Movement recognition using the accelerometer in smartphones

2 authors, including:

Sian Lun Lau
Sunway Education Group

57 PUBLICATIONS   **304** CITATIONS

Some of the authors of this publication are also working on these related projects:

Project   Empowering People with Visual Impairments through Smart Technologies View project

Project   Seam4US View project

Movement Recognition using the Accelerometer in Smartphones

Sian Lun LAU[1], Klaus DAVID[1]

[1]*Chair for Communication Technology,*
University of Kassel,
Wilhelmshöher Allee 73, Kassel, 34119, Germany
Tel: +49 561 804 6446, Fax: + 49 561 804 6360, Email: comtec@uni-kassel.de

Abstract: The area of activity recognition is essential for context-aware systems. Previous and current investigations demonstrate that the accelerometer is suitable for accurate movement and activity recognition. Since smartphones are used by people in their daily lives, they can be seen as an attractive sensor device for the purpose of activity recognition. In our work, experiments have been carried out to investigate the suitability of the built-in accelerometer by comparing the influences of classification algorithms, features and the combination of sampling rates and window sizes for features extraction have on the classification accuracy. Obtained results indicate that smartphones similar to the test device provide good accuracy in recognizing common movements.

Keywords: context-awareness, activity recognition, classification, smartphone.

1. Introduction

In the past decades, much work has been carried out in the area of situation and activity recognition. The different approaches can be summarized as follows: By attaching different sensors to the human body and its surrounding, the acquired information will be processed, analyzed, and evaluated to interpret further information. In the area of context-awareness, this obtained information is known as contexts. A context-aware system can "understand" the situation of a user with the presence of contexts and then can provide suitable service behaviour adaptation.

The previous work explored the possibilities in using wearable sensors to recognize movement and activities. While the results showed great potential, in our opinion the placing of multiple sensors on the human body is simply not practical for the everyday user. For example, in [1] and [2] multiple sensors were placed all over the body. Some of the work was carried out in laboratory environments [2][3][4]. In order to enable such approaches to be used in our daily lives, the sensor devices have to be non-obtrusive for the users targeted. In the near future, we expect some of these prototypes to be made into miniature wearable sensors. However, even at present there are already some commercially available products such as mobile phones and especially smartphones that have the potential to be used as suitable sensor devices.

Mobiles phones are common items in most countries today [5]. Their widespread usage is advantageous for the realisation of smart adaptive systems, such as context-aware systems. The number of mobile phones that have built-in sensors and multiple connectivity possibilities are increasing throughout the past years, making these mobile phones a potential interface and mediator between users and context-aware systems [6]. For example,

one can collect accelerometer sensor values and ambient sound from a user's surrounding and send this information via Wireless LAN or 3G network to a server for processing and recording. As compared to the work of a few years ago, it is possible to realise this approach using a commercially available mobile phone such as a Nokia N95 or an Apple iPhone. These mobile phones, also commonly called smartphones, provide advanced functionality beyond voice calls and short message services.

The recognition tasks performed using a smartphone can be used in various application scenarios. For example, the system can log a user's movements through out his daily life. This can be a useful method to monitor patients or elderly people living alone. The movement contexts can be used to derive higher activity contexts such as the event a person is currently attending or the activity he is undertaking. These contexts can then be applied in context-aware systems to provide adaptive services to the user.

In this paper, we will investigate the following issues:

1. How suitable are the accelerometers found in common smartphones in the area of movement recognition as compared to approaches using dedicated accelerometers?
2. What are the possible algorithms and features that can provide accurate movement recognition using smartphones?
3. Which sampling rate and window size combination for features extraction will provide better movement recognition using smartphones?

The structure of this paper is as follows: In Section 2, a summary of related work will be presented. Section 3 will elaborate on the approach and idea, and Section 4 will present the experiments performed. The results will be discussed in Section 5 and the paper will be concluded in Section 6.

2. Related work

The earlier work investigated the possibilities to use accelerometers as main sensor data for activity recognition. Work such as [1], [2], [3], [4], and [7] demonstrated that the usage of dedicated accelerometers can provide good results in the area of activity recognition. The different investigations have one thing in common - multiple accelerometers were placed on different parts of the body, either wired or wireless, and users were required to perform designated movements. The recorded accelerometer values were processed and the resulted features were evaluated using classification algorithms for potential recognition.

The ideas were expanded with the inclusion of additional sensor information. For example, in [8] a heart rate monitor was coupled with data taken from five accelerometers to detect physical activities. The team at Intel Research in Seattle and University of Washington used the multi-modal sensor board (MSB) that had accelerometer, audio, temperature, IR/visible/high-frequency light, humidity, barometric pressure and digital compass [9]. They investigated activity recognition classification of physical activities with multiple MSBs. The group in [10] used a tri-axial accelerometer together with a wearable camera to recognize human activity. Note that these dedicated accelerometers used in the above work are capable of producing sampling rate more than 100Hz and are accurate up to ±10G. In our tests, we managed to obtain sampling rates around 60Hz-70Hz using a Nokia N95 8GB. Most of the above investigations used accelerometer sampling rates from 30Hz to 50Hz, except for [1] and [2] that sampled the accelerometer data at 76.25Hz and 93Hz respectively.

The above research has obtained good results with recognition accuracies between 83% and 99%. These investigations have shown that by analyzing the accelerometer data, systems were able to provide good recognition. However, these set ups required multiple accelerometers to be worn and observed. For a typical real user, it can be rather obtrusive and troublesome to wear multiple sensors or sensor boards at specific positions. This factor also partly motivated us to use a smartphone as a non-obtrusive sensor source. The

inclusion of other sensors did give a slight improvement on accuracy rate, but two issues remain outstanding. Firstly, not all additions of different sensors improved accuracy significantly. Secondly, additional sensors also mean more data to be processed and computed. Therefore, an appropriate selection of sensors that provide good recognition will make for the ideal solution.

Recently, the three-dimensional accelerometer integrated in smartphones was also investigated as a potential sensor for movement recognition. In [11], the accelerometer of a Nokia N95 was used as a step counter. The results showed that such smartphones can provide accurate step-counts comparable to some of the commercial, dedicated step counter products, provided the phone is firmly attached to the body. The DiaTrace project [12] uses a mobile phone with accelerometers for physical activity monitoring. The proof-of concept prototype obtained a recognition accuracy of >95% for activity types of resting, walking, running, cycling and car driving. Brezmes et. al. used also the accelerometer data collected with a Nokia N95 with K-nearest neighbour algorithm to detect common movements [13]. These investigations mentioned that the accelerometer sampling rate on smartphones are generally lower as compared to dedicated sensors. The obtained results also showed the potential of smartphones being used for activity recognition. However, there was no comparison performed in identifying suitable algorithms, sampling rates, and features.

The results above have laid down a good foundation for the work here. For activity recognition using a commercial product such as smartphones, one of the many challenges is to find out the relevant criteria such as algorithm and feature extraction methods that will enable successful recognition and utilization of such context information.

3. Approach

We would like to investigate the implementation of movement recognition using smartphones based on the above mentioned approaches and experiences. The usage of a smartphone as a sensor device has the following advantages:

1. The available sensors are built-in. As long as the desired context can be derived and recognized from the data of the built-in sensors, the users are not required to use external sensors in order to collect needed information. In cases where the need occurs, additional sensors and devices can be interfaced to smartphones to extend necessary sensors other than the ones built-in. The flexibility and readiness of the smartphone as a sensor device are seen as advantages.
2. The smartphones of today have many properties that enable context-aware related implementations. Most smartphones have relatively high processing power and sufficient memory for data processing tasks. They also contain more than adequate storage space to store data and computed information. The smartphones also provide communication possibilities that allow information exchange between user and external services. The smartphone itself can be seen as a small computing device with common connectivity integrated.
3. A smartphone is likely to be with a user during his daily activities. It can be seen as a natural choice of a non-obtrusive device. The chances of users feeling awkward or uncomfortable will be much lower as compared to approaches that affect the usage habits of the users.
4. Most smartphones have also relatively long operation durations. For an average user, under normal usage patterns (some daily phone conversations and text messages), a smartphone should have at least a day's operation time before a recharge is required. With proper management for sensor data polling, a whole day sensor data collection is achievable.

In this paper, experiments were carried out to investigate movement recognition with supervised classification algorithms using a smartphone with a built-in accelerometer. As

mentioned earlier, the accelerometers in smartphones are not able to produce high sampling rates such as those in [1] and [2]. Therefore, applicable combinations of sampling rates and window sizes for feature extraction should be investigated. The details on the selection of algorithms, features, sampling rates and window sizes are presented in the next section.

It is also important to know how the location of the smartphone will affect the recognition accuracy. From the study carried out in [5], it is observed that a large number of men keep their mobile phones in their trousers, while the women in their bags. Previous work showed that position around the waist or slightly under the waist (e.g. the trouser pocket) will bring promising results. Therefore, for our first implementation, we expect the user to carry out his daily activity with his smartphone kept in a regular position, such as his trouser pocket or at his waist.

The idea of movement recognition on smartphones should provide value-added benefits especially for the user. On one hand, it can be used in a context-aware environment to provide adaptive services. On the other hand, the benefits can also further motivate users to actively interact with and use the system. Similarly, the selected recognition methods should also provide good accuracy. Constant falsely recognized contexts will eventually cause users to stop using the system.

Based on the above requirements, experiments were carried out to investigate the issues mentioned in Section 1. The following sections will elaborate the experiments carried out and the obtained results.

4. Experiments

The experiments have the goal to enable movement detection using accelerometer data collected from a smartphone. A Nokia N95 8GB smartphone was chosen as the test device. It has the Symbian S60 3rd Edition FP1 operating system and has a built-in triaxial accelerometer. The highest sampling rate achieved when all three axes readings were recorded on the phone's mass storage was around 60Hz-70Hz. The recording was performed using a script written in Python for S60.

In our experiments we have chosen five common movements, typically also used in the references of section 2, and which can be observed in daily lives - walking, standing, sitting, walking up and down the stairs. This choice of movements facilitates a comparison to previous work, i.e. how well and suitable are smartphones in movement acceleration measurements.

A Nokia N800 Internet Tablet was used to manually annotate the movements performed. The test person was required to select a corresponding movement annotation using a dedicated application on the N800 before the movement was carried out. This gave an approximation of the duration of each annotated movement. The data is later combined using a script to create an annotated record of movements with the respective accelerometer readings. As mentioned by [12], a sampling rate of 32Hz is considered as sufficient for potential recognition of body movements using accelerometer. Therefore, the accelerometer raw data was processed to produce simulated samples with sampling rates of 5Hz, 10Hz, 20Hz and 40Hz. The smaller sampling rates were chosen to investigate whether the smaller rates can produce equally good recognition accuracies.

The sliding window technique was used for feature extraction. Features were then computed from these samples using different window sizes. Selected window sizes were 5, 10, 20, 40 and 80 samples per window. The combination of different sample rates and window sizes will represent windows equivalent to 0.5, 1, 2 and 4 seconds. The experiments used windows with 50% overlap as perform in work such as [1] and [7].

The accelerometer data and the respective annotation were carried out in two settings. In both settings, the test person performed the selected actions in a sequence (e.g. sitting, walking, standing, walking, stairs up and walking). In the first setting (S1), the smartphone

was placed in the right trouser pocket at a fixed position. For the second setting (S2), the smartphone was placed in the same pocket as in S1 but without a fixed position. In this case, the phone might turn a little when the test person moves around.

The samples collected according to each setting will be classified to see how well can recognition be achieved respectively. Additionally, data from both settings will be combined to investigate whether the recognition accuracy will be greatly affected. The accuracy is determined by calculating the percentage of the number of correctly classified instances with respect to the total number of instances.

The following sub-sections will elaborate how the classifications were carried out.

4.1 Selection of features

Some common features used in previous work are mean, standard deviation, energy of the Fast Fourier Transform (FFT), or correlation. We think it is of interest to understand whether it is still possible to obtain comparable good results with a minimum number of features. Therefore, we have restricted our selection to the following four features: the mean and standard deviation of the accelerometer raw data, and the mean and standard deviation of the FFT components in the frequency domain

Three combinations of features were chosen in the classification evaluation process. The combination C1 included only average and standard deviation of accelerometer values of each axis and all three axes. For combination C2, only average and standard deviation of FFT coefficients of each axis and all three axes were taken in consideration. The combination C3 included all four features of each axis and all three axes.

4.2 Classification

In order to verify the practicability of the above decisions, the sensor data was classified using selected supervised classification methods. Decision Tree (DT), Bayesian Network (BN), Naïve Bayes (NB), K Nearest Neighbour (KNN) and Support Vector Machine (SVM) were chosen because the previous work showed good results using these algorithms. On top of these classifiers, we also have added also rule based learner (Jrip) and Sequential Minimal Optimization (SMO) for the purpose of comparison.

The classification evaluations were carried out using the 10-fold cross-validation method. This will provide an estimation how well the models will perform generally. The above classifiers were available in the Weka Toolkit [14]. Each combination of features was calculated using different combinations of sampling rate and windows size. The produced features together with the respective movement annotations were then analysed and evaluated using Weka. The classification evaluation tests were carried on a virtual machine with a 2.5Ghz CPU and 1GB RAM. In the next section, the results will be presented and discussed.

5. Results and discussion

The results were analyzed in the following manner: Firstly, the results for both settings were compared to see which classifiers gave the best recognition accuracies. Secondly, a comparison of the influence of sampling rates and window sizes was made among the best results. Thirdly, an evaluation was carried out with the built models to test an additional test set. Both first and second analyses were done based of the evaluation results from the 10-fold cross validation evaluations.

5.1 Classification evaluation and influence of features

The result of classifications showed that C1 and C3 gave best recognition accuracy up to 99.27% (Setting S1) and 96.59% (Setting S2) for all five movements. Among these best results are the combination of all three axes and the Z-axis. The latter measures the vertical movement of the test person. Therefore, these two combinations provided better results than single axis sampling of X- or Y-axis.

By comparing the best results (e.g. as shown in Table 1), we observed that the combination C1 generally gave very good recognition accuracy for all movements. Among all classifiers, the KNN gave better accuracy in many combinations. Other algorithms such as DT, NB, and BN gave also relatively good results.

Table 1: The result of the classifier evaluation for setting S1

Combination	Algorithm	Sampling Rate (Hz)	Window Size (samples)	Accuracy (%)
C1 - all axes	KNN	20	80	99.27
C1 - all axes	NB	20	80	97.81
C1 - all axes	DT	20	80	97.08
C3 - all axes	KNN	20	80	95.62
C3 - all axes	DT	20	80	95.62
C3 - all axes	BN	20	80	95.62
C1 - all axes	KNN	10	40	95.42
C1 - all axes	DT	10	40	95.42

Table 2: The result of the classifier evaluation for combined settings

Combination	Algorithm	Sampling Rate (Hz)	Window Size (samples)	Accuracy (%)
C1 - all axes	KNN	20	80	96.52
C1 - all axes	Jrip	20	80	93.53
C1 - all axes	BN	20	80	93.53
C1 - all axes	KNN	20	40	93.19
C1 - all axes	NB	20	80	93.03
C3 - Z-axis	KNN	20	80	93.03
C3 - all axes	Jrip	20	80	92.54

For the case where data from both settings were combined, the movement recognition accuracies were slightly lower. In Table 2, KNN gave rather high accuracy. Again, the combination C1 gave the best accuracies. The rule learner Jrip performed relatively well also.

As observed above, the combination C1 gave better results than C2 and C3. In other words, the mean and standard deviation of accelerometer data are features that can produce good recognition. The addition of FFT features also obtained good results, and are slightly lower in terms of accuracy as compared to the combination C1. Processing and recognition running on the smartphone itself are potentially achievable when the selected features and algorithms are simple and efficient. In this way, remote processing and evaluation can be excluded.

5.2 Influence of sampling rates and window sizes

It was observed that the best sampling rate and window size combinations are the combinations of samples rates 20Hz with 80 samples per window, 10Hz with 40 samples per window and 40Hz with 80 samples per window. The highest accuracy for above combinations was 99.27%. No matter if it is setting S1, S2 or the mixture of both, 20Hz with 80 samples worked best with algorithms such as KNN, BN, NB and Jrip. The samples

with the combinations of 20Hz with 40 samples per window and 5 Hz with 10 samples per window (2 seconds) gave also good accuracy (93.19% and 91.16% respectively) when evaluated with the KNN algorithm.

In cases where accelerometers are not able to produce higher sampling rate, we expect the maximum accuracy to be lower. In our experiments, it was possible to obtain an accuracy of up to 94.26% (Table 3).

Table 3: The result of the classifier evaluation with sampling rate of 5Hz for setting S1

Combination	Algorithm	Sampling Rate (Hz)	Window Size (samples)	Accuracy (%)
C1 - all axes	DT	5	20	94.26
C3 - all axes	Jrip	5	20	92.62
C3 - all axes	DT	5	20	92.62
C1 - all axes	KNN	5	20	91.80
C1 - all axes	KNN	5	10	91.16

The evaluation also indicated that lower sampling rates are able to provide good accuracies. It is possible to obtain >90% accuracy with sampling rates from 5Hz to 20Hz.

5.3 Accuracy of the built models

The above built models and identified combinations of sampling rate, window size and algorithms were tested against a new set of test data. Similar sequence of movements was recorded using the same smartphone and was evaluated using the models built from experiment setting S1. The results are shown in Table 4. The recognition accuracy of the built models against the test data was acceptable (up to 91.95%). In most cases, movements going up and down the stairs had lower accuracy. Other movements were almost 100% for the best results.

Table 4: The result of the classifier evaluation using test data against
the built model from setting S1

Combination	Algorithm	Sampling Rate (Hz)	Window Size (samples)	Accuracy (%)
C1 - Z-axis	KNN	10	40	91.95
C1 - Z-axis	KNN	20	80	90.70
C1 - Z-axis	NB	20	80	90.70
C1 - Z-axis	NB	10	40	89.66
C3 - Z-axis	KNN	20	80	89.53
C3 - all axes	KNN	10	40	88.64
C3 - Z-axis	KNN	10	40	88.50

For this evaluation, it is observed that the combination C1 with KNN and BN using sampling rates of 10Hz and 20Hz and window lengths of 4 seconds gave the best accuracy. This gave an estimation of how well these combinations will perform in a real implementation.

6. Conclusions

In our work we investigated the possibility of using the accelerometer sensor in a Nokia N95 8GB smartphone to recognize common movements, i.e. walking, standing, sitting, walking up and down the stairs. From the results of the evaluation, it is observed that such a smartphone can provide accurate movement recognition. Contrary to previous work, a sampling rate of 10Hz and 20Hz is sufficient to achieve good accuracy (>90%) with the combination of only mean and standard deviation of the accelerometer values. The combination of sampling rates with window sizes of 2 and 4 seconds gave higher accuracies

than smaller window sizes. In cases where the phone is placed in a trouser pocket, even if the position is not firmly fixed, it is possible to obtain accuracy up to 96.5% using the algorithm KNN.

The results from the experiments gave us promising results. It will be interesting to implement the classification algorithms such as KNN and DT on the smartphone for learning and real time recognition. Additional movements should also be investigated in order to cover more aspects of daily activities.

Acknowledgement

References

[1] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," Pervasive 2004, pp. 1–17, April 2004.

[2] N. Kern, B. Schiele, and A. Schmidt, "Recognizing context for annotating a live life recording," Personal Ubiquitous Comput., vol. 11, no. 4, pp. 251–263, 2007.

[3] K. V. Laerhoven and O. Cakmakci, "What shall we teach our pants? ," in ISWC '00: Proceedings of the 4th IEEE International Symposium on Wearable Computers, (Washington, DC, USA), p. 77, IEEE Computer Society, 2000.

[4] J. Mantyjarvi, J. Himberg, and T. Seppanen, "Recognizing human motion with multiple acceleration sensors," in Systems, Man, and Cybernetics, 2001 IEEE International Conference on, vol. 2, pp. 747–752 vol.2, 2001.

[5] Y. Cui, J. Chipchase, and F. Ichikawa, "A cross culture study on phone carrying and physical personalization," in HCI (10) (N. M. Aykin, ed.), vol. 4559 of Lecture Notes in Computer Science, pp. 483–492, Springer, 2007.

[6] C. Frank, P. Bolliger, F. Mattern, and W. Kellerer, "The sensor internet at work: Locating everyday items using mobile phones," Pervasive and Mobile Computing, vol. 4, pp. 421–447, June 2008.

[7] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," American Association for Artificial Intelligence, 2005.

[8] E. M. Tapia, S. S. Intille, W. Haskell, K. Larson, J. Wright, A. King, and R. Friedman, "Real-Time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor," in Wearable Computers, 2007 11th IEEE International Symposium on, pp. 37–40, 2007.

[9] J. Lester, T. Choudhury, and G. Borriello, "A practical approach to recognizeing physical activities," in Proceedings of the 4th International Conference, PERVASIVE 2006, Dublin, Ireland, May 7-10, 2006.

[10] Y. Cho, Y. Nam, Y. Choi, and W. Cho, "SmartBuckle: human activity recognition using a 3-axis accelerometer and a wearable camera," in Proceedings of the 2nd International Workshop on Systems and Networking Support for Health Care and Assisted Living Environments, (Breckenridge, Colorado), pp. 1–3, ACM, 2008.

[11] M. Mladenov and M. Mock, "A step counter service for java-enabled devices using a built-in accelerometer," in CAMS '09: Proceedings of the 1st International Workshop on Context-Aware Middleware and Services, (New York, NY, USA), pp. 1–5, ACM, 2009.

[12] G. Bieber, J. Voskamp, and B. Urban, "Activity recognition for everyday life on mobile phones," in HCI (6) (C. Stephanidis, ed.), vol. 5615 of Lecture Notes in Computer Science, pp. 289–296, Springer, 2009.

[13] T. Brezmes, J. Gorricho, and J. Cotrina, "Activity recognition from accelerometer data on a mobile phone," in Proceedings of the 10th International Work-Conference on Artificial Neural Networks: Part II: Distributed Computing, Artificial Intelligence, Bioinformatics, Soft Computing, and Ambient Assisted Living, (Salamanca, Spain), pp. 796–799, Springer-Verlag, 2009.

[14] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The weka data mining software: An update," SIGKDD Explorations, vol. 11, 2009.