor. The results of these works seemed to be efficient due to the absence of much noise. We refer to [BI04, RDML05, LCB06] for more details about this subject. On the other hand, in recent years instead of using wearable sensors, some studies [BGC09, DDK$^+$12, GFH09, HTM11, KWM11, SZL$^+$10, Yan09] used commercial devices (most commonly smart phones) to detect activities with the possibility of displacements of the sensors while performing a certain action. Contrary to the former case, the latter is most likely to include a lot of noise in the raw data. Further, considering the soaring number of smart phones in use today (and projected further increases), using them as a platform for activity recognition is logical. For comparison between previous studies and our work, please see Table 1. Columns in the table correspond to the following: reference to the work, the number of actions for activity recognition, the number of subjects data was collected from, sampling rate, overall best success among different classifiers' results and whether or not only specific orientations of the phone are allowed during data collection, respectively.

There are 3 main contributions of our work. First, we collected data from a large set of subjects with varying ages and gender, and the activities were carried out at different environments (see section 3.1). Secondly, we achieve, to the best of our knowledge, the highest overall success rate compared to similar studies. Lastly, we tested if the same methodology can be used for gender recognition (see section 4).

Figure 1: The phone and the application used for data collection, and one of the environments ascending/descending stairs data was collected.

3 ACTIVITY RECOGNITION SYSTEM

3.1 Activities and Subject Set

We collected data from 102 subjects. 35 of them are females and 67 of them are males. The ages vary between 19 and 55, and the average age is 30.

The activities consisted of 7 main actions: walking, running, jumping, standing, ascending stairs, descending stairs, and standing up and sitting down as one action. Each action is named as *walking, running, jumping, standing, stair-up, stair-down and updown*, respectively. 40 of the subjects were able to perform 2 sessions of data collection (of 7 main actions), at least a week apart from each other. The rest was able to perform the actions once.

Moreover, as a separate experiment, 30 subjects performed the *in-out-car* action which means walking approximately 8 steps towards a car, getting in to the car, pretending to start the engine, getting out and walking to the starting position. The reason for data collection of the in-out-car activity is to see weather complex actions could be seen as a combination of some simple actions. For example, with an appropriate segmentation of these simple actions, one can recognize the complex action that they form (see Figure 2 at the end of the paper for a sample graph of an in-out-car action, which was manually segmented and each segment was manually classified).

3.2 Data Collection

In this study, we collected data from accelerometer and gyroscope sensors via an application we developed for iOS (see Figure 1). iPhone 5 and iPhone 5s smart phones were used to collect data (see [App] for details about extracting sensors' data on iOS). Initially, several different frequency levels were tested and 100Hz was chosen to be the experimental sampling rate.

While collecting data, we tried to realize real-world scenarios as much as possible. So, we did not specify any clothing, shoes or the orientation of the phone. A session consisted of performing 7 activities (or a subset of them), each for a period of around 30 seconds, while the mobile phone remained in the subject's left or right pocket according to his/her choice of pocket. However, not all subjects performed all the actions and some subjects performed some actions for a shorter period of time. In order to eliminate noise occurrences at the start of the activity, the application delays retrieving data for 5 seconds. At least 5 different environments and different places in these environments were used. An example of these environments can be seen in Figure 1. The overall collected data was approximately 15000 seconds ($\approx$ 4.17 hours). The ratio of each action's duration to the duration of overall collected data is 0.17 for walking, 0.17 for running, 0.15 for jumping, 0.14 for standing, 0.12 for ascending stairs, 0.11 for descending stairs, and 0.14 for standing up and sitting down.

3.3 Feature Extraction and Classification

It is hard to characterize activities from the raw data obtained from sensors. To extract essential information, the raw data from both sensors was first divided into windows. We tested window sizes of 1, 2, 4 and 6 seconds. Also, the sliding window idea, which is shown to be efficient in earlier works [BI04], was adapted. Each window overlapped with the previous one with half the size of a window. e.g if the window size is 2 seconds, then two consecutive windows overlap by 1 second.

After the signal is divided into windows, a feature vector was extracted from each window (we developed a C++ code for this purpose). The original vector consisted of the values of minimum, maximum, mean, standard deviation, root mean square and zero-crossings of (x, y, z) values from the accelerometer data, and the values of minimum, maximum, mean and standard deviation of (x, y, z) values from the gyroscope data. From now on, this feature vector of size 30 will be referred as *the base feature vector*.

It was observed that no information about the shape of the data curve was included in the base feature vector. That's why the binned-average of raw data (see Table 2 for detailed explanations of all features) was added to the feature vector and it was tested if it would lead to an increase in the performances. It did not have much effect on most of the classifiers but increased the performance of *Multilayer Perceptron* classifier by about 1%.

Another idea was to use gender as a feature (i.e. adding a new binary component to the feature vector). It was thought that a female's actions would have lower levels of variation in acceleration values. This could result in a case like a female's ascending stairs could be classified as a male's walking. The tests showed that this idea led to only slight changes in most of the classifiers, but increased the performance of *Multilayer Perceptron* classifier by 1%. Detailed analyses of classifiers are omitted and we refer the reader to [DHS00] for a comprehensive study on pattern recognition and classification.

Table 2: Feature Vector

	Accelerometer	Gyroscope	Explanation
Min	(x, y, z)	(x, y, z)	The smallest values in acceleration and gyroscope data along all three axes
Max	(x, y, z)	(x, y, z)	The largest values in acceleration and gyroscope data along all three axes
Mean	(x, y, z)	(x, y, z)	The average of the values from acceleration and gyroscope data along all three axes
Standard Deviation	(x, y, z)	(x, y, z)	The standard deviation of the values from acceleration and gyroscope data along all three axes
Root Mean Square	(x, y, z)		The square root of the average of the squares of the values from accelerometer data along all three axes
Zero Crossings	(x, y, z)		The number of times the values from accelerometer data along all three axes changes sign
Binned-average	(x, y, z)		The values of accelerometer data are divided into 4 equally-sized bins and the average of the values in each bin is added to the vector

4 EXPERIMENTAL RESULTS

We used WEKA 3.7.11 data mining toolkit [WF05] to perform the tests on feature vectors extracted from each window. 6 different classifiers, most of which were also used in similar previous studies, with different feature vectors were tested. The classifiers consist of Bayesian Network, Multi-layer Perceptron, K-Star, Classification via Regression, Bagging and Logistic Model Tree. The tests were performed on the entire dataset using a 10-fold cross validation[1] with default parameters of WEKA toolkit. The default window size for all the classifiers is 2 seconds ($\approx$ 200 samples) and the default feature vector is the base feature vector of size 30.

Among 6 different classifiers we tested with the base feature vector, the success rate is always above 90% except for Bayesian Network. These success rates are given in Table 3. The results show that with the overall success rate of 98.16%, K-Star, similar to the previous studies, is the best fit for the considered paradigm of activity recognition system. Confusion matrix for K-Star classifier (using the base feature vector) with a window size of 2 seconds can be found in Table 4.

Table 3: Success Rates of Different Classifiers with the Base Feature Vector

	Classification via Regression	Bagging	Multi-layer Perceptron	K-Star	Bayesian Network	Logistic Model Tree
Overall Success Rate	90.45%	91.94%	90%	**98.16%**	77.27%	90.67

In all the classifiers, we observed that ascending stairs and descending stairs activities always had the lowest performances. They were below 80% for Multi-layer Perceptron. So, Multi-layer Perceptron's accuracy would have been much better if these actions were not included. These actions are especially mixed either with each other or with walking. Example graphs of each action is given at the end of the paper in Figure 2.

For different window size of 1, 2, 4 and 6 seconds, only 1 second window size had a small decrease in the success rate for K-Star (with the base feature vector). The comparison between different window sizes is

[1] K-fold cross validation randomly divides the dataset into K equally-sized subsets and for each subset carries out the test as follows: selected subset is used as the test set and the union of the $K-1$ remaining subsets is used as the training set.

Table 4: Confusion Matrix for K-Star Classifier with 2 seconds Window Size

		Classified as						
		walking	standing	running	stair-up	stair-down	jumping	updown
Action	walking	2509	0	1	14	19	0	2
	standing	1	2012	0	1	1	0	3
	running	25	0	2545	7	12	6	1
	stair-up	45	0	7	1715	14	1	3
	stair-down	33	0	6	15	1544	2	5
	jumping	0	0	8	1	4	2156	0
	updown	6	10	0	10	8	0	2020

given in Table 5. We also observed that binned-average feature and considering gender as a feature do not affect the success rates much. The highest increase, when they are both added to the base feature vector, was in Multi-layer Perceptron by 2%.

We also tested if a gender recognition could be performed using the same methodology, which was not considered in previous studies. A unique activity tag (a number between 1 and 7) for each activity was added to the base feature vector, and gender was to be classified. Again, the 2-second windows were used with $2/3$ of the data for training and the remainder for test and the success rate was 98.62% with K-Star.

The last experiment performed was simulating different sampling frequency levels[2]. The results show that increasing frequency level from 25Hz up to 100Hz slightly increases the performances. This suggests that a small fraction of data contains enough information for successful activity recognition. This could be useful in the cases where the power consumption of the device is critical.

Table 5: Success Rates of Different Window Sizes with K-Star and the Base Feature Vector

	1 second	2 seconds	4 seconds	6 seconds
Overall Success Rate	95.92%	98.16%	98.33%	98.01%

Table 6: Success Rates of Different Sampling Rates with K-Star, 2-second Windows and the Base Feature Vector

	25Hz	50Hz	100Hz
Overall Success Rate	97.41%	97.87%	98.16%

5 CONCLUSION

We provided detailed analyses regarding human activity recognition via smart phone sensors. On a database with more than 100 subjects, we were able to obtain recognition rates exceeding 98%. After evaluating the performance in terms of confusion matrices, effects of time-window sizes and different

[2]The simulation for 25Hz is done by omitting last 3 sampled values of every 4 consecutive data values. The simulation for 50Hz is done by omitting every other sampled value.

feature vectors in this paper, in the future we plan to work on personal identification ("smart phone biometrics") on the collected database. Further, we plan to add other complex activities (in addition to the tested "getting in and out of a car" activity), to see how the recognition rates scale for such sequences, essentially made up of basic activity atoms. If such complex activities can be recognized with acceptable accuracy, the results can be utilized for automatic social networking feeds (albeit with possible privacy issues), health-related monitoring, calculating daily calorie intakes, etc. Such activities can span entertaintment/sports (jogging, hiking, playing video games...), daily tasks (cleaning, cooking...), or occupational tasks (e.g., teaching).

References

[App] Apple. Core Motion Framework Reference. `https://developer.apple.com/library/ios/documentation/CoreMotion/Reference/CoreMotion_Reference/CoreMotion_Reference.pdf`. [Online; accessed 20-July-2014].

[BGC09] T. Brezmes, J.L. Gorricho, and J. Cotrina. Activity Recognition from Accelerometer Data on a Mobile Phone. In *Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living*, volume 5518 of *Lecture Notes in Computer Science*, pages 796–799. Springer Berlin Heidelberg, 2009.

[BI04] L. Bao and S. S. Intille. Activity Recognition from User-Annotated Acceleration Data. In *Pervasive Computing*, volume 3001 of *Lecture Notes in Computer Science*, pages 1–17. Springer Berlin Heidelberg, 2004.

[DDK+12] S. Dernbach, B. Das, N. C. Krishnan, B.L. Thomas, and D.J. Cook. Simple and Complex Activity Recognition through Smart Phones. In *Intelligent Environments (IE), 2012 8th International Conference on*, pages 214–221, June 2012.

[DHS00] R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern Classification (2nd Edition)*. Wiley-Interscience, 2000.

[GFH09] N. Györbíró, Á. Fábián, and G. Hományi. An Activity Recognition System for Mobile Phones. *Mob. Netw. Appl.*, 14(1):82–91, February 2009.

[HTM11] A. Henpraserttae, S. Thiemjarus, and S. Marukatat. Accurate Activity Recognition Using a Mobile Phone Regardless of Device Orientation and Location. In *Body Sensor Networks (BSN), 2011 International Conference on*, pages 41–46, May 2011.

[KWM11] J. R. Kwapisz, G. M. Weiss, and S. A. Moore. Activity Recognition Using Cell Phone Accelerometers. *SIGKDD Explor. Newsl.*, 12(2):74–82, March 2011.

[LCB06] J. Lester, T. Choudhury, and G. Borriello. A Practical Approach to Recognizing Physical Activities. In *Pervasive Computing*, volume 3968 of *Lecture Notes in Computer Science*, pages 1–16. Springer Berlin Heidelberg, 2006.

[RDML05] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman. Activity Recognition from Accelerometer Data. In *Proceedings of the 17th Conference on Innovative Applications of Artificial Intelligence - Volume 3*, IAAI'05, pages 1541–1546. AAAI Press, 2005.

[SZL+10] L. Sun, D. Zhang, B. Li, B. Guo, and S. Li. Activity Recognition on an Accelerometer Embedded Mobile Phone with Varying Positions and Orientations. In *Ubiquitous Intelligence and Computing*, volume 6406 of *Lecture Notes in Computer Science*, pages 548–562. Springer Berlin Heidelberg, 2010.

[WF05] I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques, Second Edition*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2005.

[Yan09] J. Yang. Toward Physical Activity Diary: Motion Recognition Using Simple Acceleration Features with Mobile Phones. In *Proceedings of the 1st International Workshop on Interactive Multimedia for Consumer Electronics*, IMCE '09, pages 1–10, New York, NY, USA, 2009. ACM.

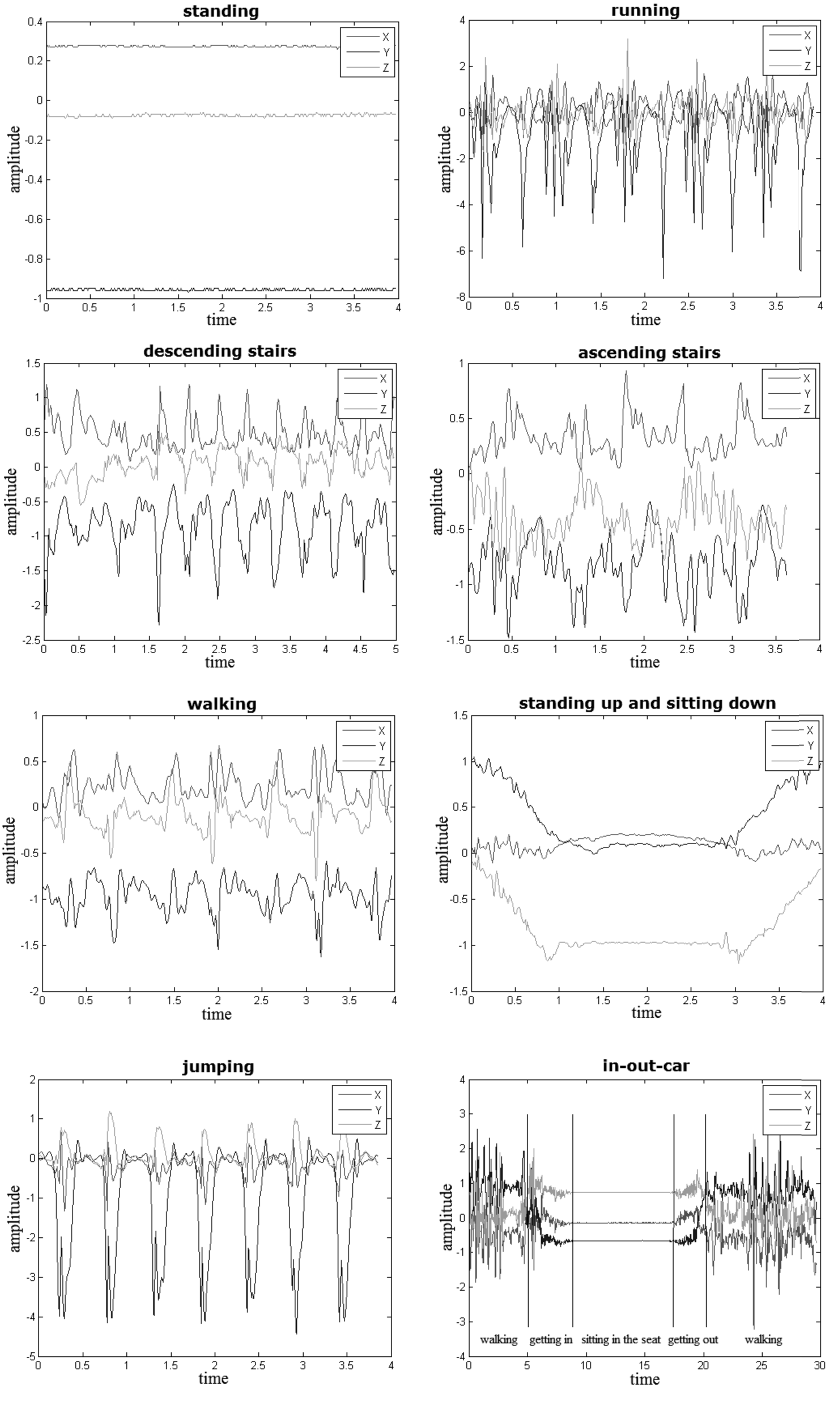

Figure 2: Graphs of main actions (performed by different subject) for a small period and of an in-out-car action