2016 IEEE International Conference on Pervasive Computing and Communications (PerCom)

On-body Localization of Wearable Devices: An Investigation of Position-Aware Activity Recognition

Timo Sztyler
University of Mannheim
Mannheim, Germany
timo@informatik.uni-mannheim.de

Heiner Stuckenschmidt
University of Mannheim
Mannheim, Germany
heiner@informatik.uni-mannheim.de

***Abstract*—Human activity recognition using mobile device sensors is an active area of research in pervasive computing. In our work, we aim at implementing activity recognition approaches that are suitable for real life situations. This paper focuses on the problem of recognizing the on-body position of the mobile device which in a real world setting is not known a priori. We present a new real world data set that has been collected from 15 participants for 8 common activities were they carried 7 wearable devices in different positions. Further, we introduce a device localization method that uses random forest classifiers to predict the device position based on acceleration data. We perform the most complete experiment in on-body device location that includes all relevant device positions for the recognition of a variety of different activities. We show that the method outperforms other approaches achieving an F-Measure of 89% across different positions. We also show that the detection of the device position consistently improves the result of activity recognition for common activities.**

I. Introduction

Activity Recognition is an active field of research in pervasive computing [1]–[4]. The development of wearable devices such as smart-phones, smart-watches and fitness wristbands feature a variety of sensors that are carried all day long by many people (compare [3]) provide new opportunities for continuous monitoring of human activities such as running, sitting, standing, or walking [5]. A problem of many existing studies on the subject is that they are conducted in a highly controlled environment. In consequence, the results of these studies often do not carry over to real world applications. Our aim is to develop robust activity recognition methods based on mobile device sensors that generate high quality results in a real world setting.

The quality of activity recognition, especially in a real world setting, depends on the on-body position of the wearable device providing the sensor data [6]. Previous studies have shown that relevant on-body positions are head, upper arm, forearm, chest/waist, thigh, and shin [2]. Further, in case of an uncontrolled environment, it is also important to detect the state transitions between these positions. In this context, the acceleration sensor is the most interesting sensor to recognize the device position due to a low power consumption that enables continuous sensing over a complete day. Moreover, a lot of studies already achieved good results using the acceleration sensor in context of activity recognition [5]–[7].

In this paper, we focus on the detection of relevant on-body positions in the context of everyday movements and activities using a single accelerator sensor. Further, we examine the impact of the recognized position information on the accuracy of the activity recognition. The contributions of our work are the following:

- We present a new real world data set for on-body position detection and position-aware activity recognition.
- We perform the most complete experiment in on-body position detection and activity recognition carried out so far.
- We present a method based on a random forest classifier that detects the on-body position with an F-measure of 89%.
- We show that the recognition method consistently improves the recognition of activities in a real world setting.

The paper is structured as follows: In Section II, the related work concerning the on-body detection of wearable devices is summarized. Then, we outline the data collection phase and present our data set. Section IV covers the process of feature extraction and classification where we introduce our approach for on-body position detection and position-aware activity recognition. In Section V, we present our experimental results and outline the effect of considering the device position. Finally, Section VI covers the conclusion and future work of this paper.

II. Related Work

Researchers have already investigated activity recognition independent of the device position [8]. However, previous studies stated that position information increases the accuracy of the activity recognition but the opinion regarding the impact of this information on their results differs significantly [1], [6], [9]. This difference is due to different sets of positions and activities considered in the different studies. Indeed, so far nobody considered all relevant body positions in context of common movements and activities in one study. Therefore,

978-1-4673-8779-8/16/$31.00 ©2016 IEEE

it is still unclear how accurate each relevant position can be detected regarding different activities.

The on-body localization problem of wearable devices plays an important role because it can help to improve the accuracy of activity recognition, to optimize the energy consumption of a device, or to increase the precision of observing the environment. This is a consequence of the results of previous studies that investigated the influence of the on-body position to determine optimal sensor placement in context of activity recognition [1], [2], [6]. In these studies, it has been shown that there are seven different body locations that behave differently in activity recognition, i.e. forearm, head, shin, thigh, upper arm and waist/chest and that dividing these body parts (e.g., head or shin) into smaller regions does not improve the accuracy [2]. Further studies have shown that the optimal sensor placement depends on the activity to be recognized [6]. As a result, the benefiting of the position information and also the feasibility to derive device positions by an accelerometer is stated.

So far, the localization problem was only addressed by a couple of researchers. Kunze et al. published one of the first approaches where he first tried to detect if the user is walking and then used specific patterns of sensor readings to derive the current device position [10]. However, this approach is limited due to the small set of selected positions and the fact that position changes are not recognized if the user does not walk. Recently, researchers investigated also the possibility to derive the positions hand, bag, and pocket from different common activities [9]. In this context, they stated that the effect of the location information on the accuracy of the activity recognition depends on the performed activity.

While these studies focused on on-body position detection with an accelerometer, several researchers also examined the possibilities to detect if the phone is located in- or out-pocket [11], in a bag [9], [12], or still worn by the same person [13]. In this context, they also used other sensors such as a microphone, light, or proximity sensor. They highlight that an accurate detection is possible but also point out that it is difficult to control the environment regarding brightness or sound level which has to be considered as the crucial problem.

III. Data Set

In this paper, we investigate the detection of the on-body position of a wearable device and its influence on the quality of activity recognition. For this purpose, we created a data set[1] which covers, among others, the acceleration data of the activities climbing stairs down (A_1) and up (A_2), jumping (A_3), lying (A_4), standing (A_5), sitting (A_6), running/jogging (A_7), and walking (A_8) of fifteen subjects (age 31.9 ± 12.4, height 173.1 ± 6.9, weight 74.1 ± 13.8, eight males and seven females). For each activity, we recorded simultaneously the acceleration of the body positions chest (P_1), forearm (P_2), head (P_3), shin (P_4), thigh (P_5), upper arm (P_6), and waist (P_7). Each subject performed each activity roughly 10 minutes except for jumping due to the physical exertion ($\sim$1.7 minutes). In detail, we recorded for each position and axes 1065 minutes. Concerning male and female, the amount of data is equally distributed. To the best of our knowledge the result is the most complete, realistic, and transparent data set for on-body position detection that is currently available.

[1] http://sensor.informatik.uni-mannheim.de/

Fig. 1. Sensor placement. The subject wears the wearable devices on the head, chest, upper arm, waist, forearm, thigh, and shin (top down).

The required data was collected using customary smartphones and a smart-watch[2] which were attached to the mentioned positions (see Figure 1). The devices were synchronized with the time service of the network provider and the accelerometer was sensed with a sampling rate of 50 Hz where the data was stored on a local SD card. The sampling rate was chosen with consideration of battery life as well as with reference to previous studies [7], [14]. The recording of the data was performed using a self-developed sensor data collector and labeling framework. The framework consists of a *Wear* (1) and *Hand* (2) application (see Figure 2) which interact with each other via Bluetooth. This application provides the possibility to control the built-in sensors, specify the sampling rate, and record several sensors simultaneously. The binary[3] and the source code[4] of this application are publicly available.

To attach the devices to the relevant body positions, common objects and clothes were used such as a sport armband case, trouser pocket, shirt pocket, or the bra. There was no further fixation of the device to closely resemble their use in everyday life. In case of the head we used a belt to avoid that the subject had to hold this device during the performance of

[2] "Samsung Galaxy S4" and "LG G Watch R"

[3] https://play.google.com/store/apps/details?id=de.unima.ar.collector

[4] https://github.com/sztyler/sensordatacollector

Fig. 2. Sensor Data Collector. The framework consists of a wear (1) and hand (2) app which allows to record each sensor that is available and provides labeling and visualization functions.

the activities. This simulates that the subject phones during the activities.

The data collection took place under realistic conditions, i.e., the subjects walked through the city, jogged in a forest, or climbed up the stairs of a guard tower of an old castle. The order of the activities was left to the subjects but they were instructed to stand idle for a few seconds before and after the activity was performed. Concerning the activities, there were no instructions. It was up to the subject, e.g., how fast they wanted to walk or how they wanted to sit. In this context, typically the subjects used their smart-phone, talked with somebody else, or were eating and drinking something while they were standing or sitting.

Each movement was recorded by a video camera to facilitate the usage of our data set also by other people. Our data set is available[1] and covers beside the mentioned acceleration data also GPS, gyroscope, light, magnetic field, and sound level data which were also recorded during the data collection phase but will not considered in the following. Besides, there is also a detailed description of each subject including images of the attached devices and a short report.

IV. Method

Following most existing work, we use a supervised approach, both for on-body localization and for activity recognition. The introduced data set was used as training data and for evaluation. In the following, we describe the features generated from the sensor data and the learning methods and strategies used in our study.

A. Feature Extraction

The essential idea behind generating features from time depended data streams is to segment the recorded data into windows and compute a feature vector for each window. Preceding studies in the context of activity recognition already examined different settings regarding the window size and meaningful features [15]. They state that overlapping windows are more suitable because they can handle transitions more accurately. Further, the window size depends on the kind of activities which should be recognized. In our context, most of the existing studies considered a size between one and three seconds [4], [8], [16]. However, so far there is no agreed set of features. Indeed, a comparison of the different but overlapping feature sets of previous studies is difficult due to the different settings and goals of the studies. Nevertheless, some researchers have compared different groups of features and also stated that frequency-based features improve the accuracy of the recognition [16].

TABLE I
Summary of considered feature methods.

	Methods
Time	Correlation coefficient (Pearson), entropy (Shannon), gravity (roll, pitch), mean, mean absolute deviation, interquartile range (type R-5), kurtosis, median, standard deviation, variance
Frequency	Energy (Fourier, Parseval), entropy (Fourier, Shannon), DC mean (Fourier)

Hence, based on these results, we use windows which overlap by half and have a length of one second. Further, we consider the most common time- and frequency-based features that were used in previous work (see Table I) where time-based feature values are transformed into frequency-based ones by applying *Discrete Fourier transform*[5]. Finally, for the experiments, we performed *attribute selection* to optimize the feature vector.

In this context, we also computed gravity-based features that provide information of the device orientation. The gravity component was extracted from the recorded accelerometer data. We applied a *low-pass filter*[6] to separate the acceleration and gravitational force to derive the gravity vectors. These vectors allow to determine the orientation of the device by computing the angles between them, also known as *roll* and *pitch* (see Figure 3). The *azimuth* angle, however, cannot be calculated because the direction of north is required. This means that it is not possible to derive if the device is back-to-front. Further, we only consider absolute value of the acceleration so that we do not distinguish if the device is upside down. Hence, we consider these four cases as the same position. To be more flexible and avoid overfitting, we also transform the roll and pitch angles in one of sixteen predefined discretized orientations. Besides, the gravity-based features are only considered in the context of on-body position detection.

The feature extraction process was performed with a self-developed framework that computes all mentioned features. The framework is available[7] and allows to specify the mentioned settings. As a result, the framework returns a list of feature vectors which are in the following further processed.

[5] The Fourier transformation can be applied with different scaling factors. We use the *JTransforms* implementation (https://github.com/wendykierp/JTransforms) which scales by 1.

[6] A low-pass filter passes values which have a lower frequency as the specified *cutoff frequency* and attenuates values that have a higher frequency.

[7] https://github.com/sztyler/sensorfeatureextraction

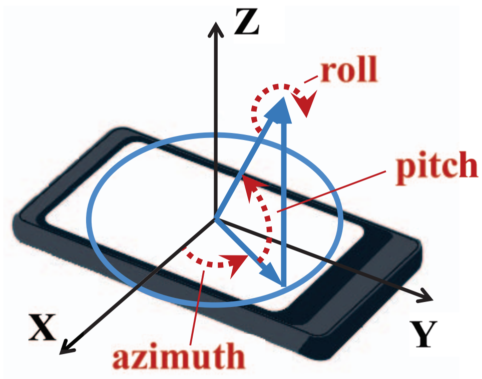


Fig. 3. The coordinate system is defined in reference to the screen. The acceleration of the device is measured along the axes. The gravity enables to compute the angle between the axes to determine the orientation (roll, pitch). To calculate azimuth, the direction of north is required.

B. Random Forest Classifiers

Decision trees have already successfully been used for activity recognition, however, it is well known that classical decision trees are sensitive to overfitting when the generated trees become very deep. In order to overcome the overfitting problem, ensemble methods have been proposed that balance the results of multiple decision trees that have been trained on different parts of the training data. Random forest classifiers are one of these ensemble methods that have been proposed by Breimann [17]. A random forest classifier is constructed in the following way:

Let $D = \{(\overline{x}_1, y_1), \cdots, (\overline{x}_n, y_n)\}$ be a learning problem with feature vectors $\overline{x}_i$ and results y_i. In a first step a number of samples $S_1, \cdots, S_m$ are drawn from D using sampling with replacement. For each sample S_i, a decision tree classifier f_i is trained using a variation of the classical decision tree learning algorithm that uses feature bagging. This means that for each branching decision in the decision tree construction only a randomly selected subset of feature vectors is taken into account. This is necessary to ensure that the different generated decision trees are uncorrelated [18]. In this context, the decision tree considers the information gain of each feature to determine the importance during the construction.

The resulting set of uncorrelated decision trees can now be used to determine the outcome for an unseen feature vector $\overline{x}'$ based on the principle of bagging. In particular, the result is determined by averaging over the predicted results of all individual decision trees as follows:

$$\hat{f}(\overline{x}') = \frac{1}{n} \sum_{i=1}^{n} f_i(\overline{x}')$$

For the case of a classification problem, the combined classifier essentially performs a majority vote over the outcomes of the individual decision trees. It has been shown that bagging prevents the overfitting problem as the combination of multiple classifiers has a significantly lower variance than an individual classifier. Due to this advantage, we focus on the random forest classifier for position but also activity recognition.

C. Position Detection

We treat position detection as a multi-class classification problem with target classes being head, upper arm, forearm, chest, waist, thigh, and shin that correspond to the relevant position according to Vahdatpour and others [2].

In initial experiments, we observed a major problem when trying to distinguish between different device positions while considering all performed activities. More precisely, data of the activities lying, standing, and sitting frequently leads to misclassification of device positions. This is caused by the fact that in context of these three activities the human body only has a slight acceleration so that the computed feature vectors are not easily distinguishable. To address this problem, we distinguish between static (standing, sitting, lying) and dynamic (climbing up/down, jumping, running, walking) activities and consider these two groups in the following as two types of activity-levels. This enables to consider different features sets. Hence, we train a classifier that distinguishes between static and dynamic activities that is used as a first step in the position detection process. A similar distinction has been made in [19] to improve the accuracy of activity recognition.

We trained both classifiers using stratified sampling combined with 10-fold cross validation to ensure that all folds cover the same ratio of classes. Further, to make the result more stable, we performed 10 runs where each time the data set was randomized and the 10-folds were recreated. The classifiers were trained and evaluated for each subject individually. Thus, we did not consider several subjects at once because of the individual behavior and the differences which result from different ages.

D. Activity Recognition

In the activity recognition phase, we aim to detect the activities climbing stairs up and down, jumping, lying, running, sitting, standing, and walking. In this context, we evaluate the impact of the information of the device position. For this purpose, we construct position-independent and position–aware activity classifiers and compare their performance on our data set.

The *position-independent* activity recognition approach simply consists of a single classifier per subject that is trained on all data independent of the device position. We expect this recognition approach to perform sub-optimal as the motion information from the sensors can be assumed to be very different in the different positions for the same activity.

The *position-aware* activity recognition approach consists of a set of individual classifiers for each device position and each subject. The classifier to be used is determined in a position recognition step that is executed before the actual activity recognition. Figure 4 provides an overview of the detection process: first the unlabeled record is classified as a dynamic or a static activity. As mentioned above, this step is necessary as we can more reliably detect the device position

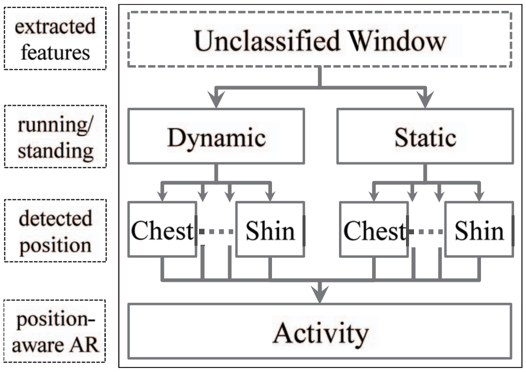


Fig. 4. Activity Recognition. The nodes illustrate the target class and the edges the applied classifier. The current window is classified as "dynamic" (climbing, jumping, running, walking) or "static" (standing, sitting, lying). Then the device position is recognized and a position specific classifier applied to derive the current activity.

if we know whether the current activity is a static or a dynamic activity. Then, the position of the device is recognized with an activity-level depended classifier that uses a feature set that has been optimized for the type of activity. Finally, the current activity is recognized by selecting and applying the classifier for the detected device position. Obviously, the performance of the position-aware activity recognition approach relies on the correct identification of the device position. Therefore, to test the feasibility of this approach, we use the results of the activity-level dependent position detection experiments - including all mistakes made - as input for the activity recognition experiments.

V. Results

In the following, we present our results and outline the conducted experiments to show the effect of the proposed device localization approach but also the influence of the derived location in context of activity recognition. The introduced methods were evaluated for each individual subject. Due to lack of space, we only present the aggregated results of all subjects. However, the individual results of each subject and classifier are available[8]. Unless otherwise specified, the provided results are based on the random forest classifier which turned out to consistently perform better than other classification techniques.

A. Position Detection

For the first experiment, we evaluated an activity-independent approach to create a baseline. Thus, we trained for each subject a single classifier on the data of all performed activities and each position. Table II shows the result and illustrates that the device position can be recognized with an F-measure of 81%. In this context, the *shin* (P_4) has the highest (88%) and the *forearm* (P_2) and *upper arm* (P_6) the lowest (79% / 78%) recognition rate. The latter highlights the problem regarding the flexibility of the arm during each activity and

[8]A complete overview of all results can be found here: http://sensor.informatik.uni-mannheim.de#results

TABLE II
Activity-independent position recognition rates for different on-body locations

Class	Precision	Recall	FP Rate	F-Measure
P_1	0.79	0.82	0.04	0.80
P_2	0.79	0.78	0.03	0.79
P_3	0.79	0.82	0.04	0.80
P_4	0.90	0.86	0.02	0.88
P_5	0.83	0.80	0.03	0.82
P_6	0.79	0.78	0.03	0.78
P_7	0.79	0.81	0.04	0.80
$avg.$	0.81	0.81	0.03	0.81

TABLE III
Activity-level dependent position recognition rates showing that the recognition performance is problematic for static activities

Activities	Precision	Recall	FP Rate	F-Measure
static	0.72	0.72	0.05	0.72
dynamic	0.89	0.89	0.02	0.89
both	0.81	0.81	0.03	0.81

also indicates that these two positions are the most problematic device locations. Examining the confusion matrix, shows that the individual positions are not mixed up. Indeed, the false-positives and the false-negatives are almost evenly distributed.

Further investigations point to the fact that the recognition rate of the correct device location is higher if the related activity is characterized by stronger acceleration. Hence, the separation between static and dynamic activities results in a significantly different recognition rates for these two kinds of activity groups (72% / 89%). As we can see in Table III the recognition rate is consistently lower for static activities (-9%).

We examined the feature set and figured out that the gravity of the device provides useful information. However, attention should be paid to the fact that our experiments also showed that the gravity vector and derived features (roll and pitch) lead to overfitting. Hence, if a classifier was trained for a specific position then the position recognition rate dropped after the device was reattached for this position. This is mainly because the orientation of the device was slightly changed by the user. Thus, the orientation seems not to be a reliable indicator of the current device position. However, investigations have shown that static activities and the device orientation are correlated. Thus, the orientation enables to separate implicit between the static activities which results in less misclassifications of the device position across these activities. In this context, we only considered the introduced discretized orientation. Table IV summarizes the results and shows that the recognition rate of the device localization in context of static activities increases by 16%.

Certainly, the usage of different feature sets for these two kinds of activity groups require the ability to separation

TABLE IV

POSITION RECOGNITION RATE FOR STATIC ACTIVITIES AND DIFFERENT FEATURE SETS SHOWING THAT ORIENTATION AND TIME-BASED FEATURES ARE NEEDED TO ACCURATE RECOGNITION

Features	Precision	Recall	FP Rate	F-Measure
time-based features	0.72	0.72	0.05	0.72
with orientation	0.88	0.88	0.02	0.88
only orientation	0.54	0.53	0.08	0.54

TABLE V

RECOGNITION RATE FOR CLASSIFIER THAT DISTINGUISHES BETWEEN STATIC AND DYNAMIC ACTIVITIES

Class	Precision	Recall	FP Rate	F-Measure
dynamic	0.98	0.96	0.02	0.97
static	0.94	0.98	0.04	0.96
avg.	0.97	0.97	0.03	0.97

TABLE VI

DETAILED RESULTS FOR THE PROPOSED POSITION RECOGNITION METHOD.

Class	Precision	Recall	FP Rate	F-Measure
P_1	0.87	0.89	0.11	0.88
P_2	0.87	0.85	0.15	0.86
P_3	0.86	0.89	0.11	0.87
P_4	0.95	0.92	0.08	0.94
P_5	0.91	0.90	0.10	0.91
P_6	0.85	0.84	0.16	0.85
P_7	0.91	0.92	0.08	0.92
$avg.$	0.89	0.89	0.11	0.89

between them. Hence, we constructed a classifier that decides to which activity group, the performed activity belongs. Table V outlines the result and clearly shows that the allocation performs very well (97%).

As a result, we evaluated the approach where we first decide if a static or dynamic activity is performed and then apply an activity-level specific position classifier. Compared to the baseline, Table VI shows that this approach has an 8% higher recognition rate. In this context, the *shin* is still the best (94%) and the arm (*forearm* and *upper arm*) the worst (86% / 85%) position. Looking at the confusion matrix still exposes an evenly distribution of the false-negatives and false-positives but certainly lower values. This indicates that the distinction of the activity-levels, more precise, the individual handling of the dimensions of the data lead to a better distinction of the device positions. Hence, the experiments show that in most of the cases it is possible to recognize the device position correctly. Thus, in general the considered positions seem not to be mixed up concerning the classification which confirms that each position provides different information for the same activity.

In summary, our position-recognition approach that makes use of a random forest classifier and distinguishes between different activity levels achieves an average performance of 89% across all positions.

TABLE VII

RESULTS OF THE BASELINE METHOD FOR ACTIVITY RECOGNITION WITHOUT POSITION INFORMATION.

Class	Precision	Recall	FP Rate	F-Measure
A_1	0.84	0.76	0.02	0.80
A_2	0.77	0.81	0.04	0.79
A_3	0.99	0.94	0.00	0.96
A_4	0.83	0.83	0.03	0.83
A_5	0.68	0.77	0.06	0.72
A_6	0.70	0.64	0.05	0.67
A_7	0.93	0.89	0.01	0.91
A_8	0.85	0.87	0.03	0.86
$avg.$	0.80	0.80	0.03	0.80

TABLE VIII

CONFUSION MATRIX FOR THE BASELINE ACTIVITY RECOGNITION METHOD WITHOUT POSITION INFORMATION.

	Predicted							
	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
A_1	4997	910	2	3	41	23	52	554
A_2	514	6758	1	36	155	108	36	784
A_3	5	2	1114	0	0	0	66	0
A_4	14	94	0	7208	512	837	63	5
A_5	20	108	0	370	6652	1231	224	12
A_6	19	117	0	1000	1798	5622	150	15
A_7	69	95	6	52	611	177	8712	22
A_8	290	741	0	3	49	23	17	7677

B. Activity Recognition

The whole idea of our work is based on the idea that knowledge about the device position improves activity recognition. We therefore also have to show that the position-aware activity recognition approach that uses the automatically detected device position outperforms the baseline approach that does not consider the device position. For this purpose, we constructed and examined the introduced position-independent activity classifier for each subject which was trained on all data of all positions. Table VII illustrates the performance of this approach and shows that the correct activity is recognized with an F-measure of 80%. However, considering the individual activities, it shows that the recognition rate is unequally distributed. Thus, *sitting* (A_6) has a significantly worse (67%) and *jumping* (A_3) a much better (96%) recognition rate. Additionally, the activities *climbing down* (A_1) and *standing* (A_5) are often confused with other activities. In this context, the related confusion matrix (see Table VIII) emphasizes that the recognized activity is often wrong if a performed activity is similar to another, i.e., *lying* (A_4), *standing* (A_5), and *sitting* (A_6) but also *climbing up* (A_1), *down* (A_2), and *walking* (A_8) are often mixed up.

In contrast, the introduced position-aware approach achieves a 4% higher F-measure. Table IX shows that in case of each

TABLE IX
RESULTS OF THE PROPOSED ACTIVITY RECOGNITION METHOD THAT USES AUTOMATICALLY DETECTED DEVICE POSITIONS.

Class	Precision	Recall	FP Rate	F-Measure
A_1	0.84	0.77	0.02	0.81
A_2	0.78	0.81	0.04	0.79
A_3	0.99	0.95	0.00	0.97
A_4	0.90	0.88	0.02	0.89
A_5	0.74	0.81	0.05	0.77
A_6	0.78	0.74	0.04	0.76
A_7	0.94	0.91	0.01	0.92
A_8	0.85	0.88	0.03	0.86
$avg.$	0.84	0.83	0.03	0.84

TABLE X
MEAN CONFUSION MATRIX: PROPOSED ACTIVITY RECOGNITION METHOD USING AUTOMATICALLY DETECTED DEVICE POSITION.

	Predicted							
	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8
A_1	5080	849	2	4	42	24	40	548
A_2	526	6820	1	26	134	87	31	768
A_3	7	5	1130	0	0	0	46	1
A_4	18	94	0	7660	324	579	57	8
A_5	19	99	0	217	7000	1020	244	15
A_6	19	112	0	582	1380	6470	141	18
A_7	70	96	11	38	535	142	8830	24
A_8	287	709	1	3	50	24	14	7720

activity, the consideration of the device localization results in a higher or equal recognition rate. Concerning the static activities, we can observe that the F-measure values increased notably. Indeed, the activities *lying* (+6%), *standing* (+5%), and *sitting* (+9%) have improved the most. In this context, the related confusion matrix (see Table X) makes clear that the problem of misclassification is not completely solved but better handled than before. For dynamic activities, the recognition rate improved slightly.

Considering the activities and positions in detail (see Table XI), it leads to the fact that there is no optimal device position. The chest, waist, thigh, and shin perform on average at best but they perform different depending on the activity. Thus, the activity *climbing stairs up* is best handled by the chest (up to 5% better) whereas the *thigh* recognizes the activity *standing* the best (up to 14% better). This confirms a statement of a previous work where they stated that the optimal sensor placement depends on the activity [6]. Further, it points out that most of the positions perform still bad regarding the static activities. This indicates that even low acceleration combined with the (predicted) device position makes it hard to distinguish between such activities. Besides, there are also activities where each position performs very well. Hence, the activities *running* ($\geq$ 91%) and *jumping* ($\geq$ 95%) are equally well recognized for all positions due to the high acceleration of the devices. These show that the acceleration

TABLE XI
RESULTS OF THE PROPOSED ACTIVITY RECOGNITION METHOD WITH KNOWN DEVICE POSITIONS.

Class	P_1	P_2	P_3	P_4	P_5	P_6	P_7
A_1	0.86	0.75	0.76	0.83	0.81	0.80	0.82
A_2	0.83	0.72	0.76	0.84	0.83	0.78	0.80
A_3	0.97	0.97	0.97	0.95	0.95	0.98	0.97
A_4	0.89	0.83	0.89	0.90	0.86	0.94	0.91
A_5	0.72	0.73	0.71	0.86	0.84	0.75	0.81
A_6	0.72	0.76	0.65	0.82	0.80	0.74	0.82
A_7	0.92	0.91	0.91	0.93	0.94	0.92	0.93
A_8	0.89	0.82	0.82	0.89	0.88	0.85	0.88
$avg.$	0.84	0.80	0.79	0.87	0.86	0.83	0.86

strength is decisive concerning the activity recognition rate and that in case of low acceleration additional information of the environment or context-related information are required.

Despite the fact that we recognized only in 89% of all cases a correct device position and compared with the position-independent approach (80%), these results indicate clearly that the consideration of the device position results in a higher activity recognition rate (84%). The results show clearly that it does not depend on the activity but on the device position if the information of the device position improves the activity recognition rate. In this context, also the individual handling of the different dimensions (e.g., device position and activity-level) leads to a better distinction of the target classes, so to a better recognition rate. Especially in the context of the static activities, these two approaches lead to a significant better recognition.

C. Comparison with other Classification Methods

In order to show the benefits of using the proposed random forest classifier, we compared its performance to the one of other common classification methods, in particular Artificial Neural Network (ANN), Decision Tree (DT), k-Nearest Neighbors (kNN), Naive Bayes (NB), and Support Vector Machine (SVM). All of these classifiers were used in previous work on activity recognition and achieved good results.

Considering the activity-level depended position recognition approach, the other classifier performed worse. Figure 5 illustrates the results and shows clearly that *Random Forest* (89%) outperforms the other classifier. In this context, NB (39%) performed the worst probably due to assumption that all features are independent. In contrast, *k-NN* (75%), *ANN* (77%), and *SVM* (78%) achieved reasonable results. We performed parameter optimization and choose a radial basis function regarding *SVM*. The *DT* (82%) performed second best but the recognition rate is much worse (-7%) than that of the *RF*. Besides, the training phase of the *RF* was one of the fastest whereas *ANN* and *SVM* took the longest.

Concerning activity recognition, we evaluated the performance of the classifier in context of position-aware activity recognition based on the recognized device positions of the random forest. Figure 6 shows that *RF* (84%) achieved the highest activity recognition rate where *NB* (61%) performed the worst. Further k-*NN* (70%) and *SVM* (71%) performed almost equal but worse than *ANN* (75%) and *DT* (76%). Besides, we also evaluated the performance of all classifier in a position-independent scenario but it expose that independent of the classifier the position-aware approach is always better.

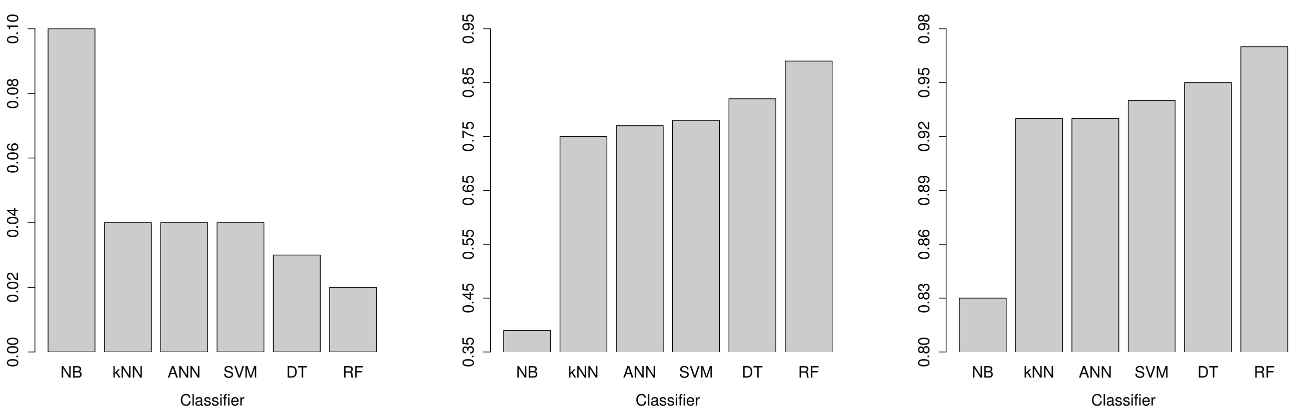
These results show that the use of the random forest classifier is not only the best classification method for determining the device position, it also outperforms all other classifiers with respect to determining the activity given a hypothesis about the position of the device.

VI. Conclusion and Future Work

In this paper, we investigated the possibility to detect the current on-body position of a wearable device in a real world scenario with a single acceleration sensor in context of several different common activities. Additionally, to evaluate the impact of the position information, we performed position-aware activity recognition where we considered the results of the on-body position detection including all mistakes. For this purpose, we created a large real world data set by recording 7 on-body positions of 15 subjects where they performed 8 different activities. Considering this data, our experiments showed that the best results were achieved with the machine learning technique random forest which detected the correct on-body device position with an F-measure of 89%.

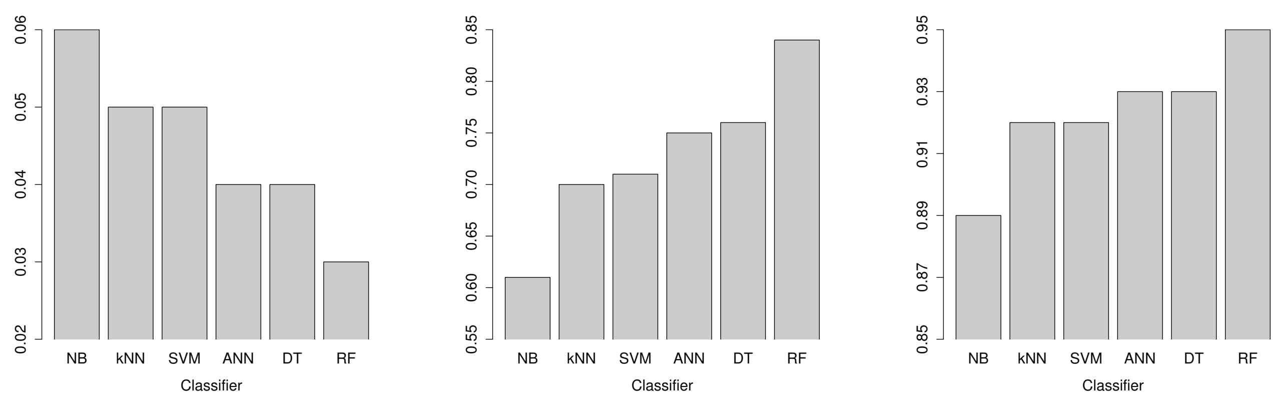
Subsequently, the activity recognition experiments were conducted. Their results show that the position-aware approach recognizes the correct device position with an F-measure of 84%. Concerning the position information, the position-aware performs 4% better then the position-unaware approach. Hence, the results provide a strong evidence for the improvement of the activity recognition rate in case that the on-body position is known.

In previous work, researchers achieved lower or equivalent recognition rates and considered less positions and activities. Thus, Coskun et al. considered the hand, trousers (thigh), and backpack and achieved a recognition rate of 85% [9]. Furthermore, Vahdatpour et al. considered the same on-body positions as we did expect the chest and considered only the activity *walking* but achieved an accuracy of 89% [2]. This indicates that the consideration of more positions and



(a) FP-Rate (b) F-measure (c) Accuracy

Fig. 5. Performance of the different classifier for position recognition in the activity-level dependent scenario.



(a) FP-Rate (b) F-measure (c) Accuracy

Fig. 6. Performance of the different classifier for position-aware activity recognition. The on-body device position was detected in a previous step by the activity-level dependent approach (Random Forest).

activities lead to a lower recognition rate as we can see in the results of our first experiments that did not distinguish between static and dynamic activities. However, as presented, due to the individual handling of different activity-level groups, our approach performs significantly better in a real world scenario where people change the orientation, device position, and activity all the time.

Considering activity recognition, Coskun et al. stated that the usefulness of the information of the device position depends on the performed activity. Further, they also stated that in general this information has a less effect on the recognition rate [9]. In contrast, Martin et al. stated that the information of the position leads to a significant improvement concerning the activity recognition [1]. In view of the fact that we considered all relevant on-body positions and several different and common activities, our result provides strong evidence concerning the positive influence of the position information.

As future work, we plan to investigate two aspects. The first aspect focuses on improving the position and activity recognition and to reduce the effort concerning the training-phase. Several people already wear two devices (e.g. smart-watch and smart-phone) which perhaps enable to compute cross-position features that identify the performed activity with higher accuracy. In this context, we want also evaluate if it is possible to reduce to effort of constructing individual classifier by considering groups of people. The second aspect, focus on deriving more precise activities. For instance, which kind of task is performed during sitting. Thus, we want to combine the recognized activity with context-related information and activity-dependent analyzing techniques to derive if a person is eating or driving a car.

REFERENCES

[1] H. Martín, A. M. Bernardos, J. Iglesias, and J. R. Casar, "Activity logging using lightweight classification techniques in mobile devices," *Personal and ubiquitous computing*, vol. 17, no. 4, pp. 675–695, 2013.

[2] A. Vahdatpour, N. Amini, and M. Sarrafzadeh, "On-body device localization for health and medical monitoring applications," in *Pervasive Computing and Communications (PerCom), 2011 IEEE International Conference on*. IEEE, 2011, pp. 37–44.

[3] L. Barkhuus and V. E. Polichar, "Empowerment through seamfulness: smart phones in everyday life," *Personal and Ubiquitous Computing*, vol. 15, no. 6, pp. 629–639, 2011.

[4] 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*. Springer, 2010, pp. 548–562.

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

[6] L. Atallah, B. Lo, R. King, and G.-Z. Yang, "Sensor positioning for activity recognition using wearable accelerometers," *Biomedical Circuits and Systems, IEEE Transactions on*, vol. 5, no. 4, pp. 320–329, 2011.

[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, pp. 74–82, 2011.

[8] 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.

[9] D. Coskun, O. D. Incel, and A. Ozgovde, "Phone position/placement detection using accelerometer: Impact on activity recognition," in *Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP), 2015 IEEE Tenth International Conference on*. IEEE, 2015, pp. 1–6.

[10] K. Kunze, P. Lukowicz, H. Junker, and G. Tröster, "Where am i: Recognizing on-body positions of wearable sensors," in *Location-and context-awareness*. Springer, 2005, pp. 264–275.

[11] E. Miluzzo, M. Papandrea, N. D. Lane, H. Lu, and A. T. Campbell, "Pocket, bag, hand, etc.-automatically detecting phone context through discovery," *Proc. PhoneSense 2010*, pp. 21–25, 2010.

[12] J.-g. Park, A. Patel, D. Curtis, S. Teller, and J. Ledlie, "Online pose classification and walking speed estimation using handheld devices," in *Proceedings of the 2012 ACM Conference on Ubiquitous Computing*. ACM, 2012, pp. 113–122.

[13] C. T. Cornelius and D. F. Kotz, "Recognizing whether sensors are on the same body," *Pervasive and Mobile Computing*, vol. 8, no. 6, pp. 822–836, 2012.

[14] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *AAAI*, vol. 5, 2005, pp. 1541–1546.

[15] O. D. Lara and M. A. Labrador, "A survey on human activity recognition using wearable sensors," *Communications Surveys & Tutorials, IEEE*, vol. 15, no. 3, pp. 1192–1209, 2013.

[16] S. J. Preece, J. Y. Goulermas, L. P. Kenney, and D. Howard, "A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data," *Biomedical Engineering, IEEE Transactions on*, vol. 56, no. 3, pp. 871–879, 2009.

[17] L. Breiman, "Random forests," *Machine learning*, vol. 45, no. 1, pp. 5–32, 2001.

[18] T. K. Ho, "The random subspace method for constructing decision forests," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 20, no. 8, pp. 832–844, 1998.

[19] J.-Y. Yang, J.-S. Wang, and Y.-P. Chen, "Using acceleration measurements for activity recognition: An effective learning algorithm for constructing neural classifiers," *Pattern recognition letters*, vol. 29, no. 16, pp. 2213–2220, 2008.