Fine-grained Accelerometer-based Smartphone Carrying States Recognition during Walking

Kaori Fujinami, Tsubasa Saeki, Yinghuan Li, Tsuyoshi Ishikawa, Takuya Jimbo, Daigo Nagase, and Koji Sato
Department of Computer and Information Sciences, Tokyo University of Agriculture and Technology
2-24-16 Naka-cho, Koganei, Tokyo 184-8588

Abstract—Due to the dependency of our daily lives on smartphones, the states of the device have impact on the quality of services offered through a smartphone. In this article, we focus on the carrying states of the device while the user is walking, in which 17 states, e.g., in the front-left trouser pocket, calling phone in the right hand, in a backpack are subjects to recognition based on supervised learning with accelerometer-derived features. A large-scale data collection from 70 persons with three walking speeds allows reliable evaluation regarding suitable features and classifiers model, the feature selection method, robustness of localization against unknown person, and effect of walking speed in training a classifier. Person-independent evaluation shows that average F-measures of 17 class classification and merged 9 class classification were 0.823 and 0.913, respectively.

Keywords—*Smartphone; on-body localization; accelerometer; machine learning; feature selection; wearable computing*

I. Introduction

Our daily lives heavily depend on smartphones that provides not only phone calling functionality, but also ubiquitous access to the Internet and replacement of objects for specific purposes, e.g., camera, pedometer, etc. as software. Various sensors are embedded into the device, which allows the a system to extract a user's and/or a device's context such as engaging activity [19], [23], [26] and a person/device location [16], [24], identity of pedestrian [28], environmental conditions around a user [8], [10], [15], [27], and so on, which contributes to provide appropriate information/services to a user based on the context.

According to a phone carrying survey, 17% of people determine the position of storing a mobile phone based on *contextual restrictions*, e.g. no pocket in the T-shirt, too large phone size for a pants pocket, comfort for an ongoing activity [4]. These factors are variable throughout the day, and thus users change their positions in a day. This suggests that a context, *on-body device position*, has great potentials in improving the usability of a smartphone and the quality of sensor-dependent services, facilitating human-human communication, the reduction of unnecessary energy consumption, etc. [7]. Note that the position is not an exact 3D coordinate, but the names of the parts of the body, clothes and items to carry the device during walking such as "inside a chest pocket", "inside a bag", and "calling (attaching to the ear)".

In this article, we propose a machine learning-based classifier and classification features to identify 17 storing positions of a smartphone on the body against a segment of data, i.e., window, obtained while a person is walking. The contribution of this article is summarized as follows:

- Classification features suitable for classifying 17 classes are specified, in which we show a subset-based feature evaluation is superior to a collection of individual "good" features.

- We show raw acceleration signal shows better classification performance than linear and vertical component of acceleration signals.

- A large scale user independent classification performance evaluation is presented, in which 70 persons provided acceleration signals of smartphone carrying during walking.

- The effect of heterogeneity of walking speed in training a classifier is evaluated, in which a training dataset with various speed can build more robust classifier than training with single, i.e., normal, speed data only.

The remaining part of this article is organized as follows. In Section II examines related work regarding on-body position sensing. Section III describes about dataset used in this study. The localization method is presented in Section IV, in which the notion of *series* is introduced, and classification features are presented. Section V shows experiments from various aspects including suitable features and classification model, the feature selection method, robustness of localization against unknown person, and effect of walking speed in training a classifier. Finally, in Section VI, we conclude the article.

II. Related Work

On-body position sensing is getting attention to researchers in machine learning and ubiquitous computing communities [7], [25], [29]. A research direction is on the type of a device which is actually realized or intended to be utilized in the future as wearable devices [18], [21], [29] or a smartphone [1], [5]–[7], [12], [14], [22], [25], [30]. The type of a device relates to the selection of target positions. In the wearable device approach, the target positions range from the head to the ankle including fine-grained discrimination such as upper arm vs. forearm and shin vs. thigh [29]. A device is usually attached firmly using a belt or a special mounting fixture. This indicates that the direction of the device might not change so irregularly within a specific activity in a frequent manner, given that small displacement might occur during activities [17]. By contrast, a smartphone terminal is usually stored into containers such as the pockets of jacket, chest and trousers pockets and a wide variety of bags, as well as in a user's hand, hanging from the neck and on a table as surveyed in [4], [30]. In this case, the degree of freedom of irregular movement in a large container, e.g., jacket pocket, handbag,

would increase. Another aspect is the modality of sensing, in which an accelerometer is dominant due to its low power operation and the availability in most commercial smartphones and wearable devices. Incel [14] shows an extensive study on acceleration-based phone localization, in which recognition features are proposed that represent the movement, rotation and orientation of devices during diverse activities of a person such as walking, sitting and biking. Fujinami proposed 63 classifier-independent features for 9 on-body phone positions including bags during walking, which selected based on as what are more predictive of classes and less correlated with each other [7]. Shi et al. [25], Alanezi et al. [1], and Incel [14] utilized a gyroscope in combination with an accelerometer. They reported that the combined approach slightly improved the accuracy [1], [14]; however, considering the power-hungry nature of a gyroscope [32], the improvement would not be the major reason for utilizing a gyroscope.

Regarding the evaluation method, n-fold cross validation is often utilized [1], [12], [18], [22], [25], [29], which utilizes $(n$-$1)/n$ of dataset for training a classifier and $1/n$ for testing the classifier; it tends to result in good recognition performance because the training dataset may contain $(n$-$1)/n$ of data from each person in theory, and hence the classifier "knows" about the subjects in advance. By contrast, Leave-One-Subject-Out (LOSO) cross validation is carried out by testing a dataset from a particular person with a classifier that is trained without a dataset from the person. So, LOSO-CV is regarded as a fairer and practical test method, and recently getting attention [1], [7], [14], [30]. To validate the generalization of a classification model, the number of subjects is important, i.e., small number of subjects fail in capturing the characteristic of the population. Incel [14] carried out a performance evaluation using LOSO-CV against an integrated dataset from 35 persons in total; however, the number of persons varies between positions (35 persons for trouser pocket, 25 for backpack, 15 for hand and 10 for messenger bag, jacket, belt and wrist), and the average number is 15.6. Fujinami utilized LOSO-CV using dataset from 20 persons, in which data from 9 positions were equally collected [7]. By contrast, we equally collect data from 17 positions on the body of 70 persons including three states in hands, i.e., swinging, talking on the phone, watching the screen from both left and right hands, as well as carrying items, i.e., bags, which is a unique aspect of our work. In existing work, the type of a bag is not clearly defined [30] or limited to a messenger bag [14], [30]. We consider that the scale of experiment in this article, i.e., 17 positions of 70 persons, is the largest one in the literature.

III. DATASET

A. Target Positions

We targeted 11 popular positions as shown in Fig. 1, among which both the left and right sides of three types of "hand", trousers front/back pockets, and jacket pockets were collected. Three type of "hand" correspond to calling, watching the screen in the portrait direction and swinging during walking. In total, 17 classes are defined and analyzed.

B. Sensor Modality

The three-axis accelerometer employed in this study is a primary sensor embedded into almost all of today's smart-
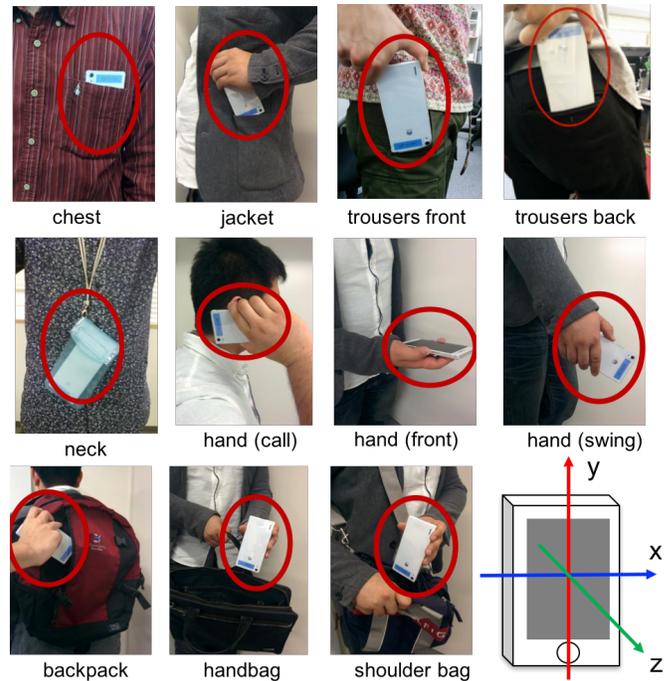

Fig. 1. Target storing positions.

phones. The signals can be used to characterize the movement patterns generated while a person is walking. Although the combination of an accelerometer and a gyroscope slightly improved the classification accuracy [1], [14], because a gyroscope is power-hungry sensor [32]. Typical waveforms of the target classes are presented in Fig. 2.

C. Data Collection

In data collection, 70 subjects (53 male and 17 female, undergraduate or graduate students at the age of 20's) were recruited with a 2,000 yen equivalent worth of remuneration. The subjects carried Huawei Ascend P7 smartphone terminals running Android 4.4 in 2 to 5 positions simultaneously, and were asked to walk about 5 min per position in the campus of our university including straight ways and corners at three walking speeds, i.e., slow, normal and fast. The speeds could be chosen by the subjects themselves; however, the order of the trial in the walking speed was kept constant such that fast, normal and slow. The subjects may get tired as the experiments proceeds. So, we consider that it is preferable to start walking with fast speed. We collected raw acceleration signals from Android API at the speed of `SENSOR_DELAY_FASTEST`. The sampled data from Android sensor system are added to an internal queue of our data collection system and polled at 50 Hz. Note that the data on a phones being carried in trousers pockets covered four orientations: downward and upward and with the display surface facing towards and away from the body.

IV. ON-BODY LOCALIZATION SYSTEM

A. Overview

The localization is carried out to recognize a class of a position from the 17 candidate positions based on the similarity
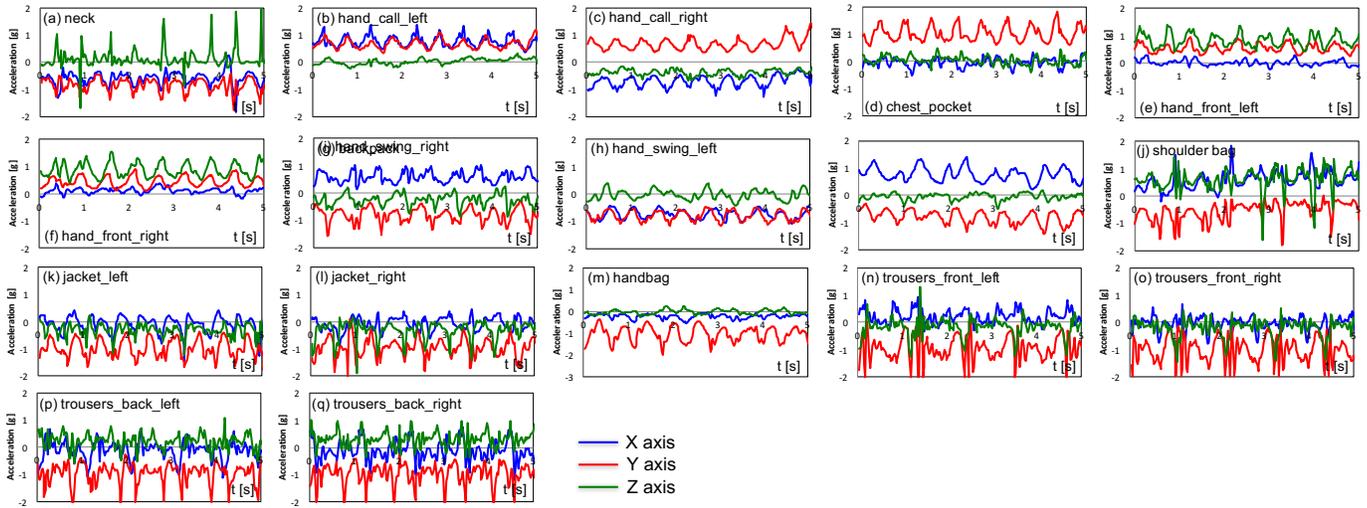
Fig. 2. Signal variations in acceleration during walking.

of patterns of acceleration signals. The recognition process is carried out window-by-window, in which a window consists of a certain number of sampled acceleration signals. In line with the principles of Vahdatpour, et al. [29], Fujinami [7], and Mannini, et al. [21], primarily recognizes the storing position of a smartphone while a person is walking. In this article, we assume that a segment representing a person is walking is already identified.

B. Signal Series and Axis

The term "series" indicates the type of basic time series data, which includes *raw* acceleration signal, *linear* acceleration component, and *vertical* acceleration component. As described in Section III-C, the raw acceleration signal is what is just obtained from accelerometer. In this section, we present the other two series. Here, the notation $a_{s,a,i}$ represents the i-th sample in a window of a-axis in the s-signal series. The coordinates in the definition of the three series is illustrated in Fig. 3.
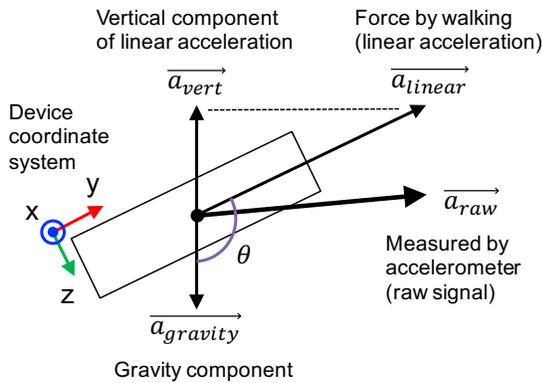

Fig. 3. The definition of three series.

1) Linear Component: Linear acceleration is obtained by removing gravity components from the measured signals. Sophisticated linear acceleration signal estimation methods have

been proposed by combining gyroscope and magnetometer [13]; however, we utilize only accelerometer for the same reason as the choice of an accelerometer as a modality of storing position recognition. We adopted the method proposed by Cho et al. [3], in which the gravity components are approximated by the mean of raw acceleration signals (1) in a window, and the linear components are obtained by subtracting the gravity component from the raw acceleration signals (2). Here, $\overline{\vec{a}_{raw}}$ is a vector of the mean raw acceleration signals of x, y and z axes in a window. Also, $\vec{a}_{linear}$ and $\vec{a}_{raw}$ indicate a vector of a sample of linear acceleration signal and raw acceleration signal in a window, respectively.

$$\vec{a}_{gravity} = \overline{\vec{a}_{raw}} \qquad (1)$$

$$\vec{a}_{linear} = \vec{a}_{raw} - \vec{a}_{gravity} \qquad (2)$$

2) Vertical Component: The vertical component is obtained by decomposing the linear component based on the component of gravity in each axis (3) [13].

$$
\begin{aligned}
\vec{a}_{vertical} &= |\vec{a}_{vertical}| \frac{\vec{a}_{gravity}}{|\vec{a}_{gravity}|} \\
&= (|\vec{a}_{linear}| cos\theta) \frac{\vec{a}_{gravity}}{|\vec{a}_{gravity}|} \qquad (3)
\end{aligned}
$$

Here, $cos\theta$ is obtained based on the definition of *inner product* ($\cdot$) as represented by (4).

$$cos\theta = \frac{\vec{a}_{linear} \cdot \vec{a}_{gravity}}{|\vec{a}_{linear}||\vec{a}_{gravity}|} \qquad (4)$$

Then, (3) is represented with the gravity and the linear components by assigning (4) as represented by (5).

$$
\begin{aligned}
\vec{a}_{vertical} &= (|\vec{a}_{linear}| \frac{\vec{a}_{linear} \cdot \vec{a}_{gravity}}{|\vec{a}_{linear}||\vec{a}_{gravity}|}) \frac{\vec{a}_{gravity}}{|\vec{a}_{gravity}|} \\
&= (\frac{\vec{a}_{linear} \cdot \vec{a}_{gravity}}{\vec{a}_{gravity} \cdot \vec{a}_{gravity}}) \vec{a}_{gravity} \qquad (5)
\end{aligned}
$$

In addition to the three axes, i.e., x, y and z, we introduce the magnitude of the three-axes signals (m) as the forth dimension for series s as shown in (6).

$$a_{s,m} = |\vec{a}_s| \qquad (6)$$

C. Recognition Features

Recognition features play very important role on determining the performance of a recognition system. In this section, we describe the definition of the candidates of features. The localization is carried out to recognize a class of a position from the 17 candidate positions based on the similarity of patterns of acceleration signals. The recognition process is carried out window-by-window, in which a window consists of a certain number of sampled acceleration signals. A feature vector is obtained per window, in which features are calculated against the three series of acceleration signals.

We take an approach of listing up candidates of features from literature [1], [7], [14], [30] and observation of waveforms (Fig. 2), and selecting relevant and non-redundant features based on a machine learning technique. We systematically calculate the candidates of features from a window of four-dimensional vector of raw acceleration signals by the combination of feature types and the axes. Totally, 326 features are obtained (72 types $\times$ 4 axes for individual axes, 5 (one for time domain and four for frequency domain) types $\times$ 6 ($=_4C_2$) pairs for correlation-based features and two types $\times$ 4 ($=_4C_3$) triples for features obtained from combination of three axes). Tables I, II, and III show the features calculated from the four axes individually, the features regarding the correlation of two axes, i.e., *correlation coefficient*, and the features representing the relationship among three axes, respectively. The feature selection is described in Section V-B.

Regarding the subscript L, M and H, the frequency spectrum is equally divided into three "frequency ranges", which correspond to 0.001-5.000 Hz, 5.001-10.000 Hz and 10.001-25.000 Hz, respectively. In addition, the subscript all indicates the entire frequency range of 0.20-25.00 Hz. Note that a feature maxSdev$_F$ is obtained in a way similar to "sliding window average"; a subwindow with 2.9 Hz range is created in an entire frequency spectrum to calculate standard deviation ($sdev$); the subwindow is slid by 0.1 Hz throughout the frequency spectrum; and the maximum $sdev$ is found. maxSdevF$_F$ is the central frequency of a particular subwindow that gives maxSdev$_F$. The size and sliding-width (0.1 Hz) of subwindow were heuristically determined. A feature calculated as the sum of squared values of frequency components is sumPower$_F$ (a.k.a "FFT energy" in [9]) [2]. The FFT entropy (entropy$_F$) is then calculated as the normalized information entropy of FFT component values of acceleration signals, which represents the distribution of frequency components in the frequency domain [2].

V. EXPERIMENT

In this section, we describe experiments from various aspects.

A. Condition

The window size is set to 256 samples, i.e., 5.12 seconds, with the sliding of every 128 samples (overlapping 50 %).

Throughout the experiment, we utilized a machine learning toolkit Weka 3.7.13 [20] running on Apple Mac Pro (3.5 GHz 6-Core Intel Xeon E5, 32 GB RAM, OS X El Capitan). Table IV summarizes average number of recognition instances, i.e., feature vectors, and standard deviation per person.

TABLE IV. AVERAGE NUMBER AND STANDARD DEVIATION (S.D.) OF RECOGNITION INSTANCES PER PERSON

Carrying state	Average (S.D.)	Carrying state	Average (S.D.)
bag_backpack	369.7 (32.6)	jacket_left	366.6 (27.7)
bag_handbag	362.5 (24.8)	jacket_right	368.5 (25.8)
bag_shoulderbag	373.6 (31.4)	neck	372.5 (30.9)
chest_pocket	367.5 (29.9)	trousers_back_left	366.7 (26.0)
hand_call_left	365.6 (29.3)	trousers_back_right	367.6 (26.2)
hand_call_right	370.2 (33.5)	trousers_front_left	365.9 (25.2)
hand_front_left	367.0 (28.6)	trousers_front_right	368.8 (26.6)
hand_front_right	366.3 (28.7)		
hand_swing_left	369.8 (30.7)	total	6253.9 (427.6)
hand_swing_right	364.9 (27.6)		

B. Feature Selection

1) Methodology: Feature selection consists of three phases: feature subset evaluation, feature subset search, and series selection. As feature subset evaluation, we utilized a correlation-based feature selection (CFS) [11], which is called `CfsSubsetEval` in Weka. CFS has a heuristic evaluation function *merit*, which can specify subset of features that are highly correlated with classes, i.e., more predictive of classes, but uncorrelated with each other, i.e., more concise. As described in Section IV-C, a large number of features were listed up, which may contain redundant features. So, we consider that the capability of CFS is suitable for this problem. Formula (7) defines the heuristic merit M_s of a feature subset S that contains k features, in which $\overline{r_{cf}}$ is the mean feature-class correlation and $\overline{r_{ff}}$ is the mean feature-feature inter-correlation. For more detail, please refer to [11]. M_S acts as a ranking on feature subsets in the search space of all possible feature subsets. Note that CFS is a classifier-independent method of feature selection.

$$M_S = \frac{k\overline{r_{cf}}}{\sqrt{k + k(k-1)\overline{r_{ff}}}} \qquad (7)$$

To find the subset of features based on the CFS evaluation, we initially attempted to utilize the forward greedy stepwise search against entire feature set (`GreedyStepwise` in Weka). The method searches the best feature subsets, which begins with no features and greedily adds features one by one. However, the computation ended up with out of memory error. So, we needed to take another approach, which finds a subset with much smaller number of features than entire dataset, i.e., 326 features, with lightweight computation at first and applies the greedy stepwise search on the subset. As a lightweight computation of searching the space of feature subsets, we utilized, `BestFirst` in Weka, a greedy hill-climbing method augmented with a backtracking facility. Setting the number of consecutive non-improving samples allowed controlling the level of backtracking done. In this experiment, the number was set to five.

The dataset of selected feature subset for each "series" is evaluated by 10 fold cross-validation (10 fold CV) to specify

TABLE I.   CLASSIFICATION FEATURES (x, y AND z AXES AND THE MAGNITUDE (m) OF THE THREE AXES)

Name	Description or Formula
$mean_T$	Average of time-series data
var_T	Variance of time-series data
max_T	Maximum value of time-series data
min_T	Minimum value of time-series data
$range_T$	Difference between max and min, i.e., $max_T - min_T$
$skew_T$	Skewness of time-series data, i.e., $\dfrac{\frac{1}{N}\sum_{i=1}^{N}(a_i - mean_T)^3}{var_T^{\frac{3}{2}}}$
$kurto_T$	Kurtosis of time-series data, i.e., $\dfrac{\frac{1}{N}\sum_{i=1}^{N}(a_i - mean_T)^4}{var_T^{\frac{4}{2}}}$
RMS_T	Root mean square of time-series data, i.e., $\sqrt{\frac{1}{N}\sum_{i=1}^{N}a_i^2}$
$absMean_T$	Absolute value of $mean_T$, i.e., $|mean_T|$
$absMax_T$	Absolute value of max_T, i.e., $|max_T|$
$absMin_T$	Absolute value of min_T, i.e., $|min_T|$
$meanAbsD_T$	Averaged absolute value of successive value's difference, i.e., $\frac{1}{N-1}\sum_{i=1}^{N-1}|a_{i+1} - a_i|$
$meanXing_T$	The number of crossing the mean value
$1^{st}Q_T$	1st quartile (1/4 smallest value) of time-series data
$3^{rd}Q_T$	3rd quartile (3/4 smallest value) of time-series data
IQR_T	Inter-quartile range of time-series data, i.e., $3^{rd}Q_T - 1^{st}Q_T$
$energy_{F,\{all|L|M|H\}}$	Sum of energy spectrum, i.e., $\sum_{i=1}^{N/2}f_i^2$
$entropy_{F,\{all|L|M|H\}}$	Frequency entropy, i.e., $-\sum_{i=1}^{N/2}p_i \times log_2 p_i$, where $p_i = f_i^2 / \sum_{i=1}^{N/2}f_i^2$
$max_{F,\{all|L|M|H\}}$	Maximum value in an entire frequency spectrum
$maxF_{F,\{all|L|M|H\}}$	Frequency that gives max_F
$meanF_{F,\{all|L|M|H\}}$	Mean frequency, i.e., $\frac{\Delta f}{N/2}\sum_{i=1}^{N/2}(f_i \times i)$
$1^{st}Q_{F,\{all|L|M|H\}}$	1st quartile (1/4 smallest) frequency spectrum
$3^{rd}Q_{F,\{all|L|M|H\}}$	3rd quartile (3/4 smallest) frequency spectrum
$IQR_{F,\{all|L|M|H\}}$	Inter-quartile range of frequency spectrum, i.e., $3^{rd}Q_F - 1^{st}Q_F$
$1^{st}QF_{F,\{all|L|M|H\}}$	Frequency that gives $1^{st}Q_F$
$3^{rd}QF_{F,\{all|L|M|H\}}$	Frequency that gives $3^{rd}Q_F$
$var_{F,\{all|L|M|H\}}$	Variance in the low-frequency range
$maxSdev_{F,\{all|L|M|H\}}$	Maximum standard deviation in subwindows in frequency spectrum
$maxSdevF_{F,\{all|L|M|H\}}$	Central frequency of subwindow that gives $maxSdev_F$
$cepDenst_{F,\{all|L|M|H\}}$	Cepstrum density, i.e., $\frac{1}{N/2}\sum_{i=1}^{N/2}|Cep_i|^2$, where Cep_i is the i-th element of cepstrum coefficient

TABLE II.   CLASSIFICATION FEATURES BASED ON CORRELATION COEFFICIENTS BETWEEN TWO AXES

Name	Description
$corr_T$	Pearson's correlation coefficient of signals from two axes in time-series data
$corr_{F,\{all|L|M|H\}}$	Correlation coefficient in an entire frequency spectrum

TABLE III.   CLASSIFICATION FEATURES OBTAINED FROM THREE AXES ($i, j, k \in \{x, y, z, m\}, i \neq j \neq k$)

Name	Description
$max3axes_T$	Max of the max of 3 out of 4 axes, i.e., $max(max_{i,T}, max_{j,T}, max_{k,T})$
$min3axes_T$	Min of the min of 3 out of 4 axes, i.e., $min(min_{i,T}, min_{j,T}, min_{k,T})$

the best feature subset for later analysis. The RandomForest classifier with 10 trees is utilized as a base classifier for the cross-validation. The classification result is evaluated by F-measure. F-measure is a harmonic mean between recall and precision. F-measure for class i is defined by (8), which is averaged over 17 classes. The recall and precision for class i are represented by (9) and (10), where $N_{correct_i}$, N_{tested_i}, and N_{judged_i} represent the number of cases correctly classified into $class_i$, the number of test cases in $class_i$, and the number of cases classified into $class_i$, respectively, while i corresponds to either one of 17 classes.

$$F\text{-}measure_i = \frac{2}{1/recall_i + 1/precision_i} \qquad (8)$$

$$recall_i = \frac{N_{correct_i}}{N_{tested_i}} \qquad (9)$$

$$precision_i = \frac{N_{correct_i}}{N_{judged_i}} \qquad (10)$$

2) Result: The `BestFirst` search filtered out about 70 features for each series as initial "meaningful" features. We then applied `GreedyStepwise` search against these features to understand the best combination in particular number of features. Fig. 4 shows the relationship between the number of features and the merit score M_s. Here, the number of features increases in the order of adding to the feature subset. As shown in the figure, the increase of M_s becomes saturated at a particular number of features. This indicates that the redundancy of features increased and/or the predictiveness of an added feature decreases after a particular number of features. The merit scores of "raw" series are larger than the other two series in almost all cases of the number of features. This suggests that "raw" series contains more predictive and less redundant than the features from the other two series and may performs best.
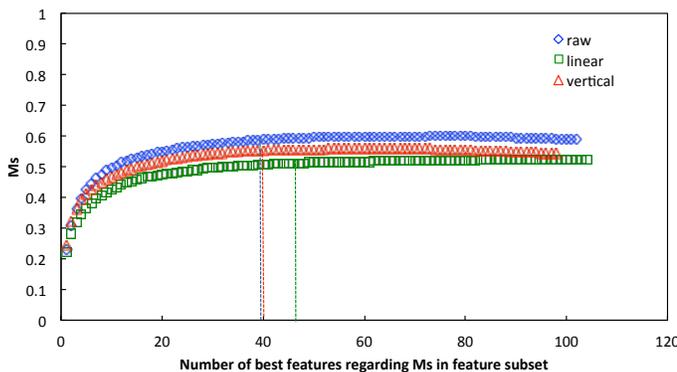

Fig. 4. Relationship between the size of feature subset and merit score of the subset (partially).

We utilized 40, 45 and 40 features for "raw", "linear" and "vertical" data series near the saturation points, respectively, in series selection. Table V summarizes average F-measures of the three series. As shown in the table, "raw" series performed the best in the three series using selected feature subsets. Table VI summarizes the selected features for "raw" data series.

TABLE V. F-MEASURE FOR EACH DATA SERIES

series	raw	linear	vertical
F-measure	0.980	0.940	0.948

TABLE VI. SELECTED FEATURES FOR "RAW" DATA SERIES. "#" REPRESENTS THE ORDER OF ADDING TO THE FEATURE SUBSET, WHILE M_s INDICATES THE MERIT SCORE OF THE SUBSET

#	Name	M_s	#	Name	M_s
1	$mean_{T,y}$	0.231	21	$var_{F,all,z}$	0.554
2	$energy_{F,M,y}$	0.309	22	$3^{rd}Q_{T,z}$	0.557
3	$1^{st}Q_{T,x}$	0.363	23	$min3axes_{T,xym}$	0.560
4	$mean_{T,z}$	0.398	24	$3^{rd}Q_{F,M,x}$	0.562
5	$3^{rd}Q_{F,M,z}$	0.426	25	$corr_{F,L,ym}$	0.564
6	$3^{rd}Q_{T,x}$	0.445	26	$corr_{T,xy}$	0.566
7	$entropy_{F,H,z}$	0.462	27	$meanXing_{T,m}$	0.568
8	$corr_{T,ym}$	0.476	28	$max_{T,x}$	0.570
9	$meanAbsD_{T,x}$	0.489	29	$kurto_{T,z}$	0.572
10	$maxSdev_{F,L,z}$	0.498	30	$skew_{T,y}$	0.574
11	$meanXing_{T,x}$	0.506	31	$corr_{T,zm}$	0.576
12	$meanF_{F,L,y}$	0.514	32	$mean_{T,x}$	0.577
13	$1^{st}Q_{T,z}$	0.519	33	$3^{rd}Q_{F,H,y}$	0.579
14	$meanXing_{T,y}$	0.525	34	$corr_{F,M,ym}$	0.581
15	$cepDenst_{F,M,x}$	0.531	35	$cepDenst_{F,M,m}$	0.582
16	$maxF_{F,all,x}$	0.535	36	$meanXing_{T,z}$	0.584
17	$skew_{T,z}$	0.539	37	$maxF_{F,all,z}$	0.585
18	$mean_{T,m}$	0.543	38	$cepDenst_{F,M,z}$	0.586
19	$corr_{T,xm}$	0.547	39	$entropy_{F,H,x}$	0.587
20	$corr_{F,L,zm}$	0.550	40	$1^{st}Q_{F,H,z}$	0.588

C. Classifier Selection

To find the best classifier, we compare popular classifiers by taking into account not only correctness of classification, but also the computational load for running on a smartphone.

1) Methodology: We carried out 10 fold CVs against Naïve Bayes, Multi-Layer Perceptron (MLP), J48 Tree and RandomForest classifiers using 40 features from raw dataset. We also measured the elapsed time to complete one fold of evaluation (test) that contains approximately 44,000 instances. Note that different numbers of trees in RandomForest were tested, i.e., 10, 50 and 100. The Support Vector Machines (SVM) has not been tested because it is parameter sensitive.

2) Result and Analysis: Fig. 5 shows the average F-measure of the classifiers and elapsed time for testing dataset per one fold. As shown in the figure, three types of Random-Forest performed better than the others. Paired t-tests showed significant difference between RandomForest with 10 trees and Naïve Bayes, MLP, and J48 with p<.05 (t(9)=224.80, t(9)=94.83, and t(9)=45.99, respectively). Also, the performance gets better as the number of threes in creased from 0.980 to 0.987. To determine the number of trees, it is important to consider the trade-off between the classification performance and processing workload. As shown in Fig. 5, the number of trees in RandomForest influences the processing speed because of the nature of the algorithm. Although paired t-tests showed that RandomForest with 10 trees was significantly lower in F-measure than RandomForest with 50 and 100 trees with p<.05 (t(9)=-44.03 and t(9)=-47,19, respectively), we took the number of trees 10 for the following experiments by taking the processing speed in processing on the smartphone. Hereinafter, RandomForest with 10 trees is utilized as a classifier in this article.

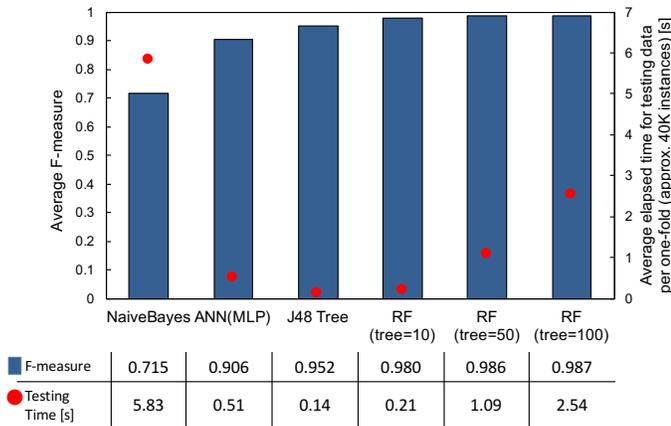|	F-measure	Testing Time [s]
NaiveBayes	0.715	5.83
ANN(MLP)	0.906	0.51
J48 Tree	0.952	0.14
RF (tree=10)	0.980	0.21
RF (tree=50)	0.986	1.09
RF (tree=100)	0.987	2.54

Fig. 5. The difference of F-measure in various classifiers. Note on the acronyms of classifiers: ANN = Artificial Neural Networks, MLP = Multi-Layer Perceptron, RF=RandomForest.

D. Feature Subset Evaluation vs. Individual Feature Collection

In Section V-B, 40 features were selected using CFS, which allows us to find subset of features that are more predictive of classes yet less correlated with each other. In this section, we evaluate the subset evaluation approach by comparing a collection of individual "good" features.

1) Methodology: The contribution of each feature is evaluated based on information gain (IG). IG is commonly used in feature selection, where the gain of information provided by a particular feature is calculated by subtracting a conditional entropy with that feature from the entropy under random guess [31]. So, the more informative feature has the higher IG.

After specifying the same number of features as those obtained by CFS method, i.e., 40 features, 10 fold CVs are performed against these two feature subsets, and F-measures are compared.

2) Result and Analysis: Table VII summarizes top-40 informative features based on IG. The features derived from x-axis show their effectiveness by appearing 7 in top-10 features. The table also shows the order of adding to the subset of 40 features obtained by CFS, which shows that individually informative feature, i.e., high IG, is not always selected in early stage (or not at all) of adding to CFS-based feature subset, i.e., low CFS value. This is natural because CFS is designed to take into account the redundancy among features and find the best combination of features, while IG is used to represent the informativeness of individual features.

Regarding the classification performance, the F-measures obtained from classifiers trained by IG-based features and CFS-based features are 0.967 and 0.980, respectively, and CFS-based feature subset is significantly contributive in classification compared to IG-based one (t(9)=83.06, p<.05). Therefore, suppose that the same number of features is utilized, we consider that the approach of feature subset evaluation was effective in building better classifier than collecting individual features with good evaluation results.

Rank	Name	IG [bit]	CFS	Rank	Name	IG [bit]	CFS
1	$mean_{T,y}$	1.21	1	20	$3^{rd}Q_{F,H,y}$	0.97	33
2	$mean_{T,x}$	1.13	32	20	$meanF_{F,M,y}$	0.97	–
3	$1^{st}Q_{T,x}$	1.10	3	20	$energy_{F,M,z}$	0.97	–
4	$3^{rd}Q_{T,x}$	1.09	6	24	$3^{rd}Q_{F,all,y}$	0.96	–
5	$RMS_{T,x}$	1.08	–	24	$3^{rd}Q_{F,M,z}$	0.96	5
6	$cepDenst_{F,all,x}$	1.07	–	24	$var_{F,M,z}$	0.96	–
7	$cepDenst_{F,H,x}$	1.05	–	24	$meanF_{F,M,z}$	0.96	–
8	$corr_{T,ym}$	1.03	8	24	$meanAbsD_{T,y}$	0.96	–
9	$energy_{F,M,y}$	1.03	2	29	$var_{F,L,y}$	0.95	–
10	$cepDenst_{F,L,x}$	1.01	–	30	$meanAbsD_{T,x}$	0.94	9
11	$cepDenst_{F,M,x}$	1.00	15	30	$maxSdev_{F,L,x}$	0.94	–
11	$max_{F,all,x}$	1.00	–	30	$maxSdev_{F,all,x}$	0.94	–
11	$max_{F,H,x}$	1.00	–	30	$maxSdev_{F,H,x}$	0.94	–
11	$max_{F,L,x}$	1.00	–	34	$IQR_{F,M,y}$	0.93	–
15	$3^{rd}Q_{F,M,y}$	0.99	–	34	$1^{st}Q_{T,z}$	0.93	13
15	$mean_{T,z}$	0.99	4	36	$meanF_{F,all,y}$	0.94	–
17	$3^{rd}Q_{T,y}$	0.98	–	36	$3^{rd}Q_{T,z}$	0.94	22
17	$meanAbsD_{T,z}$	0.98	–	36	$1^{st}Q_{F,all,z}$	0.94	–
19	$var_{F,M,y}$	0.97	–	36	$entropy_{F,H,z}$	0.94	7
20	$meanF_{F,L,y}$	0.97	12	40	$1^{st}Q_{T,y}$	0.93	–

E. Recognition against Unknown Person

As described in Section II, LOSO-CV is regarded as a fairer and more practical test method under a condition in which individual difference exists. In this section, we apply LOSO-CV to 70 subjects, which we consider the largest case in on-body smartphone localization.

1) Methodology: The dataset from one subject is treated as a test set, while the dataset from remaining 69 subjects are utilized for training a classifier. The train-and-test process is iterated 70 times.

2) Result and Analysis: Table VIII shows the detail of the classification result in the form of confusion matrix. The average F-measure in the classification of 17 classes against 70 subjects is 0.823. Although the value decreased by 0.157 from the one by 10 fold-CV, we consider that the performance is rather good given that there are 17 classes. Especially, by taking into account that no data from person for testing are included in the training data, it is surprising that left and right sides in "hand_call" and "hand_swing" were separated with very high F-measure (>0.93), in which clear differences in x-axis are observed as shown in Fig. 2. "Neck" also has high F-measure (0.932). A smartphone hanging from the neck is hit by the user's body as he/she walks forward, which causes strong impact on z-axis (Fig. 2(a)). However, as shown in Table VIII, the discriminations of left and right sides of "hand_front", "trousers_back", and "trousers_front" are often confused with each other. As shown in Fig. 2, less differences are observed in the left and right sides of theses classes than the successful cases. Also, the confusion within "bags" is slightly observed.

So, we merged the left and the right sides into one class, e.g., "hand_call_left" and "hand_call_right" → "hand_call", against "hand_call", "hand_front" and "hand_swing". Also, three types of bags are merged into one "bag" class. The result of the merging is summarized in Table IX, in which the mean F-measure is 0.913 (increased by 0.090 from original 17 classes). Furthermore, three subclasses of hand and two subclasses of trousers pockets, i.e., front and back, are merged

into single class "hand" and "trousers pocket", respectively, resulting in six class classification, which is shown in Table X. As shown in these tables, merging of multiple classes into a single class increases the performance metrics. Application designers should consider the required resolution, i.e., the level of detail of position recognition, for their target applications.

F. Effect of Various Walking Speed in Classifier Training

As described in Section III-C, we collected data with three walking speeds based on the decision of the subjects. The above experiments were carried out with dataset that contains all walking speeds. Training a classifier with single, i.e., "normal", speed is easy for the participants in data collection; however, it may sacrifice the robustness against different speed. Data collection process can be simplified if no difference exists in the robustness between classifiers modeled with heterogeneous speed and single speed. In this section, we explore the effect of walking speed in classifier training.

1) Methodology: The experiment follows LOSO-CV principle with a slight difference in walking speed between training and test datasets. More specifically, two classifiers for 17 class classification are trained using 1) dataset that contains all speeds and 2) dataset with only "normal" speed, in which training a classifier with "normal" speed is a traditional approach. Here, a dataset obtained from a test subject is excluded from the training dataset. Meanwhile, the dataset for test is either "slow", "normal", and "fast" speed. For example, a combination of "normal" speed for training with "fast" speed for testing represents a case where a person is walking faster than what the classifier knows.

In training classifiers with three walking speeds, we reduced the size of dataset to 1/3 so that it can become similar size to that of "normal" speed to avoid the bias of the number of training instances. Actually, three sets of 1/3 sampled dataset are applied, and F-measures are averaged. Regarding the classification features for training with "normal" speed, we selected dedicated ones in the same way as with all speeds (Section V-B) because we consider that suitable set of features can be different from each other due to the variation of walking speed in "all speed" case.

2) Result and Analysis: Table XI summarizes average F-measures in different combinations of walking speed in training and test datasets. Paired t-tests regarding the heterogeneity in training datasets showed that using three walking speeds performed better classification than using single, i.e., "normal", speed ($p<.05$) in all cases of walking speeds in test datasets ($t(69)=-2.34$, $t(69)=2.64$, and $t(69)=5.30$ for "slow", "normal", and "fast", respectively). The result shows that, in building a classifier, heterogeneity of walking speed is important for robust classifier.

TABLE XI. AVERAGE F-MEASURES IN DIFFERENT COMBINATIONS OF TRAINING AND TEST DATASETS

Trained\Tested	Slow	Normal	Fast
All	0.799	0.814	0.787
Normal	0.786	0.801	0.758

VI. CONCLUSION

In this article, we proposed a machine learning-based classifier and classification features to identify 17 states of a smartphone while the user is walking. A large-scale data collection from 70 persons were carried out with three different walking speeds to evaluate the effect of heterogeneity of walking speed in training a classifier. The following results were obtained:

- In feature calculation, we introduced three series of acceleration signals, *raw*, *linear*, and *vertical* components, in which the *raw* acceleration series showed the highest classification performance in the three series.

- 40 features in the *raw* series were selected from 326 candidates features based on correlation-based feature subset evaluation. The comparison with a subset by collecting individually informative features based on information gain showed that the subset evaluation method was superior to the collection-based method with the same number of features.

- Person-independent evaluation (LOSO-CV) showed that an average F-measure of 17 class classification was 0.823, while 9 class classification by merging left and right sides into one class showed an average F-measure of 0.913.

- Comparison of the heterogeneity of walking speeds in training dataset showed that the classifier built from various walking speed allowed us to realize more robust classifier than using a classifier with a single walking speed (normal speed).

We consider that the F-measure of 0.824 for 17 class classification has still room for improvement by using suitable classifier to address "classifier compatibility" issue as suggested in [7]. In addition, the classification in the experiment was carried out against a window, which means that decisions of successive windows may differ due to mis-classification. For practical recognition, we will investigate temporal smoothing techniques to smoothen such "discontinuity" of recognition results. We have already developed a mechanism to identify a segment of walking, to which these future investigation will be integrated.

REFERENCES

[1] K. Alanezi and S. Mishra, "Design, implementation and evaluation of a smartphone position discovery service for accurate context sensing," *Computers & Electrical Engineering*, vol. 44, pp. 307–323, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S004579061500021X

[2] L. Bao and S. S. Intille, "Activity recognition from user-annotated acceleration data," in *Proceedings of the 2nd International Conference on Pervasive Computing (Pervasive 2004)*, vol. LNCS 3001. Springer-Verlag, 2004, pp. 1–17.

[3] S.-J. Cho, E. Choi, W.-C. Bang, J. Yang, J. Sohn, D. Y. Kim, Y.-B. Lee, and S. Kim, "Two-stage Recognition of Raw Acceleration Signals for 3-D Gesture-Understanding Cell Phones," in *Proceedings of the Tenth International Workshop on Frontiers in Handwriting Recognition*, 2006.

TABLE VIII.   Confusion Matrix of LOSO-CV against 70 Subjects, in which the Row and the Column Indicate the Labeled and the Predicted Classes, Respectively. The Number of Instances in Each Class is Normalized so that the Sum of Each Row becomes 100

Labeled\Predicted	a.	b.	c.	d.	e.	f.	g.	h.	i.	j.	k.	l.	m.	n.	o.	p.	q.
a. bag_backpack	78	0	6	4	0	2	0	0	0	0	4	3	0	0	1	0	0
b. bag_handbag	0	94	0	1	1	1	0	1	0	0	1	1	0	0	0	0	0
c. bag_shoulderbag	5	0	87	1	0	0	1	0	0	0	2	2	2	0	0	0	0
d. chest_pocket	2	2	0	89	1	1	0	0	0	0	1	2	0	0	0	0	0
e. hand_call_left	0	0	0	1	97	0	0	0	0	0	0	0	0	0	0	0	0
f. hand_call_right	0	0	0	1	0	96	0	2	0	0	0	0	0	0	0	0	0
g. hand_front_left	0	0	1	2	0	0	86	10	1	0	0	0	0	0	0	0	0
h. hand_front_right	0	0	0	1	0	3	13	77	0	6	0	0	0	0	0	0	0
i. hand_swing_left	0	0	0	0	2	0	0	0	96	0	0	0	0	0	0	0	0
j. hand_swing_right	0	0	0	0	0	1	0	3	1	93	0	0	0	0	0	0	0
k. jacket_left	3	1	1	3	1	0	0	0	1	0	72	11	2	0	0	2	2
l. jacket_right	3	1	1	4	0	0	0	0	0	1	8	71	3	1	1	2	2
m. neck	0	0	1	0	0	0	0	0	0	0	1	1	96	0	0	0	0
n. trousers_back_left	1	0	0	2	0	0	0	0	0	0	1	2	1	66	23	2	2
o. trousers_back_right	1	0	1	1	0	0	0	0	0	0	1	2	0	25	67	1	1
p. trousers_front_left	0	0	0	1	0	0	0	0	0	0	2	2	1	2	0	75	17
q. trousers_front_right	0	0	0	1	0	0	0	0	0	0	1	4	0	2	2	27	62
Recall (average: 0.825)	0.784	0.937	0.869	0.887	0.975	0.963	0.856	0.767	0.958	0.931	0.718	0.715	0.959	0.660	0.672	0.749	0.622
Precision (average: 0.824)	0.815	0.954	0.875	0.801	0.937	0.909	0.842	0.832	0.959	0.927	0.759	0.697	0.907	0.680	0.710	0.682	0.715
F-measure (average: 0.823)	0.799	0.945	0.872	0.842	0.956	0.935	0.849	0.798	0.959	0.929	0.738	0.706	0.932	0.670	0.690	0.714	0.666

TABLE IX.   Performance of LOSO-CV for Merged 9 Classes against 70 Subjects

Labeled\Predicted	bag	chest_pocket	hand_call	hand_front	hand_swing	jacket_pocket	neck	trousers_back	trousers_front	average
Recall	0.900	0.887	0.969	0.926	0.949	0.815	0.959	0.903	0.908	0.913
Precision	0.862	0.897	0.920	0.956	0.949	0.821	0.953	0.940	0.921	0.913
F-measure	0.881	0.892	0.944	0.941	0.949	0.818	0.956	0.921	0.914	0.913

TABLE X.   Performance of LOSO-CV for Merged 6 Classes against 70 Subjects

Labeled\Predicted	bag	chest_pocket	hand	jacket_pocket	neck	trousers_pocket	average
Recall	0.900	0.887	0.982	0.815	0.959	0.935	0.913
Precision	0.884	0.922	0.928	0.857	0.959	0.925	0.912
F-measure	0.892	0.904	0.954	0.835	0.959	0.930	0.912

[4] Y. Cui, J. Chipchase, and F. Ichikawa, "A Cross Culture Study on Phone Carrying and Physical Personalization," in *Proceedings of HCI International 2007*, 2007, pp. 483–492.

[5] I. Diaconita, A. Reinhardt, D. Christin, and C. Rensing, "Inferring Smartphone Positions Based on Collecting the Environment's Response to Vibration Motor Actuation," in *Proceedings of the 11th ACM Symposium on QoS and Security for Wireless and Mobile Networks*, ser. Q2SWinet '15. New York, NY, USA: ACM, 2015, pp. 99–106. [Online]. Available: http://doi.acm.org/10.1145/2815317.2815342

[6] I. Diaconita, A. Reinhardt, F. Englert, D. Christin, and R. Steinmetz, "Do you hear what {I} hear? Using acoustic probing to detect smartphone locations," in *2014 {IEEE} International Conference on Pervasive Computing and Communication Workshops, PerCom 2014 Workshops, Budapest, Hungary, March 24-28, 2014*, 2014, pp. 1–9. [Online]. Available: http://dx.doi.org/10.1109/PerComW.2014.6815157

[7] K. Fujinami, "On-Body Smartphone Localization with an Accelerometer," *Information*, vol. 7, no. 2, p. 21, 2016. [Online]. Available: http://www.mdpi.com/2078-2489/7/2/21

[8] K. Fujinami, "Smartphone-based Environmental Sensing Using Device Location as Metadata," *International Journal on Smart Sensing and Intelligent Systems*, vol. 9, no. 4, 2016. [Online]. Available: http://s2is.org/Issues/v9/n4/papers/abstract31.pdf

[9] K. Fujinami and S. Kouchi, "Recognizing a Mobile Phone's Storing Position as a Context of a Device and a User," in *Mobile and Ubiquitous Systems: Computing, Networking, and Services*, ser. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, K. Zheng, M. Li, and H. Jiang, Eds. Springer Berlin Heidelberg, 2013, vol. 120, pp. 76–88. [Online]. Available: http://dx.doi.org/10.1007/978-3-642-40238-8{_}7

[10] J. Goldman, K. Shilton, J. Burke, D. Estrin, M. Hansen, N. Ramanathan, S. Reddy, V. Samanta, M. Srivastava, and R. West, "Participatory Sensing: A citizen-powered approach to illuminating the patterns that shape our world," *Foresight and Governance Project, White Paper*, 2009.

[11] M. A. Hall, "Correlation-based Feature Selection for Machine Learning," Ph.D. dissertation, The University of Waikato, 1999.

[12] C. Harrison and S. E. Hudson, "Lightweight material detection for placement-aware mobile computing," in *UIST '08: Proceedings of the 21st annual ACM symposium on User interface software and technology*. New York, NY, USA: ACM, 2008, pp. 279–282.

[13] S. Hemminki, P. Nurmi, and S. Tarkoma, "Gravity and Linear Acceleration Estimation on Mobile Devices," in *Proceedings of the 11th International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services*, ser. MOBIQUITOUS '14. ICST, Brussels, Belgium, Belgium: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering), 2014, pp. 50–59. [Online]. Available: http://dx.doi.org/10.4108/icst.mobiquitous. 2014.258034

[14] O. D. Incel, "Analysis of Movement, Orientation and Rotation-Based Sensing for Phone Placement Recognition," *Sensors*, vol. 15, no. 10, p. 25474, 2015. [Online]. Available: http://www.mdpi.com/1424-8220/ 15/10/25474

[15] T. Ishikawa and K. Fujinami, "Smartphone-Based Pedestrian's Avoidance Behavior Recognition towards Opportunistic Road Anomaly Detection," *ISPRS International Journal of Geo-Information*, vol. 5, no. 10, p. 182, oct 2016. [Online]. Available: http://www.mdpi.com/ 2220-9964/5/10/182

[16] T. Jimbo and K. Fujinami, "Detecting mischoice of public transportation route based on smartphone and GIS," in *UbiComp and ISWC 2015 - Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and the Proceedings of the 2015*

ACM International Symposium on Wearable Computers, 2015, pp. 165–168.

[17] K. Kunze and P. Lukowicz, "Dealing with Sensor Displacement in Motion-based Onbody Activity Recognition Systems," in *Proceedings of the 10th International Conference on Ubiquitous Computing*, ser. UbiComp '08. New York, NY, USA: ACM, 2008, pp. 20–29. [Online]. Available: http://doi.acm.org/10.1145/1409635.1409639

[18] K. Kunze, P. Lukowicz, H. Junker, and G. Tröster, "Where am I: Recognizing On-body Positions of Wearable Sensors," in *Proceedings of International Workshop on Location- and Context-Awareness (LoCA 2005), LNCS 3479*, 2005, pp. 264–275.

[19] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, "Activity recognition using cell phone accelerometers," *SIGKDD Explor. Newsl.*, vol. 12, no. 2, pp. 74–82, 2011. [Online]. Available: http://doi.acm.org/10.1145/1964897.1964918

[20] Machine Learning Group at University of Waikato, "Weka 3 - Data Mining with Open Source Machine Learning Software in Java," http://www.cs.waikato.ac.nz/ml/weka/. [Online]. Available: http://www.cs.waikato.ac.nz/ml/weka/

[21] A. Mannini, A. M. Sabatini, and S. S. Intille, "Accelerometry-based recognition of the placement sites of a wearable sensor," *Pervasive and Mobile Computing*, vol. 21, pp. 62–74, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1574119215001108

[22] E. Miluzzo, M. Papandrea, N. D. Lane, H. Lu, and A. T. Campbell, "Pocket, Bag, Hand, etc.-Automatically Detecting Phone Context through Discovery," in *Proc. of the First International Workshop on Sensing for App Phones (PhoneSense'10)*, 2010.

[23] S. Pirttikangas, K. Fujinami, and T. Nakajima, "Feature Selection and Activity Recognition from Wearable Sensors," in *2006 International Symposium on Ubiquitous Computing Systems (UCS2006)*, 2006, pp. 516–527.

[24] A. Rai, K. K. Chintalapudi, V. N. Padmanabhan, and R. Sen, "Zee: Zero-effort Crowdsourcing for Indoor Localization," in *Proceedings of the 18th Annual International Conference on Mobile Computing and Networking*, ser. Mobicom '12. New York, NY, USA: ACM, 2012,

pp. 293–304. [Online]. Available: http://doi.acm.org/10.1145/2348543.2348580

[25] Y. Shi, Y. Shi, and J. Liu, "A rotation based method for detecting on-body positions of mobile devices," in *Proceedings of the 13th international conference on Ubiquitous computing*, ser. UbiComp '11. ACM, 2011, pp. 559–560.

[26] M. Shoaib, S. Bosch, O. D. Incel, H. Scholten, and P. J. M. Havinga, "Fusion of Smartphone Motion Sensors for Physical Activity Recognition," *Sensors*, vol. 14, no. 6, p. 10146, 2014. [Online]. Available: http://www.mdpi.com/1424-8220/14/6/10146

[27] M. Stevens and E. D'Hondt, "Crowdsourcing of Pollution Data using Smartphones," in *1st Ubiquitous Crowdsourcing Workshop at UbiComp*, 2010.

[28] D. Sugimori, T. Iwamoto, and M. Matsumoto, "A Study about Identification of Pedestrian by Using 3-Axis Accelerometer," in *Embedded and Real-Time Computing Systems and Applications (RTCSA), 2011 IEEE 17th International Conference on*, 2011, pp. 134–137.

[29] A. Vahdatpour, N. Amini, and M. Sarrafzadeh, "On-body device localization for health and medical monitoring applications," in *Proceedings of the 2011 IEEE International Conference on Pervasive Computing and Communications*, ser. PERCOM '11. IEEE Computer Society, 2011, pp. 37–44.

[30] J. Wiese, T. S. Saponas, and A. J. B. Brush, "Phoneprioception: Enabling Mobile Phones to Infer Where They Are Kept," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ser. CHI '13. New York, NY, USA: ACM, 2013, pp. 2157–2166. [Online]. Available: http://doi.acm.org/10.1145/2470654.2481296

[31] I. H. Witten, E. Frank, and M. A. Hall, *Data Mining: Practical Machine Learning Tools and Techniques*, third edit ed. San Francisco, CA, USA: Morgan Kaufmann Publishers, 2011.

[32] L. Zhang, P. H. Pathak, M. Wu, Y. Zhao, and P. Mohapatra, "AccelWord: Energy Efficient Hotword Detection Through Accelerometer," in *Proceedings of the 13th Annual International Conference on Mobile Systems, Applications, and Services*, ser. MobiSys '15. New York, NY, USA: ACM, 2015, pp. 301–315. [Online]. Available: http://doi.acm.org/10.1145/2742647.2742658