Real-Time Activity Recognition Utilizing Dynamically On-Body Placed Smartphones

Marc Kurz, Bernhard Hiesl, Erik Sonnleitner

University of Applied Sciences Upper Austria
Faculty for Informatics, Communications and Media
Department of Mobility and Energy
4232 Hagenberg, Austria
{firstname.lastname}@fh-hagenberg.at

Abstract—**This work-in-progress paper presents a project that deals with real-time recognition of people's activities by utilizing commercial smartphones. One important and crucial aspect is that the phone can be placed on various on-body positions (with different orientation and rotation), thus the system has to adapt autonomously to the current phone-location. There are many possibilities to purse the smartphone - for example, facing downwards to your body, facing up away from your body, left or right pocket or even back pocket. The application has to adapt to these variations of the on-body positions to fulfill the activity recognition task in real-time. The activities that are considered in this paper are five common modes of locomotion: (i) standing, (ii) walking, (iii) running, (iv) stairs-up and (v) stairs-down. The paper presents the problem definition, reflects related work on this topic, showing the relevance of the project, and discusses the intended approach, expected results and already conducted work.**

***Keywords*–Activity recognition; Mobile sensing; Self-adaptation; Adaptive application; Adaptive real-time strategies.**

I. Introduction

Nowadays mobile phones are manufactured with strong processors, good cameras, sharp displays and precise sensors (e.g., accelerometer, gyroscope, magnetometer, GPS, etc.). The composition of these elements gives mobile platforms a vast playground for mobile developers to create applications in areas such as social media, entertainment and personal usage. More and more people tend to use their mobile phone on a daily basis, which transforms the device into a constant companion. Therefore, applications running on mobile phones could gather a huge amount of information about the user. For example, this might include the usage of the smartphone, current context or even the correct recognition of the activity the user is performing [1][2].

This paper presents a work-in-progress project dealing with the real-time recognition of people's activities utilizing a commercial smartphone. This idea is not new and has been subject to research in numerous previous publications (e.g., [3][4][5][6][7][8][9][10], see Section II - Related Work - for further details). The challenging and novel aspect is that the recognition of activities shall be rotation-, orientation- and position-independent - thus the system has to autonomously adapt to changes in the phone's position and orientation regarding the on-body placement. The five considered on-body positions for the smartphone are illustrated in Figure 1 (i.e., left-, right- front pocket, left-, right-, back-pocket, shirt-pocket). We have identified those five body positions since they are realistically used by users to carry the phone when not actively used. Furthermore, the phone could be rotated and oriented differently - thus the placement of the phone on the body of the user is very important. Common machine learning technologies are usually trained under laboratory conditions, presuming constant conditions (like position, location, etc.). This is the challenging aspect, since we also do not want to force the user to conduct a preliminary calibration - the system shall be able to self-adapt to the phone's position and orientation autonomously upon the recognition task. The relevant activities that are currently of interest are - for the sake of simplicity - reduced to five modes of locomotion: (i) standing, (ii) walking, (iii) running, (iv) stairs-up and (v) stairs-down. Possible real-world use cases and scenarios for such applications could be (i) personal logging (sports monitoring - e.g., answering questions whether the user has been physically active enough during a period of time), (ii) health-care, (iii) recommender systems, etc.

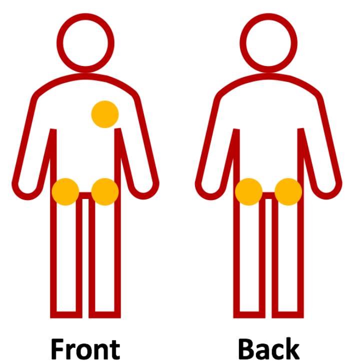

Figure 1. Possible on-body phone positions (i.e., left-, right- front pocket, left-, right-, back-pocket, shirt-pocket).

The research challenge can be summarized as follows:

> *How can highly accurate real-time activity recognition be realized utilizing a dynamically (i.e., rotation, position and orientation independent) on-body placed commercial smartphone?*

This paper provides a work-in-progress summary tackling this research challenge discussing related work (Section II), the methodological approach (Section III), an overview of the current status and preliminary results (Section IV). The paper closes with a conclusion and an outlook to future work in Section V.

 ISBN: 978-1-61208-706-1

II. RELATED WORK

Prior investigations have proven that activity recognition using mobile phones or accelerometers placed on the body is feasible in a decent way [3][4][5]. However, nowadays there are countless variations of different implementations worldwide. In this paper distinction is made between activity recognition that is position and orientation dependent and activity recognition that is position and orientation independent.

Even though Bao and Intille [4] wrote their paper in 2004, it is still one of the most cited academic sources in the activity recognition domain. The reason for this is that they developed the first semi-naturalistic activity recognition system using five bi-axial acceleration sensors mounted on different body positions. Semi-naturalistic means that participants were not supervised executing the activities they were asked to perform. Therefore, activities were performed in a more natural way and it is possible that the execution of the activity varied as participants did not feel as subjects being observed. Furthermore, Bao and Intille [4] indicate that only two accelerometers and low-level features (mean, min, max) are sufficient to recognize an activity correctly using Decision Tree algorithms. Another approach worth mentioning is demonstrated by Ravi et al. [11], who use a single triaxial accelerometer near the pelvic region for the approach. Instead of decision trees, a meta level classifier classifies the activities. Other authors such as Kwapisz et al. [7], Anjum et al. [3] and Derawi and Bours [6] use mobile phones with built in tri-axial accelerometers in their position and orientation dependent activity recognition approaches. Kwapisz et al. [7], for example, provide only one position and orientation of the mobile phone. Furthermore, Anjum et al. [3] also stated that after evaluation of different classifiers including K-Nearest Neighbors (KNN), Support Vector Machines (SVM) and Naive Bayes, Decision Trees are performing best on single tri-axial accelerometers.

Similar approaches, which correspond to the topic of this paper, were developed by Yang [12], Sung et al. [13], Henpraserttae et al. [14] and Ustev and Durmaz Incel [15]. However, the solution of each investigation was implemented in a different unique way. To build a position and orientation independent system, Sun et al. [13], for example collected sensor data with varying positions and orientations using a mobile phone. Yang [12] instead extracts the vertical and horizontal movement of the mobile phone. Therefore, the application is no longer limited with respect to orientation. Additionally, Yang [12] used Decision Trees for the classification process. Henpraserttae et al. [14] have a more computational approach to achieve a position and orientation independent activity recognition system. They apply a projection-based method. The data acquisition transports each new record into the same coordinate system by applying a matrix multiplication with a reference matrix gathered with the magnetometer.

However, each of the named projects needs model updates when it comes to phone positions which have not been considered yet. The self-adaptiveness towards the orientation, position and location of the on-body phone placement is the novel aspect differentiating this research work from related work.

III. METHODOLOGY

Prior to implementation, machine learning technologies and suited algorithms are being evaluated with focus on the special purpose of adapting autonomously to the on-body sensor position. Furthermore, an Android app has been implemented for collecting data of subjects in order to analyse the specific activities and the corresponding algorithms for the different steps of the recognition process (i.e., signal processing, feature extraction, classification). This analysis has to be done with respect to the dynamic aspect of the sensor placement. The following Sections III-A to III-D summarize the methodology for the different relevant aspects. In Section IV, an overview of the current status and preliminary results is given.

A. Data Collection

Sensor data was collected for 15 subjects performing the 5 activities carrying the the smartphone on the five different on-body locations as illustrated in Figure 1, using 4 different orientations each. Since each recording took around 10 seconds, the total amount of recorded sensor data is approx. 4,5hrs. For recording, a simple phone application was implemented that allows for selecting the recorded activity (i.e., instant labelling of the sensor data is possible with this approach) and the on-body position of the phone (see Figure 2).

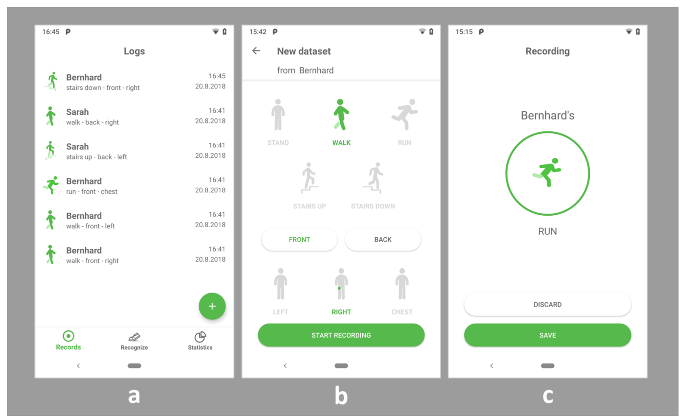

Figure 2. The application used for recording the data. (a) shows an overview of recorded data, (b) shows the selection of activity and phone-location prior to recording, (c) shows the screen to save or discard a recorded session.

The recorded modalities of the sensor data are (i) accelerometer, (ii) gyroscope and (iii) magnetometer at a rate of 100Hz. We did not consider further modalities since these three are so common that every commercial smartphone is capable of delivering these modalities. An example of raw recorded acceleration data for the activity *walk* is depicted in Figure 3.

B. Feature Extraction

As related work and projects show, low level features (which are easily computable) are significant enough to classify the correct activity [4][7][16]. For the feature extraction, a sliding window approach will be implemented and evaluated with the length of two seconds and an overlapping of 50%. Prior investigations show that this is a suitable sliding window size for activity recognition [4][13]. Usually, within two seconds at least one repetition of an activity should be completed. Inside of each window, the features as listed below will be extracted for each acceleration axis:

- **Mean:** The mean value of each window and each axis.
- **Max:** The max value of each window and each axis.

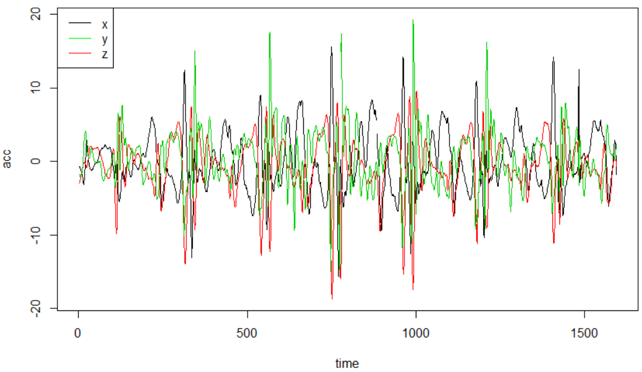

Figure 3. Exemplary data representing the activity *walk*.

- **Min:** The min value of each window and each axis.
- **Standard deviation, Variance:** The variance/standard deviation of each window and each axis.
- **Correlation:** The correlation between each pair of axes.
- **Energy:** The energy is calculated as the sum of the squared discrete Fast Fourier Transform (FFT).
- **Entropy:** The entropy is calculated as the normalized information entropy of the discrete FFT.

Subsequently, the feature vector will be used to train a machine learning model using Weka 3.8 [17]. Again, previous investigations regarding this topic show that low level features are significant enough to classify the correct activity [4][7][16].

C. Classification

The classification of activities combines the data collection and feature extraction and tries to map the recorded sensor data into an output class (i.e., the recognized activity). A first evaluation (related work research and first tryouts with the recorded dataset) of algorithms for classification shows that the following seem to be promising methods for our approach:

- K-Nearest Neighbors (KNN)
- Decision Tree (J48)
- Naive Bayes
- Support Vector Machines (SVM)

Since we have two crucial requirements for our application, namely (i) real-time recognition of activities, and (ii) adaptation to the dynamic placement of the smartphone, the algorithms for feature extraction and classification shall be robust, easily implementable and performant. Furthermore, the models that build the base for the classification process need to be small enough to be processed and calculated on mobile devices. This is the reason why we consider rather "classical" approaches (e.g., KNN, SVM, etc.) rather than recent methods like neural networks [18].

Once the classification models are created, they will be integrated into the system. As already mentioned, the classification process shall be executed in real-time. Therefore, recording and feature extraction will be needed in this process as well. In order to achieve this, the sliding window was used again. After the user starts the automatic recognition of the application, the sliding window algorithm calculates and extracts all required features of the recorded acceleration data. Due to the window size of the algorithm, it takes about two seconds to recognize the activity in real-time.

D. Dynamic Adaptation

The dynamic adaptation of the system tackles the problem of different positions, locations and orientations of the on-body placed smartphone. Every person might carry the phone differently, meaning that the sensor data might look different. This is obviously a problem for classification algorithms since supervised learning methods work in a way that they compare trained models with the real-time (preprocessed) sensor data. To achieve highly accurate real-time activity recognition with data coming from dynamically placed phones, the system has to be capable of adapting to this position-, location-, and orientation-dynamic. Compared with the previously mentioned algorithms for feature extraction and classification the following approaches seem to be promising for dynamic adaptation:

(i) Magnitude of acceleration data.
(ii) 2D projection of sensor data, respectively horizontal/vertical acceleration.
(iii) Matrix multiplication to normalize the sensor data.
(iv) Quaternions as representation of rotations in 3D space compared with Euler angles [19].

Again, it is important to achieve the activity recognition task in real-time (i.e. getting instant feedback within 1-2 seconds) and that the user is not being forced to artificially mount the smartphone at a previously defined position. Furthermore, forcing the user to execute a calibration task prior to the recognition of activities is not desired.

IV. Current Status & Preliminary Results

The current status of the project is that the sensor data from 15 subjects (8 male, 7 female, all between 22 and 30 years old) has been recorded and analysed. Feature extraction and classification methodologies have been analysed and we were able to limit the number of suitable algorithms. We have learnt that for our specific use cases of real-time recognition and dynamic adaptation, computationally expensive algorithms like neural networks [18] are not suited.

We have implemented and used a recording app (see Figure 2), which collected sensor data from a smartphone. Subjects participating in this study were asked to put the phone in their pockets in different positions and orientations while performing the five varying activities (modes of locomotion). To achieve the best approach concerning performance and recognition, different classifiers such as KNN, Naive Bayes, SVM and Decision Tree were developed using the machine learning program Weka 3.8 [17]. The trained models were compared with each other. After evaluation it became evident that Decision Tree was the most suitable model applicable to this project. The reason for this assumption is based on the model's overall result of 94% accuracy. Nevertheless, KNN, SVM and Naive Bayes will be further evaluated and considered since related work shows that these algorithms are also performing well. Evaluations of the real-time approach have shown that this will be realized with the help of the sliding window procedure, which has a fixed size of two seconds and an overlapping of 50%.

The flow of recognizing activities (see Figure 4) is currently being developed and evaluated in order to achieve the real-time requirement. As already mentioned, by not considering the dynamic-aspect of on-body phone placement, the accuracy of the recognition task utilizing standard features and rather easy classification algorithms is around 94%. We are currently in the process of fine-tuning our models, algorithms and other parameters (like the sliding window approach) to further increase the accuracy.

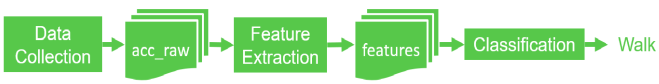

Figure 4. The subsequent steps of the real-time activity recognition task [1].

In parallel, work on the dynamic aspect has been started to achieve the system's self-adaptiveness regarding the dynamic placement of the smartphone. As mentioned, some promising approaches are currently being investigated (i.e., (i) magnitude, (ii) 2D projection, (iii) normalization of sensor data, and (iv) quaternions).

V. Conclusion and Outlook

This work-in-progress paper presents a project for recognizing people's activities at real-time utilizing nothing more than a commercial smartphone. For the sake of simplicity, the activities are reduced to the five most common modes of locomotion, which are (i) standing, (ii) walking, (iii) running, (iv) stairs-up and (v) stairs-down. Besides the requirement of recognizing the activities in real-time, another crucial aspect is the fact that the smartphone does not need to be mounted on the body of persons at a special location/position with a specific orientation. The placement of the phone shall be done in a natural way in whatever pocket the user is comfortable with. The system shall be capable of autonomously adapting to the dynamic on-body placement of the smartphone to achieve highly accurate activity recognition. Preliminary results and evaluations towards the real-time capability are promising. In order to evaluate methods and algorithms, a dataset consisting of 15 subjects performing the activities has been recorded and analysed. The dynamic placement aspect is currently subject to research, whereas different methodological approaches seem to be promising: (i) magnitude of acceleration data, (ii) 2D projection of sensor data, (iii) normalization of data by applying matrix manipulations, and (iv) quaternions as representation of rotations.

Besides future work to investigate the research challenge it is intended to extend the application further. Particularly, it is intended to include a higher number of activities, which can be considered in recording and recognition as well as to integrate more customer oriented activities. Furthermore, the current implementation only uses offline trained models. In the future, users could be able to train their personal model. In other words, future aims include the modification of the current implementation, which should allow the user to record activities, label activities and train personal models on the mobile phone.

References

[1] M. Kurz, "Opportunistic activity recognition methodologies," Ph.D. dissertation, Johannes Kepler University Linz, Jun. 2013.

[2] M. Kurz, G. Hölzl, and A. Ferscha, "Dynamic adaptation of opportunistic sensor configurations for continuous and accurate activity recognition," in Fourth International Conference on Adaptive and Self-Adaptive Systems and Applications (ADAPTIVE 2012), Nice, France, 2012.

[3] A. Anjum and M. U. Ilyas, "Activity recognition using smartphone sensors," 2013 IEEE 10th Consumer Communications and Networking Conference (CCNC), 2013, pp. 914–919. [Online]. Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6488584

[4] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in International Conference on Pervasive Computing. Springer, 2004, pp. 1–17.

[5] G. Bieber, J. Voskamp, and B. Urban, "Activity recognition for everyday life on mobile phones," in International Conference on Universal Access in Human-Computer Interaction. Springer, 2009, pp. 289–296.

[6] M. Derawi and P. Bours, "Gait and activity recognition using commercial phones," computers & security, vol. 39, 2013, pp. 137–144.

[7] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," ACM SigKDD Explorations Newsletter, vol. 12, no. 2, 2011, pp. 74–82.

[8] M. del Rosario, S. Redmond, and N. Lovell, "Tracking the evolution of smartphone sensing for monitoring human movement," Sensors, vol. 15, no. 8, 2015, pp. 18 901–18 933.

[9] K. Kunze, P. Lukowicz, H. Junker, and G. Tröster, "Where am i: Recognizing on-body positions of wearable sensors," in International Symposium on Location-and Context-Awareness. Springer, 2005, pp. 264–275.

[10] Y. He and Y. Li, "Physical activity recognition utilizing the built-in kinematic sensors of a smartphone," International Journal of Distributed Sensor Networks, vol. 9, no. 4, 2013, p. 481580.

[11] N. Ravi, N. Dandekar, P. Mysore, and M. M. L. Littman, "Activity Recognition from Accelerometer Data," Proc. The 17th Conference on Innovative Applications of Artificial Intelligence (IAAI'05), vol. 3, 2005, pp. 1541–1546. [Online]. Available: http://www.aaai.org/Papers/IAAI/2005/IAAI05-013

[12] 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. ACM, 2009, pp. 1–10.

[13] 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 International conference on ubiquitous intelligence and computing. Springer, 2010, pp. 548–562.

[14] 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. IEEE, 2011, pp. 41–46.

[15] Y. E. Ustev, O. Durmaz Incel, and C. Ersoy, "User, device and orientation independent human activity recognition on mobile phones: Challenges and a proposal," in Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication. ACM, 2013, pp. 1427–1436.

[16] X. Su, H. Tong, and P. Ji, "Activity recognition with smartphone sensors," Tsinghua science and technology, vol. 19, no. 3, 2014, pp. 235–249.

[17] M. Hall, E. Frank, G. Holmes, B. Pfahringer, P. Reutemann, and I. H. Witten, "The WEKA data mining software," ACM SIGKDD Explorations Newsletter, vol. 11, no. 1, 2009, p. 10. [Online]. Available: http://portal.acm.org/citation.cfm?doid=1656274.1656278

[18] J. Yang, M. N. Nguyen, P. P. San, X. Li, and S. Krishnaswamy, "Deep convolutional neural networks on multichannel time series for human activity recognition." in Ijcai, vol. 15, 2015, pp. 3995–4001.

[19] S. Sempena, N. U. Maulidevi, and P. R. Aryan, "Human action recognition using dynamic time warping," in Electrical Engineering and Informatics (ICEEI), 2011 International Conference on. IEEE, 2011, pp. 1–5.