A Comparative Study on Human Activity Recognition Using Inertial Sensors in a Smartphone

Aiguo Wang, Guilin Chen, Jing Yang, Shenghui Zhao, and Chih-Yung Chang, *Member, IEEE*

***Abstract*— Activity recognition plays an essential role in bridging the gap between the low-level sensor data and the high-level applications in ambient-assisted living systems. With the aim to obtain satisfactory recognition rate and adapt to various application scenarios, a variety of sensors have been exploited, among which, smartphone-embedded inertial sensors are widely applied due to its convenience, low cost, and intrusiveness. In this paper, we explore the power of triaxial accelerometer and gyroscope built-in a smartphone in recognizing human physical activities in situations, where they are used simultaneously or separately. A novel feature selection approach is then proposed in order to select a subset of discriminant features, construct an online activity recognizer with better generalization ability, and reduce the smartphone power consumption. Experimental results on a publicly available data set show that the fusion of both accelerometer and gyroscope data contributes to obtain better recognition performance than that of using single source data, and that the proposed feature selector outperforms three other comparative approaches in terms of four performance measures. In addition, great improvement in time performance can be achieved with an effective feature selector, indicating the way of power saving and its applicability to real-world activity recognition.**

***Index Terms*— Activity recognition, smartphone, accelerometer, gyroscope, feature selection.**

I. Introduction

ALONG with the development of sensor and sensor network technology, a variety of advanced applications are emerging in a large number of fields, ranging from pervasive computing, security and surveillance to vehicle network and healthcare. Also, sensor technology can potentially facilitate many applications in a home setting such as activity reminder, fall detection, rehabilitation instruction, and wellness evaluation in assisted living systems [1]–[3], where activity recognition plays an increasingly important role in bridging the gap between the low-level simple sensor data and high-level meaningful applications, especially in the healthcare related fields [4], [5]. In the application of wellness evaluation, for example, considering the fact that activities of daily living are closely related to the health state of an individual [2], then we can deduce his/her health conditions by capturing the activities performed by them in terms of location, start time, frequency and duration. If we examine their activities for a long period of time, we can detect the health state changes that are indicated by changes in the activities, such as the Alzheimer's disease characterized by long-time sitting or lying down, and sleep disorders with fragmented and shorter sleeping stages. Activity recognition using various sensors, however, is not a trivial due to the intrinsic nature of human activity. Different people may perform the same activity in a different way, and even the same person may perform an activity in a different manner at different time, and there are situations where concurrent and interleaved activities occur that makes it difficult to obtain robust activity recognizer [6]. Consequently, activity recognition is a challenging but active research area that has been drawing the attentions of researchers from the community of data mining, pervasive computing, medical and healthcare [2], [6], [7].

In activity recognition, different types of sensing technologies have been explored to improve the recognition rate and adapt to a variety of application scenarios. Generally, they can be broadly grouped into three categories: vision based approach, environment interactive sensor based approach and wearable sensor based approach [6], [8]. Vision-based approaches mainly employ a camera or video to monitor and recognize different activities [9]. Although they can provide better recognition rate, the use of camera or video is not practical in many indoor environments particularly when the privacy issue is considered. In addition, vision-based approaches often suffer from illumination variations, ambient occlusion and background change that greatly limits their actual use [9], [10].

For environment interactive sensor-based approaches, they recognizes human physical activities by capturing the interaction between the subjects and objects under the assumption that there exists the underlying relation between objects and activities [11]. For example, if a sensor embedded in a chair is triggered, we infer that an individual is sitting on it; and if we monitor that sensors that are placed on a bed are fired for a period of time, someone is probably sleeping. Such schemes could recognize human daily living activities such

Manuscript received January 29, 2016; revised March 16, 2016; accepted March 21, 2016. Date of publication March 23, 2016; date of current version April 26, 2016. This work was supported by the National Natural Science Foundation of China under Grant 61305064 and Grant 61472057. The associate editor coordinating the review of this paper and approving it for publication was Prof. Zeynep Celik-Butler. *(Corresponding author: Guilin Chen.)*

A. Wang is with Chuzhou University, Chuzhou 239000, China, and also with the Hefei University of Technology, Hefei 230009, China (e-mail: wangaiguo2546@163.com).

G. Chen and S. Zhao are with Chuzhou University, Chuzhou 239000, China (e-mail: glchen@chzu.edu.cn; zsh@chzu.edu.cn).

J. Yang is with the Hefei University of Technology, Hefei 230009, China (e-mail: jsjyj0801@163.com).

C.-Y. Chang is with Tamkang University, Taipei 25137, Taiwan (e-mail: cychang@mail.tku.edu.tw).

Digital Object Identifier 10.1109/JSEN.2016.2545708

as washing, eating, bashing, sleeping and watching, but they are generally costly and often limited to indoor scenarios [6]. Also, a complex process is required to set up and maintain the system in order to make the whole system work normally. Other issues include how to effectively and efficiently deploy these sensors in appropriate places and appliances without causing much inconvenience to users and how to correctly map sensor readings to corresponding activity label in the process of prediction model construction [11], [12].

Unlike environment interactive sensor based approach, the miniature and flexibility of sensors makes it possible for an individual to wear or carry mobile devices that are embedded with various sensing units [13]. This kind of devices are suitable for both indoor and outdoor settings, and can be worn on many parts of the human body that ranges from the upper (e.g. arm and wrist) to the lower (e.g. leg and ankle) [14], [15]. Besides that, an individual can take more than one mobile devices at the same time [16]. Correspondingly, a wealth of researches have been conducted to explore the potential of wearable devices for activity recognition in a pervasive and ubiquitous way. Among these wearable devices, smartphones have the benefit that it releases the users from wearing additional sensing components, therefore, activity recognition based on smartphones is promising and attractive due to its non-intrusiveness and high acceptance and adherence in daily life [17]. Also, modern smartphones have powerful computing and communication capabilities that enables us to process computation tasks locally and interact with remote server efficiently [18]. In addition, smartphones often contain several built-in sensors such as accelerometer, gyroscope, GPS and camera, which we can use to gather device-related information and its context [19]. Such a configuration and the increasing adoption of smartphones allow us to use smartphones to collect and analyze activity-related raw data and provide an alternative and economic way for activity recognition.

With the aim to have a comprehensive understanding of activity recognition with inertial sensors built-in a smartphone, in this study, we plan to investigate the power of accelerometer and gyroscope in activity recognition when they are used simultaneously or separately, and further explore the usefulness of feature selection methods in improving recognition performance. In particular, the main contributions of this paper are summarized as follows. (1) *Activity recognition framework*: we present an activity recognition framework that works in an offline training and online prediction scheme. This helps optimize the parameter setting and feature selection of the activity recognizer effectively and make online prediction efficiently. In addition, it is a general framework that can be applied to various classification and regression tasks. (2) *Feature selection method*: we explore the use of feature selection in activity recognition, and propose a novel feature selection in order to improve the recognition rate and reduce the time cost in prediction and associated power consumption of smartphone. We also compare the proposed method with three other state-of-the-art feature selectors. (3) *Choice of inertial sensors*: we evaluate and compare the power of the triaxial accelerometer and triaxial gyroscope, and show their synthetic effects in recognizing human physical activities. Besides, we perform feature selection on the data collected from different sensors to comprehensively analyze the power of inertial sensors. (4) *Experimental evaluation*: we conduct extensive experiments and obtain several valuable results that can help researchers make better decisions in utilizing wearable and mobile devices for activity recognition. For example, we observe that both the gyroscope and accelerometer perform well in discriminating static activities from dynamic activities, and that the fusion of gyroscope and accelerometer contributes to the improvement in recognition rate.

The rest of this paper is organized in the following way. Section III presents the related work in activity recognition and feature selection. Section III gives the proposed activity recognition framework, details the data preprocessing and feature extraction procedures, as well as illustrates the proposed feature selection method. In Section IV, we first describe the experimental setup and performance measures, and then present the experimental results from several aspects. The last section concludes this study with a brief summary and points out future research work.

II. Related Work

A large number of researchers have conducted considerable work in exploring different sensing technologies and proposed a number of methods to model and recognize human activities [13], [20]. In actual use, there are a variety of sensors that are available for use, among which, wearable devices that are embedded with accelerometer, gyroscope, GPS and other sensors have proven to be effective and are gaining popularity. Particularly, researchers have constructed various activity recognition systems that utilized accelerometers to infer body-position, due to the fact that an accelerometer can offer us acceleration and velocity information largely associated with different human physical activities [21]–[24]. For instance, Bao and Intille used five small biaxial accelerometers that were worn simultaneously on different parts of the body (four limb positions plus the right hip) to collect sensor data when an individual performed daily tasks [21]. They then extracted both time-domain and frequency-domain features and constructed a classification model to recognize twenty activities [21]. In experiments, they collected experimental data from twenty volunteers and compared the recognition rate of several different classifiers. Experimental results showed that decision tree classifier can obtain the best performance with an accuracy of 84.0%. Tapia et al. proposed to implement a real-time system that can recognize physical activities as well as their intensities using a heart rate monitor and five triaxial accelerometers placed on right arm, right leg and the waist [7]. They applied the system to recognize thirty physical gymnasium activities and obtained a 94.6% subject-dependent recognition rate and 56.3% subject-independent accuracy, and can obtain 80.6% accuracy without differentiating the activity intensities. Although the use of multiple accelerometers that are simultaneously attached to a human body leads to a better recognition rate [25], [26], this kind of approach definitely causes inconvenience to users because of the complex system configuration. Hence, this may prevent users from getting involved in a long-term setting, and motivates the use of a single sensor [27].

Rather than use multiple sensors, for example, Ravi et al. carried out a study to explore the possibility of activity recognition with a single triaxial accelerometer [14]. In the phase of data collection, they used an accelerometer mounted onto the pelvic region of an individual to collect raw accelerometer data about eight activities, including standing, walking, running, upstairs, and downstairs, sitting, vacuuming, and brushing teeth. They then proposed to construct a meta-level classifier that consisted of decision table, decision tree, k-nearest-neighbor, support vector machine, and naïve bayes. To evaluate the effectiveness, extensive experiments were conducted under four different dataset settings, and their results showed that in comparison with base-level classifier, meta-level classifier with plurality voting generally tends to achieve better performance. Khan et al. also investigated the power of a single triaxial accelerometer in human activity recognition [23]. In their study, the accelerometer was required to attach to the chest of an individual in a particular orientation. They then proposed a hierarchical prediction model to classify static, dynamic and transitional activity, applied the model to recognize fifteen activities in a natural setting and obtained satisfactory recognition accuracy. In comparison with the use of multiple accelerometers, activity recognition with a single accelerometer may be unable to discriminant similar activities, such as downstairs and upstairs [28].

In addition, there are studies that explored the fusion of different kinds of sensors for human activity recognition [29], [30]. For example, Lee and Masc built a system that consisted of a biaxial accelerometer, a gyroscope and a digital compass to infer a user's location and classify siting, walking and standing activities [31]. Yun et al. proposed to use a triaxial accelerometer, a triaxial angular rate sensor and a magnetometer to collect foot motion related data for estimating foot kinematics [32]. Their system was able to extract information about foot orientation, velocity, acceleration, position and gait phase, and obtain relatively low estimation error. Huynh et al. explored the use of an accelerometer and a gyroscope to construct a wireless and wearable fall detection system and proposed a critical threshold based activity recognition algorithm to differentiate falls from non-falls [30]. In their study, the sensing unit was attached to the chest center to collect motion data under different daily activities such as standing, walking, sitting, running and four fall schemes (forward, backward, right and left sideway).

Nowadays, with increasing power of smartphones in computation and communication, they are often embedded with GPS, accelerometers, gyroscopes, and a digital compass, and are able to process computation tasks locally and interact with remote server efficiently. Particularly, among the smartphone sensors, GPS that are commonly used in location-based services, can locate one's current position and track the trace. The triaxial accelerometer measures the proper acceleration and can be used to obtain the acceleration of three orthogonal axes: forward acceleration in y-axis, horizontal movement acceleration in x-axis, and vertical acceleration in z-axis. It contains information to recognize different activities, for example, discriminating walking upstairs from walking downstairs with the acceleration values of z-axis in the two different cases. For the triaxial gyroscope, it provides the angular acceleration information of the smart phone from three different views, and can be used to estimate the orientation and rotation of the movement with the help of the pitch, roll and yaw attitude angles. Consequently, smartphones provide us an alternative way for activity recognition [33]. Besides, the use of smartphones can release users from wearing additional sensing components for data collection and procession, and is feasible for long-term activity monitoring due to its relatively low intrusiveness and high adherence in daily life [34]. For example, Dernbach et al. demonstrated the possibility of using the inertial sensor data collected from android-based smart phones to recognize simple activities such as biking, climbing, sitting, walking, running and standing, as well as complex activities such as cleaning, cooking, washing and watering [17]. Chiang et al. used an android-based smartphone embedded with an accelerometer and GPS to record activity patterns for self-wellness management. In the study, they considered nine activities and evaluated the system with four different classifiers [35]. Anjum and Ilyas built a smartphone application to detect seven human physical activities and further to estimate calories consumption [19], and they collected 510 activity traces and evaluated the effectiveness with the k-nearest-neighbor, naïve bayes, and support vector machine and decision tree classifiers. For smartphone-based activity recognition, besides pursing high accuracy, energy saving and real-time response are two important and critical factors in improving user experience [36].

In terms of accuracy and energy consumption, the quality and the number of used features matters. There are a plenty of researches that aim to extract important features in both time-domain and frequency-domain from the raw senor signal with prior knowledge [23], [37], [38], while few studies, to the best of our knowledge, optimize the selection of features in a data-driven way. In data analysis, a subset of best individual feature may not be a best subset of features because of the redundancy among features, and redundant features may deteriorate the performance of k-nearest-neighbor classifier [39]. Feature selection aims to reduce the feature dimensionality by selecting a small subset of discriminant features. It helps improve the classification accuracy and the generalization ability of a classifier, lower the time cost and lengthen the battery usage time in activity recognition [40]. Accordingly, how to select powerful features remains a research topic, and an effective feature selection algorithm is desirable in actual use.

III. Activity Recognition With Smartphone Sensors

A. Proposed Activity Recognition Model

There are a variety of classification models available for use. According to the availability of data labels, we can broadly group them into three categories: supervised learning models working with labelled data, unsupervised learning models without data labels, and semi-supervised learning techniques

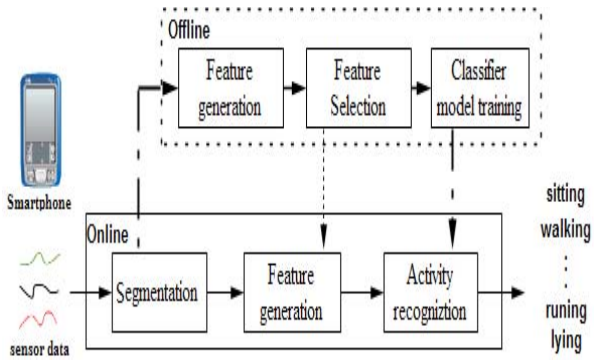


Fig. 1. Proposed smartphone-based activity recognition framework.

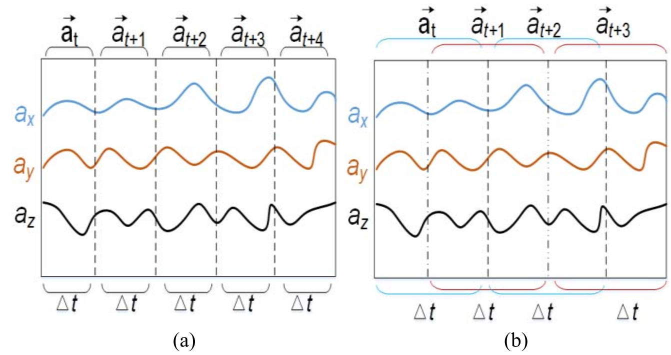


Fig. 2. Two types of sliding window techniques. (a) Non-overlapping. (b) Overlapping.

that can utilize both labelled and non-labelled data [33]. The widely applied approach to activity recognition is to apply the supervised learning with an explicit training phase in order to pursue high accuracy. In activity recognition, supervised learning typically consists of three stages. In the first phase, a stream of sensor data is divided into segments where a sliding window technique is used. Specifically, a window with a fixed/varied length or fixed/varied number of sensor events is shifted along the stream with (non-) overlapping between adjacent segments [6], [41]. The next step is to extract features from the segments and represent the raw sensor data with feature vectors, and then construct and train the classification model over these extracted features. The last task is to use the obtained classifier to associate a stream of sensor data with a predefined activity. In the setting of smartphone-based activity recognition, considering the limited processing power and battery life [36], we propose the following scheme: training the classification model offline and recognizing activities online. Fig. 1 presents the proposed activity recognition framework. In this study, we focus our attention on the use of the built-in triaxial accelerometer and gyroscope. Therefore, there are two types of data generated by a smartphone.

In the offline training phase, we first extract features from the collected sensor data. In this stage, various time-domain and frequency-domain features such as mean, standard deviation, entropy, energy and correlation coefficient, are to be extracted from the accelerometer and gyroscope in each segment and form a feature vector to represent the segment. The feature vector together with corresponding label constitute a training sample. We then conduct feature selection on the training set to select a subset of discriminant features from the original feature space in order to remove irrelevant and redundant features. Last, we can train and optimize an activity recognizer, and deploy the model on the smartphone rather than re-train the classifier. In the online prediction phase, activity recognition is performed on the smartphone in order to make real-time response. The sensor data generated by the inertial sensors is first segmented into frames using the same processing techniques as the offline stage. We then extract features from the segment in accordance with the ones obtained via the feature selection algorithm, and get a feature vector. The activity recognizer finally gives the activity label of interest associated with the feature vector.

B. Feature Extraction

In activity recognition, we are required to divide the time-series sensor data into segments before extracting features, since standard classification models are not suitable for this type of data. In handling time-series data, the sliding window technique is widely used and has proven effective [6], [42]. According to whether there exists overlapping part between two consecutive windows, it has two classical schemes: sliding window without overlapping, and sliding window with overlapping. For the purpose of illustration, Fig. 2 presents two schemes with an example of segmenting the accelerometer signal, where a_x, a_y, and a_z represent the three components of a triaxial accelerometer, Δt means the window length, and $\overrightarrow{a_t}$ refers to the readings of a_x, a_y, and a_z in the period of time $[t, t + \Delta t]$. In the case of non-overlapping, $\overrightarrow{a_t}$ and $\overrightarrow{a_{t+1}}$ come from different periods of time, as shown in Fig. 2(a). For the overlapping situation, $\overrightarrow{a_t}$ and $\overrightarrow{a_{t+1}}$ share parts of the sensors readings. Generally, the sliding window with overlapping has better smoothness property and is more suitable for analyzing continuous data.

Next, we extract informative features from the each window before training the activity recognizer. In actual use, researchers try to extract as many as features in the time-domain as well as in the frequency-domain in order to capture the slight differences between different activities. For the triaxial accelerometer and triaxial gyroscope embedded in a smartphone, in order to compare their power and evaluate their synthetic effects in recognizing activity, we note the window related to the accelerometer as $\overrightarrow{a_t}$, and the gyroscope-related window as $\overrightarrow{b_t}$. $\overrightarrow{a_t}$ and $\overrightarrow{b_t}$ may contain different number of readings, but are of equal window size Δt. Assume that the sampling frequency is p, then there are $N = p^* \Delta t$ readings in a window. For the triaxial accelerometer, each readings consists of a_x, a_y, and a_z corresponding to the three axes, and there are N readings for each axis, which can form a vector with length N. We can extract a variety of statistical features that are previously shown to be effective for activity recognition. In the time-domain, according to the N readings in a single window, for each axis, we can calculate the mean value, standard deviation, maximal value, minimal value, median absolute deviation, signal magnitude area, energy measure, signal entropy, interquartile range, autoregression coefficients with different orders, time in milliseconds between peaks, and other measures [14], [23], [43], [44]. In addition, we can extract useful features when considering the relationship of different axes, such as the correlation coefficients between two signals, the angle between two vectors, and the uncorrelated energy between two axes, and the mean value in a single window [43]. Besides these, we can apply the Fast Fourier Transform algorithm (FFT) to extract the frequency components [33]. In the frequency domain, for each axis, commonly used features include, but not limited to, the frequency component with largest magnitude, the weighted average of the frequency components, skewness of the signal, kurtosis of the signal, and the signal energy and entropy [33], [43]. Also, we can extract those statistical features as in the time domain. Similarly, for the triaxial gyroscope, each readings consists of b_x, b_y, and b_z corresponding to the three axes. We can extract the statistical features from both the time- and frequency-domain. When using the accelerometer and gyroscope simultaneously, we can obtain extra features $h(\overrightarrow{a_t}, \overrightarrow{b_t})$. For example, we can calculate the angle between the body gyroscope and the gravity acceleration signal [43].

Particularly, to indicate that the features are extracted from which one of the two sensing units, we note the accelerometer related features as $F_a = f(\overrightarrow{a_t})$, gyroscope related features as $F_b = g(\overrightarrow{b_t})$, and hybrid features as $F = F_a \oplus F_b \oplus h(\overrightarrow{a_t}, \overrightarrow{b_t})$, where f, g, and h are named the feature extraction functions, $\oplus$ is the concatenation function to merge several feature vectors into a single vector. Therefore, we can obtain three subsets of features for further analysis.

C. Proposed Feature Selection Method

In the feature extraction stage, researchers tend to extract a large range of time-domain and frequency-domain features from (each) components of the accelerometer and gyroscope in order to better characterize the original sensor signal. This may introduce irrelevant and redundant features and deteriorate the performance of the activity recognizer, therefore, a feature selection process is favorable in order to obtain better recognition rate and reduce the time cost in online prediction. Generally, feature selection methods can be categorized into two groups: filter-based and wrapper-based methods [39], [45]. Filter-based methods are independent from a classifier, and have lower computational complexity and better generalization ability. Because filter methods evaluate the quality of a feature or a subset of features by using only the intrinsic properties of the training samples, they are flexible in combination with a variety of classifiers. In contrast to filter methods, wrapper methods are specific to a given classifier to evaluate the quality of a candidate subset, and these methods tend to obtain better classification performance than the filter methods at the cost of a high time complexity [29].

In this study, we propose to hybridize the filter and wrapper methods to achieve the tradeoff between time cost and accuracy. The proposed method consists of two stages. In the first phase, we evaluate the importance of each feature in the original feature space F using ReliefF algorithm and rank all the feature in descending order. ReliefF is distance-based filter measure and has great power in selecting discriminant

Algorithm 1. Proposed feature selection approach FW

Input: *Data* with feature set F and class label C, *count*;

Output: S; //selected feature subset

1 ranked feature list R using a filter method:

2 $S = \{R_1\}$; //select the first feature of R

3 acc = evaluate(*classifier*, $Data^{\downarrow\{S\cup C\}}$);

4 **for** i = 2 **to** n **do**

5   S_{new} = update(copy(S), add(R_i));

6   [acc_{new}, num] = evaluate(*classifier*, $Data^{\downarrow\{S_{new}\cup C\}}$);

7   **if** ($acc_{new} > acc$ && $num \geq count$) **then**

8         $S = [S, R_i]$;

9         $acc = acc_{new}$;

10 **return** S;

Fig. 3. Proposed feature selection method.

features [45]. However, the obtained feature ranking fails to consider the interaction between features. Consequently, ReliefF may obtain a subset of high ranking features with redundancy among them, and discard the features that are ranked behind but can improve the recognition rate when combined with others. To alleviate this problem, in the second stage, we further use the sequential forward selection scheme and take the classification accuracy as the feature inclusion metric to find the best combination of features in a wrapper way. Specifically, after using ReliefF, we can obtain a sequence of ranked features R according to their relevance to the target variable. The first feature in R is the most relevant to the target and the last one is the least relevant to the target. Then, starting from the first feature, we incrementally add feature f from the sequence of ranked features to the selected subset of features S by testing whether f satisfying the inclusion criteria. If f satisfying the criteria, then we select f into S, update the current classification accuracy, and evaluate the next candidate feature in R; otherwise, we skip f and evaluate the next candidate features. Continue the above procedure until all the features have been evaluated. Fig. 3 presents the pseudo-code of the proposed feature selection algorithm. For illustration purpose, we note the proposed hybrid method as FW. In this study, we adopt the following criteria to decide whether a new feature f is added to the selected feature subset S: (1) a five-fold cross validation scheme is employed, and (2) the new feature f is included only if the average accuracy of the five-fold cross validation over $Data^{\downarrow\{S\cup f\cup C\}}$ is better than that of the five-fold cross validation over $Data^{\downarrow\{S\cup C\}}$ and at least *count* of the five-folds works well. *count* is actually a counter for recording how many times the five classification accuracy obtain over $Data^{\downarrow\{S\cup f\cup C\}}$ is better than the average accuracy of the five-fold cross validation over $Data^{\downarrow\{S\cup C\}}$. C is the target variable, and $Data^{\downarrow\{S\cup C\}}$ represents projecting the dataset *Data* over its attributes S and C. For better control of noise and over-fitting, the recommended empirical values for *count* are two or three. In Fig. 3, line 1 is the filter stage, lines 2-9 represents the wrapper stage, and line 6 indicates the inclusion criteria. The returned items of the function *evaluate*(*classifier*, $Data^{\downarrow\{S_{new}\cup C\}}$) include the average classification accuracy S_{new} over $Data^{\downarrow\{S\cup f\cup C\}}$ and the number *num* indicates how many times the five classification accuracies over $Data^{\downarrow\{S \cup f \cup C\}}$ are better than the average classification accuracy over $Data^{\downarrow\{S \cup C\}}$. *classifier* stands for the used learning algorithm in wrapper methods. Such a method enables us to select relevant features efficiently and avoids the situation where low ranked features have no chance to be chosen.

IV. Experimental Results and Analysis

A. Experimental Setup and Experimental Dataset

Since one of our tasks is to show and compare the power of triaxial accelerometer and triaxial gyroscope embedded in a smartphone in human activity recognition when they are used simultaneously or separately, we plan to evaluate the recognition performance in the following three settings: (1) only sensor data generated from the gyroscope component are used; (2) only the triaxial accelerometer related data are considered; (3) both accelerometer and gyroscope data are involved. Obviously, we treat the accelerometer and gyroscope separately in the former two cases, and consider their synergic effects in the last case, which enables us to extract and obtain a discriminant subset of features to differentiate very similar activities such as sitting and standing. Consequently, such a setting is reasonable and sufficient to fulfill the task.

Our next task is to explore the use of feature selection methods in building an effective activity recognizer. In wrapper-based feature selection, a classifier is required and used to measure the quality of a candidate feature. In addition, to evaluate the power of a feature selection method, a classifier is also required to evaluate the quality of the finally selected feature subset. In our study, two commonly used learning algorithms with different metrics are applied, i.e. naïve bayes and k-nearest-neighbor. Particularly, naïve bayes classifier is sensitive to redundant features and irrelevant features deteriorate the performance of k-nearest-neighbor classifier. Also, we use the same learning algorithm for selecting features and verifying the goodness of obtained feature subset. Besides, to show the benefits and superiority of the proposed feature selection method FW in improving recognition rate and time performance, we compare it with three other feature selection methods: principal component analysis (PCA) [46], fast correlation-based filter (FCBF) [40] and wrapper method with sequential forward selection (Wrapper) [35]. For FW, to better control noise and overfitting, *count* usually takes the value of 2 or 3, and we set $count = 2$ in our study since it leads to a better performance. In addition, we note the method without using feature selection as "Original".

PCA is a feature extraction method, and it reduces the feature dimensionality by transforming the original data to a lower dimensional feature space. PCA first calculates the data covariance matrix, and its associated eigenvalues $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m)$ and eigenvectors $\nu = (\nu_1, \nu_2, \ldots, \nu_m)$. In the dimensionality reduction step, the eigenvectors pertaining to the k ($k \leq m$) largest eigenvalues are kept [46]. In our study, we choose to keep $threshold = 99.0\%$ variance information of the raw data, and determine the minimal value of k using (1).

$$\sum_{j=1}^{k} \lambda_j / \sum_{j=1}^{m} \lambda_j \geq threshold \tag{1}$$

FCBF is a filter-based feature selection [40]. It obtains the final feature subset with a two-stage procedure. In the first phase, FCBF filters out those features whose relevance with the target class is less than a predefined threshold γ. It further eliminates redundant features using Markov blanket technique in the later phase. For FCBF, we set $\gamma = 0$, indicating that irrelevant features are removed in the first stage.

Wrapper adopts a greedy search strategy to incrementally select features. Starting from an empty set, Wrapper first selects the feature that is most relevant to the target class, and then searches for the next candidate feature that most contributes to the enhancement of the classification accuracy. Continue the process until there is no improvement in accuracy or there is no remaining candidate feature. In feature selection, Wrapper evaluates only $O((S+1)N)$ candidate feature subsets if S features are finally selected and $O(N^2)$ feature subsets in the worst case on a data set with N features.

In order to evaluate the performance of the proposed activity recognition framework, we conducted extensive experiments on a publicly available dataset that was collected by the triaxial accelerometer and triaxial gyroscope embedded in a Samsung Galaxy S II smartphone [43]. The smartphone, running the Android operating system, is equipped with a dual-core cortex-A9 CPU (1.2 GHz processor and 1GB RAM), and has a removable Li-Ion 1650 mAh battery and can work for several hours continuously. The acceleration is measured with the STMicroelectronics K3DH 3-axis accelerometer, which is accurate to ± 2G with resolution $= 0.0625$ (G is the gravitational constant), and the angular velocity is obtained by the K3G gyroscope sensor with the maximum range of 8.72665 and resolution $= 0.000305433$. The dataset was collected when each of the thirty volunteers aged between nineteen and forty-eight performed activities with a smartphone attached to the waist. The accelerometer and gyroscope worked at a sampling rate 50Hz.The task on this dataset is to distinguish the following six human physical activities: *walking*, *upstairs*, *downstairs*, *sitting*, *standing* and *lying*. These activities are closely related to the daily activity level and functional health of an individual, and can be an indicator for the occurrence of falls. The dataset was pre-processed and segmented with a fixed-length sliding window of 2.56 seconds and fifty percent overlap between two adjacent segments. Therefore, there were 128 sensor readings in a single window. In addition, we extract a range of features associated with the accelerometer and gyroscope in both time domain and frequency domain. Consequently, 348 features were extracted from accelerometer data with 164 time-domain features and 184 frequency-domain features. For the gyroscope component, 106 features were extracted in time domain and 105 features were from the frequency domain. When the accelerometer and gyroscope are used, two extra features were extracted in the time domain. Also, each sample was manually labeled with its activity label. In total, there are 7352 training samples and 2947 test samples.

TABLE I
EXPERIMENTAL DATA DESCRIPTION

Data	Number of features		
	time-domain	frequency-domain	total
Gyroscope	106	105	211
Accelerometer	164	184	348
Both	272	289	561

TABLE II
CONFUSION MATRIX FOR A THREE-CLASS CLASSIFICATION PROBLEM

		True Label			
		1	2	3	sum
Inferred Label	1	TP_{11}	FP_{12}	FP_{13}	NI_1
	2	FP_{21}	TP_{22}	FP_{23}	NI_2
	3	FP_{31}	FP_{32}	TP_{33}	NI_3
	sum	NT_1	NT_2	NT_3	total

TABLE III
EXPERIMENTAL RESULTS OF THE NUMBER OF SELECTED FEATURES AND CLASSIFICATION ACCURACY USING NAÏVE BAYES

Dataset	Original		PCA		FCBF		Wrapper		FW	
	acc	#fea	acc	#fea	acc	#fea	acc	#fea	acc	#fea
Gyroscope	50.8	211	66.6	78	63.0	15	68.4	18	69.8	46
Accelerometer	81.1	348	83.3	104	85.6	19	87.9	18	88.1	78
Both	77.0	561	85.6	178	88.2	34	89.5	23	90.1	74

TABLE IV
PRECISION COMPARISON USING NAÏVE BAYES

Dataset	Precision				
	Original	PCA	FCBF	Wrapper	FW
Gyroscope	55.0	66.8	65.7	68.6	71.1
Accelerometer	82.6	83.1	86.1	88.3	88.2
Both	79.2	85.3	88.4	90.0	90.4

Table I summarizes the dataset, and “Both” indicates accelerometer and gyroscope.

B. Performance Measures

To evaluate the effectiveness of the proposed feature selector and show its superiority over the comparative methods, we compare them in terms of the number of selected features, the obtained classification performance, and the actual running time in prediction. In the evaluation of the activity recognizer, a confusion matrix that contains the actual outputs and predicted outputs is applicable to evaluate the classification performance [47]. Table II presents an example of confusion matrix for activity recognition in the case of three activities. Accordingly, we can use the following four measures to show the classification performance, and the higher precision, recall, F1 and accuracy, the better the constructed recognizer.

(1) *Precision* represents the weighted average of the fraction of the inferred activity labels that are correctly predicted for each activity class. For a classification problem with C classes, *Precision* can be calculated with (2).

$$\text{Precision} = \frac{1}{C}\sum_{i=1}^{C}\frac{TP_{ii}}{NI_i}, \tag{2}$$

where TP_{ii} is the number of test samples that are corrected classified for the inferred label i; NI_i shows the total number of test samples that are classified as label i, and equals the sum of the number in corresponding row.

$$NI_i = TP_{ii} + \sum_{j=1, j\neq i}^{C} FP_{ij} \tag{3}$$

(2) *Recall* refers to the weighted average of the fraction of the true activity labels that are correctly classified for each activity class. For a classification problem with C classes, we can measure *Recall* using (4).

$$\text{Recall} = \frac{1}{C}\sum_{i=1}^{C}\frac{TP_{ii}}{NT_i}, \tag{4}$$

where NT_i indicates the number of test samples with true label i, and can be obtained by totaling the number of corresponding column.

$$NT_i = TP_{ii} + \sum_{j=1, j\neq i}^{C} FP_{ji} \tag{5}$$

(3) $F1$ provides a way to combine precision and recall into a single metric, is calculated using (6). $F1$ takes a real number between 0 and 1, and 1 indicates that the classifier can correctly classify all test samples.

$$\text{F1} = \frac{2 * precision * recall}{precision + recall} \tag{6}$$

(4) *Accuracy* means the probability of correctly classifying each sample, and equals the number of samples that are correctly grouped. Equation (7) present the formula, where *total* indicates the total number of test samples.

$$\text{Accuracy} = \frac{\sum_{i=1}^{C} TP_{ii}}{total} = \frac{\sum_{i=1}^{C} TP_{ii}}{\sum_{i=1}^{C} NI_i} = \frac{\sum_{i=1}^{C} TP_{ii}}{\sum_{i=1}^{C} NT_i}. \tag{7}$$

C. Experimental Results

We conducted experiments on a desktop with a 3.2GHz processor and 4G memory storage. For the KNN classifier, 1NN was used as the learning function. Tables III-V present the experimental results of the number of used features, accuracy, precision and F1 in the case of naïve bayes. Tables VI-VIII present the experimental results using k-nearest-neighbor. The last row “Both” in each table refers to the results obtained by using both accelerometer and gyroscope data.

1) Accelerometer vs. Gyroscope vs. Both: In this subsection, we analyze the power of accelerometer and gyroscope when they are used separately and simultaneously in the situation where the feature selection method is not involved. This helps determine whether there exist irrelevant and redundant ones in the extracted features. According to the experimental results

TABLE V
F1 COMPARISON USING NAÏVE BAYES

Dataset	F1				
	Original	PCA	FCBF	Wrapper	FW
Gyroscope	53.4	66.8	64.5	68.6	70.5
Accelerometer	81.5	83.0	85.6	87.9	88.0
Both	78.0	85.3	88.1	89.6	90.1

TABLE VI
EXPERIMENTAL RESULTS OF THE NUMBER OF SELECTED FEATURES AND CLASSIFICATION ACCURACY USING k-NEAREST-NEIGHBOR

Dataset	Original		PCA		FCBF		Wrapper		FW	
	acc	#fea	acc	#fea	acc	#fea	acc	#fea	acc	#fea
Gyroscope	65.8	211	65.0	78	52.8	15	33.6	1	64.2	96
Accelerometer	84.3	348	83.6	104	80.0	19	83.0	12	85.1	58
Both	87.8	561	87.9	178	82.9	34	83.7	21	87.8	66

TABLE VII
PRECISION COMPARISON USING k-NEAREST-NEIGHBOR

Dataset	Precision				
	Original	PCA	FCBF	Wrapper	FW
Gyroscope	66.4	65.7	53.4	33.8	64.9
Accelerometer	84.5	83.8	79.8	82.8	85.4
Both	88.0	87.9	82.7	83.7	88.1

TABLE VIII
F1 COMPARISON USING k-NEAREST-NEIGHBOR

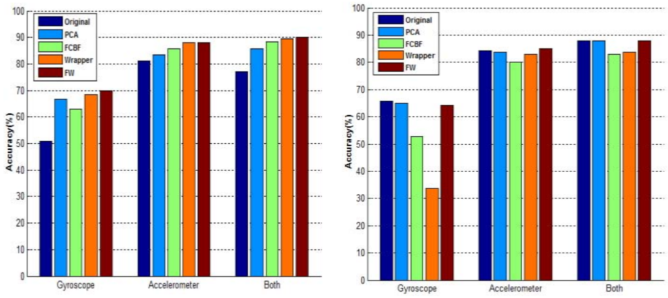
Dataset	F1				
	Original	PCA	FCBF	Wrapper	FW
Gyroscope	66.3	65.5	53.2	33.7	64.7
Accelerometer	84.3	83.5	79.6	82.7	85.1
Both	87.7	87.7	82.6	83.5	87.7

shown in Tables III-VIII. We can observe that the use of the triaxial accelerometer consistently obtains better performance than that of using the triaxial gyroscope in terms of accuracy, precision, and F1. Specifically, for naïve bayes, the use of accelerometer obtains 81.1% accuracy, 82.6% precision and 81.5% F1, while the use of gyroscope only obtains 50.8% accuracy, 55.0% precision and 53.4% F1. In the case of 1-nearest-neighbor classifier, the use of gyroscope achieves 65.8% accuracy, 66.4% precision and 66.3% F1, respectively, which is significantly lower than that of the accelerometer with 84.3% accuracy, 84.5% precision and 84.3% F1. This indicates that the acceleration contains more discriminant and valuable information than the angular velocity in human activity recognition, demonstrating the superiority of accelerometer over gyroscope.

Furthermore, it has been known to us that the gyroscope and accelerometer can provide relatively novel features to each other. For example, the accelerometer has the capacity to obtain the acceleration, while the gyroscope can measure the angular velocity. However, from Tables III-VIII, we can observe that the fusion of both accelerometer and gyroscope data fails to guarantee better performance all the time. Specifically, in the case of naïve bayes, we obtained 77.0% accuracy with the accelerometer and gyroscope, which was lower than 81.1% of the accelerometer. Also, in terms of F1, using accelerometer and gyroscope contributes to the enhancement in the activity recognition rate. Accordingly, the data fusion improves the accuracy, precision and F1 to 87.8%, 88.0% and 87.7%, respectively, from the corresponding 65.8%, 66.4% and 66.3% of the gyroscope and the 84.3%, 84.5% and 84.3% of the accelerometer. The possible reason for this outcome is that there exists redundant features provided by the accelerometer and gyroscope and that naïve bayes classifier is built under the assumption of conditional independence among features.

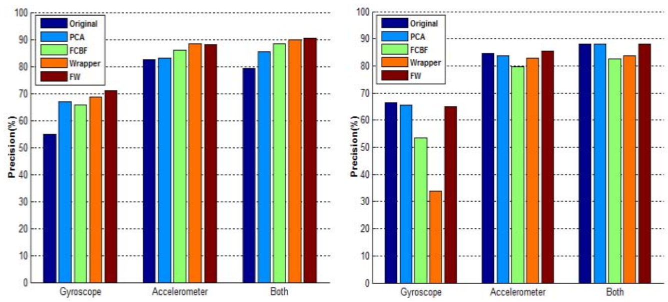
2) With vs. Without Feature Selection: In this subsection, we show the benefits of feature selection methods and evaluate the effectiveness of the proposed feature selector. Tables III-VIII present the results that are obtained without feature selection as well as the results obtained with our proposed feature selection method and its three comparative ones. According to the experimental results, we can observe that the application of feature selection method not only greatly reduces the feature dimensionality, but also achieves better classification performance except for the situation where k-nearest-neighbor classifier is used on the gyroscope data. Specifically, for naïve bayes, in terms of the number of selected features, PCA can reduce the number of gyroscope-related features from 211 to 78, the number of accelerometer-related features from 348 to 104, and reduce the total number of features to 178. In terms of classification performance, PCA can obtain 66.6% accuracy on the gyroscope data, 83.3% accuracy on the accelerometer data, and 85.2% accuracy over the combined data. FCBF can improve the accuracy to 63.0%, 85.6% and 88.2% on the gyroscope data, accelerometer data and the combined data, respectively. Significantly, FCBF uses 15, 19 and 34 features in achieving such a high classification. For Wrapper, the number of selected features is similar to that of FCBF, and it can achieve slightly better classification performance. In comparison with PCA, FCBF and Wrapper, our proposed method FW can obtain the best classification performance, although the number of selected features is more than the ones that are obtained by FCBF and Wrapper. In the aspect of precision, recall and F1, we can draw conclusions similar to accuracy. For k-nearest-neighbor, FW selects 96, 58 and 66 features from the gyroscope data, accelerometer data and the combined data, respectively. Although the number of features selected by FCBF and Wrapper is less than that of FW, FW outperforms them in terms accuracy, precision, recall and F1. For example, FW can obtain 87.7% F1, which is 5.1% higher than that of FCBF and 4.2% higher than that of Wrapper. In comparison with Original and PCA methods, FW can select a more compact subset of features. This helps to construct an activity recognizer with better generalization ability.

Besides, we can also observe that the fusion of gyroscope data and accelerometer data consistently obtains better classification performance than that of using only gyroscope data or accelerometer data in the situation where feature selection is performed. In the case of naïve bayes, for FCBF, it obtains 88.2% accuracy, 88.4% precision and 88.1% F1 over



(a) (b)

Fig. 4. Accuracy comparison of the different feature selection methods. (a) Using Naïve Bayes classifier. (b) Using k-nearest-neighbor classifier.



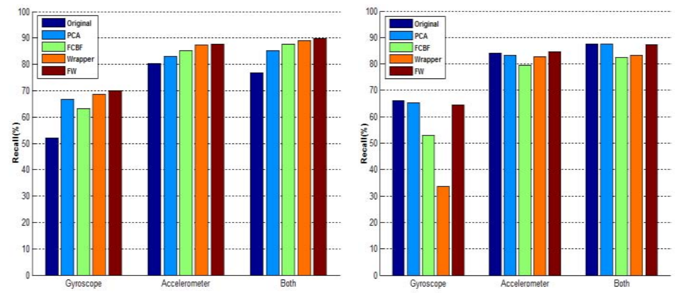
(a) (b)

Fig. 5. Precision comparison of the different feature selection methods. (a) Using Naïve Bayes classifier. (b) Using k-nearest-neighbor classifier.

the combined data, which are higher than that obtained over the gyroscope data as well as over the accelerometer data. For Wrapper, it obtains 83.7% accuracy over the combined data, which is 0.7% higher than the one over the accelerometer data and 50.1% higher than the one over the gyroscope data. For FW, it improves the classification accuracy from 77.0% to 90.1% over the hybrid data, and obtains 88.1% accuracy over the accelerometer data. The k-nearest-neighbor classifier can obtain similar results to naïve bayes. This indicates that directly working on the hybrid data without feature selection cannot guarantee achieving better classification performance than that over single-source data, and that an effective feature selector is desirable to eliminate irrelevant and redundant features.

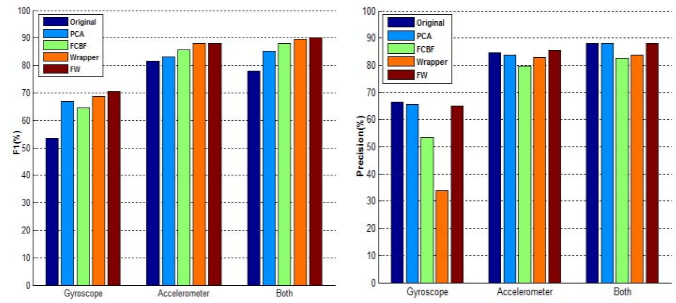
In order to present an intuitive impression of the experimental results obtained with different feature selection methods over the three different types of data, Figs. 4–7 show the results of accuracy, precision, recall and F1, respectively. In each figure, we present the results of k-nearest-neighbor as well as naïve bayes. Generally, we can come to three conclusions: (1) recognizing human activities with the accelerometer can obtain better performance in terms of the four metrics; (2) hybridizing the accelerometer and gyroscope data has superiority over the single-source data in achieving better recognition rates; (3) the proposed feature selector FW outperforms its comparative ones in accuracy, precision, recall and F1, which demonstrates the effectiveness of FW.

To better show the synthetic effects of the accelerometer and gyroscope, we present the number of selected features over the combined data for naïve bayes and k-nearest-neighbor, and give how many features of them are from the accelerometer and the gyroscope, respectively. Table IX presents



(a) (b)

Fig. 6. Recall comparison of the different feature selection methods. (a) Using Naïve Bayes classifier. (b) Using k-nearest-neighbor classifier.



(a) (b)

Fig. 7. F1 comparison of the different feature selection methods. (a) Using Naïve Bayes classifier. (b) Using k-nearest-neighbor classifier.

TABLE IX

NUMBER OF SELECTED FEATURES OVER THE COMBINED DATA

Method	#both	#accelerometer	#gyroscope
Naïve bayes	74	52	22
K-nearest-neighbor	66	48	18

corresponding results, in which the second column "#both" means the total number of selected features, the third column "#accelerometer" indicates how many of them are related to the accelerometer, and the fourth column "#gyroscope" corresponds to the gyroscope. We can observe from Table IX that the finally selected features come from both accelerometer and gyroscope, which partially explains the superiority of the fusion of accelerometer and gyroscope data over the single sensing component in human activity recognition. We can also see that the two feature subsets selected by naïve bayes and k-nearest-neighbor are different, implying that our proposed method is specific to the learning algorithm.

Additionally, in order to gain a better insight into the activity recognition problem and the proposed feature selection method FW, the corresponding confusion matrix was constructed. Tables X-XII show the results using naïve bayes and FW over the gyroscope, accelerometer and the hybrid data, respectively. Experimental results about k-nearest-neighbor are presented in Tables XIII-XV. According to the results, we can observe that the single use of gyroscope or accelerometer can distinguish dynamic activity (walking, upstairs and downstairs) from static activity (sitting, standing and lying) with a high accuracy. Specifically, for naïve bayes, activity recognition with the hybrid data makes 15 classification errors out of the 2947 test samples, compared to the 28 errors of using

TABLE X

CONFUSION MATRIX FOR ACTIVITY RECOGNITION WITH GYROSCOPE USING NAÏVE BAYES AND FW

	Walking	Upstairs	Downstairs	Sit	Stand	Lying
Walking	**429**	54	98	0	0	0
Upstairs	16	**390**	46	6	13	6
Downstairs	51	27	**276**	0	3	0
Sit	0	0	0	**252**	39	151
Stand	0	0	0	186	**448**	118
Lying	0	0	0	47	29	**262**

TABLE XI

CONFUSION MATRIX FOR ACTIVITY RECOGNITION WITH ACCELEROMETER USING NAÏVE BAYES AND FW

	Walking	Upstairs	Downstairs	Sit	Stand	Lying
Walking	**470**	33	28	0	0	0
Upstairs	5	**424**	51	3	4	0
Downstairs	21	14	**341**	0	0	0
Sit	0	0	0	3**59**	64	0
Stand	0	0	0	129	**464**	0
Lying	0	0	0	0	0	**537**

TABLE XII

CONFUSION MATRIX FOR ACTIVITY RECOGNITION WITH ACCELEROMETER AND GYROSCOPE USING NAÏVE BAYES AND FW

	Walking	Upstairs	Downstairs	Sit	Stand	Lying
Walking	**473**	17	28	0	0	0
Upstairs	2	**444**	54	4	10	1
Downstairs	21	10	**338**	0	0	0
Sit	0	0	0	**381**	38	0
Stand	0	0	0	106	**484**	0
Lying	0	0	0	0	0	**536**

TABLE XIII

CONFUSION MATRIX FOR ACTIVITY RECOGNITION WITH GYROSCOPE USING k-NEAREST-NEIGHBOR AND FW

	Walking	Upstairs	Downstairs	Sit	Stand	Lying
Walking	**351**	100	81	0	1	0
Upstairs	41	**303**	43	1	1	1
Downstairs	103	68	**296**	0	1	0
Sitting	0	0	0	**285**	76	162
Standing	1	0	0	149	**370**	87
Lying	0	0	0	56	83	**287**

TABLE XIV

CONFUSION MATRIX FOR ACTIVITY RECOGNITION WITH ACCELEROMETER USING k-NEAREST-NEIGHBOR AND FW

	Walking	Upstairs	Downstairs	Sit	Stand	Lying
Walking	**454**	84	43	0	0	0
Upstairs	25	**375**	53	0	0	0
Downstairs	17	12	**324**	0	0	0
Sit	0	0	0	**384**	98	0
Stand	0	0	0	107	**434**	0
Lying	0	0	0	0	0	**537**

TABLE XV

CONFUSION MATRIX FOR ACTIVITY RECOGNITION WITH ACCELEROMETER AND GYROSCOPE USING k-NEAREST-NEIGHBOR AND FW

	Walking	Upstairs	Downstairs	Sit	Stand	Lying
Walking	**446**	74	52	0	0	0
Upstairs	21	**385**	36	1	0	0
Downstairs	29	12	**332**	0	0	0
Sit	0	0	0	**395**	39	0
Stand	0	0	0	95	**493**	0
Lying	0	0	0	0	0	**537**

the gyroscope and 7 errors of using the accelerometer. For k-nearest-neighbor, it makes 5, 0 and 1 errors when conducting on the gyroscope data, accelerometer data and the hybrid data, respectively. Furthermore, we can also observe that the accelerometer has the superiority over the gyroscope in distinguishing lying from sitting and standing. For example, for naïve bayes, in Tables X-XI, the use of gyroscope mistakenly predicts 151 sitting test samples and 118 standing test samples to be with lying labels, and classifies 47 lying test sample to sitting and 29 lying test samples to standing. In contrast, the use of accelerometer can accurately differentiate between lying and other two activities. In addition, we can observe that combining gyroscope and accelerometer data contributes to classify similar activities, such as sitting vs. standing, upstairs vs. downstairs. For instance, according to Tables X-XII, we can see that it makes 225 errors over the gyroscope data, 193 errors over the accelerometer data in distinguishing between sitting and standing, and reduces it to 144 errors with the combined data. Results in Tables XIII-XV show that in differentiating upstairs and downstairs, the combined data results in 48 errors in comparison to the 111 errors over the gyroscope data and the 65 errors over the accelerometer data.

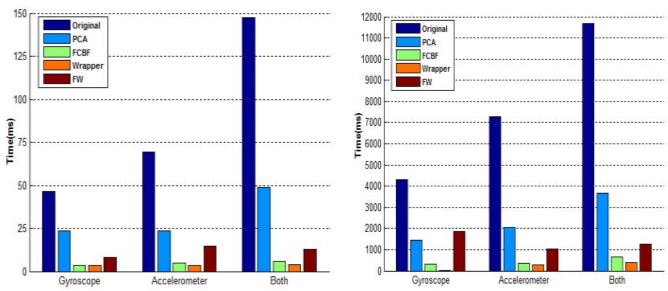
3) Time Cost(s) Comparison: In the online activity recognition, besides accuracy, the time complexity in making prediction is another important factor that is worthy of consideration. For the case of smartphone, the higher time cost definitely consumes more battery energy. Consequently, this inevitably shortens the time of service and affects user experience. In this subsection, we investigate and compare the time performance of the four feature selection methods. We also present the experimental results without using feature selection as a comparison, as shown in Tables XVI and XVII. Meanwhile, we provide the time cost comparison in the form of graph, as shown in Fig. 8. According to the experimental results, we can observe that the use of feature selection methods helps reduce the actual time cost significantly over all of the three different types of data. Furthermore, we can observe that FCBF and Wrapper are more efficient than PCA and FW, and that FW outperforms PCA. Over the accelerometer data, it costs 69.4 seconds with all the features,

TABLE XVI
TIME COST(S) COMPARISON USING NAÏVE BAYES

Dataset	Original	PCA	FCBF	Wrapper	FW
Gyroscope	46.6	23.6	3.5	3.6	8.4
Accelerometer	69.4	23.4	4.7	3.6	14.5
Both	147.5	48.6	5.7	4.1	13.0

TABLE XVII
TIME COST(S) COMPARISON USING k-NEAREST-NEIGHBOR

Dataset	Original	PCA	FCBF	Wrapper	FW
Gyroscope	4306.7	1455.1	325.8	14.1	1855.6
Accelerometer	7295.2	2057.7	373.1	276.2	1052.4
Both	11693.7	3675.9	655.7	410.0	1249.3



(a) (b)

Fig. 8. Time cost comparison of the different feature selection methods. (a) Using Naïve Bayes classifier. (b) Using k-nearest-neighbor classifier.

while FW reduces it to 14.5 seconds in comparison with 23.4 seconds of PCA, 4.7 seconds of FCBF and 3.6 seconds of Wrapper. Also, working in a feature space with higher dimensionality, activity recognition with all the features costs 147.5 seconds. In contrast, it takes PCA 48.6 seconds, FCBF 5.7 seconds and Wrapper 4.1 seconds, and FW is 11.0 times faster than the Original. Similarly, for k-nearest-neighbor, as shown in Table XVII and Fig. 8(b), compared with the Original method, this case leads to 3.2-, 17.2-, 33.3-, and 5.6-fold improvements in the average time cost(s) for PCA, FCBF, Wrapper and FW, respectively.

Overall, according to the experimental results and in-depth analysis, we conclude that FW obtains comparable performance to PCA in terms of accuracy, precision, recall and F1 while having lower time complexity. Also, in comparison with FCBF and Wrapper, although FCBF and Wrapper have better time performance than FW, FW consistently achieves better classification performance than FCBF and Wrapper. Considering that the primary goal is to pursue high accuracy, therefore, FW is a better choice towards a tradeoff between accuracy and time cost.

V. CONCLUSION

In assisted living systems, human activity recognition helps bridge the gap between the low-level sensor data and the high-level human-centric applications, and plays an increasing important role in improving the life quality and promoting health at an individual as well as the population level. Among the various available sensing components, the inertial sensing units built-in a smartphone are widely used due to its convenience, high adherence and low intrusiveness. To explore the power of the triaxial accelerometer and gyroscope in activity recognition, in this study, we conducted an extensive research. Considering the limited processing power and battery energy of a smartphone, we first presented an activity recognition framework that works in an offline training and online prediction scheme. This enables us to optimize the selection of features and corresponding model parameters. We then discussed the sliding window technique and illustrated the features that can be extracted in the time domain and frequency domain. With the aim to find a subset of discriminant features, we further proposed a novel feature selection method that takes the advantages of filter and wrapper methods. Last, extensive experiments and analysis were conducted to evaluate the effectiveness of the proposed feature selector and to compare the power of accelerometer and gyroscope where they were used simultaneously or separately. Experimental results show that the triaxial accelerometer can provide more discriminant information than the triaxial gyroscope, and that the fusion of accelerometer and gyroscope data contributes to obtain better classification performance. In addition, an effective feature selection method can significantly reduce the feature dimensionality and further improve the recognition rate and time performance, and our proposed method FW makes a better tradeoff between the accuracy and time cost.

For the future work, we plan to conduct further research in the following lines. First, since feature selection has the capacity to improve the activity recognizer in terms of generalization ability and time complexity, exploring other effective feature selectors and comparing them with the proposed one in this study remains a topic for future research. Second, although we tested the effectiveness of the proposed approach in recognizing six physical activity, it is actually a general framework that can be applied to other situations such as activity recognition with other sensors, and other classification and regression problems. So one of the further work involves applying the proposed approach in related areas. Third, because of the inter-subject variability, it is often difficult to build a robust classification model with good generalization ability. Correspondingly, one possible avenue is to explore how to transfer the knowledge obtained on one subject to other individuals. Finally, accurate activity recognition helps perceive the state of an individual and facilitates the design and development of human-centric applications. Therefore, we intend to develop an assisted living system in order to understand the user behavior patterns and promote health in a home setting.

REFERENCES

[1] H. Ghayvat, J. Liu, S. C. Mukhopadhyay, and X. Gui, "Wellness sensor networks: A proposal and implementation for smart home for assisted living," *IEEE Sensors J.*, vol. 15, no. 12, pp. 7341–7348, Dec. 2015.

[2] N. K. Suryadevara and S. C. Mukhopadhyay, "Determining wellness through an ambient assisted living environment," *IEEE Intell. Syst.*, vol. 29, no. 3, pp. 30–37, May/Jun. 2014.

[3] S. Chernbumroong, S. Cang, A. Atkins, and H. Yu, "Elderly activities recognition and classification for applications in assisted living," *Expert Syst. Appl.*, vol. 40, no. 5, pp. 1662–1674, Apr. 2013.

[4] K. K. B. Peetoom, M. A. S. Lexis, M. Joore, C. D. Dirksen, and L. P. De Witte, "Literature review on monitoring technologies and their outcomes in independently living elderly people," *Disab. Rehabil. Assist. Technol.*, vol. 10, no. 4, pp. 271–294, Jul. 2015.

[5] F. J. Ordóñez, P. de Toledo, and A. Sanchis, “Sensor-based Bayesian detection of anomalous living patterns in a home setting,” *Pers. Ubiquitous Comput.*, vol. 19, no. 2, pp. 259–270, Feb. 2014.
[6] N. C. Krishnan and D. J. Cook, “Activity recognition on streaming sensor data,” *Pervasive Mobile Comput.*, vol. 10, pp. 138–154, Feb. 2014.
[7] E. M. Tapia *et al.*, “Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor,” in *Proc. 11th IEEE Int. Symp. Wearable Comput.*, Boston, MA, USA, Oct. 2007, pp. 37–40.
[8] O. D. Lara and M. A. Labrador, “A survey on human activity recognition using wearable sensors,” *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1192–1209, Jul. 2013.
[9] X. Ma, H. Wang, B. Xue, M. Zhou, B. Ji, and Y. Li, “Depth-based human fall detection via shape features and improved extreme learning machine,” *IEEE J. Biomed. Health Inform.*, vol. 18, no. 6, pp. 1915–1922, Nov. 2014.
[10] J. M. Chaquet, E. J. Carmona, and A. Fernández-Caballero, “A survey of video datasets for human action and activity recognition,” *Comput. Vis. Image Understand.*, vol. 117, no. 6, pp. 633–659, Jun. 2013.
[11] E. M. Tapia, S. S. Intille, and K. Larson, “Activity recognition in the home using simple and ubiquitous sensors,” in *Pervasive Computing*. Cambridge, MA, USA: Springer, 2004, pp. 158–175.
[12] Q. Zhang, M. Karunanithi, R. Rana, and J. Liu, “Determination of activities of daily living of independent living older people using environmentally placed sensors,” in *Proc. 35th EMBC*, Osaka, Japan, Jul. 2013, pp. 7044–7047.
[13] L. Chen, J. Hoey, C. D. Nugent, D. J. Cook, and Z. Yu, “Sensor-based activity recognition,” *IEEE Trans. Syst., Man, Cybern. C, Appl. Rev.*, vol. 42, no. 6, pp. 790–808, Nov. 2012.
[14] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, “Activity recognition from accelerometer data,” in *Proc. 17th IAAI*, Pittsburgh, PA, USA, 2005, pp. 1541–1546.
[15] K. Ouchi and M. Doi, “Indoor-outdoor activity recognition by a smartphone,” in *Proc. UBICOMP*, Pittsburgh, PA, USA, 2012, p. 537.
[16] L. Gao, A. K. Bourke, and J. Nelson, “Evaluation of accelerometer based multi-sensor versus single-sensor activity recognition systems,” *Med. Eng. Phys.*, vol. 36, no. 6, pp. 779–785, Jun. 2014.
[17] S. Dernbach, B. Das, N. C. Krishnan, B. L. Thomas, and D. J. Cook, “Simple and complex activity recognition through smart phones,” in *Proc. 8th IE*, Guanajuato, Mexico, Jun. 2012, pp. 214–221.
[18] W. Wu, S. Dasgupta, E. E. Ramirez, C. Peterson, and G. J. Norman, “Classification accuracies of physical activities using smartphone motion sensors,” *J. Med. Internet Res.*, vol. 14, no. 5, p. e130, Sep. 2014.
[19] A. Anjum and M. U. Ilyas, “Activity recognition using smartphone sensors,” in *Proc. Consum. Commun. Netw. Conf.*, Las Vegas, NV, USA, Jan. 2013, pp. 914–919.
[20] L. Chen, C. D. Nugent, and H. Wang, “A knowledge-driven approach to activity recognition in smart homes,” *IEEE Trans. Knowl. Data Eng.*, vol. 24, no. 6, pp. 961–974, Jun. 2012.
[21] L. Bao and S. S. Intille, “Activity recognition from user-annotated acceleration data,” in *Pervasive Computing* (Lecture Notes in Computer Science), vol. 3001. Vienna, Austria: Springer, Apr. 2004, pp. 1–17.
[22] N. C. Krishnan, D. Colbry, C. Juillard, and S. Panchanathan, “Real time human activity recognition using tri-axial accelerometers,” in *Proc. Sensors Signals Inf. Process. Workshop*, Sedona, AZ, USA, 2008, pp. 3337–3340.
[23] A. M. Khan, Y.-K. Lee, S. Y. Lee, and T.-S. Kim, “A triaxial accelerometer-based physical-activity recognition via augmented-signal features and a hierarchical recognizer,” *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 5, pp. 1166–1172, Sep. 2010.
[24] Y. Zheng, W.-K. Wong, X. Guan, and S. Trost, “Physical activity recognition from accelerometer data using a multi-scale ensemble method,” in *Proc. 25th IAAI*, Bellevue, WA, USA, 2013, pp. 1575–1581.
[25] J. R. Kwapisz, G. M. Weiss, and S. A. Moore, “Activity recognition using cell phone accelerometers,” *ACM SIGKDD Explorations Newslett.*, vol. 12, no. 2, pp. 74–82, Dec. 2010.
[26] A. Reiss, G. Hendeby, and D. Stricker, “A competitive approach for human activity recognition on smartphones,” in *Proc. ESANN*, Bruges, Belgium, 2013, pp. 455–460.
[27] E. Garcia-Ceja, R. F. Brena, J. C. Carrasco-Jimenez, and L. Garrido, “Long-term activity recognition from wristwatch accelerometer data,” *Sensors*, vol. 14, no. 12, pp. 22500–22524, Dec. 2014.
[28] U. Maurer, A. Smailagic, D. P. Siewiorek, and M. Deisher, “Activity recognition and monitoring using multiple sensors on different body positions,” in *Proc. Workshop BSN*, Cambridge, MA, USA, Apr. 2006, pp. 113–116.
[29] Y. Zhang, S. Markovic, I. Sapir, R. C. Wagenaar, and T. D. C. Little, “Continuous functional activity monitoring based on wearable tri-axial accelerometer and gyroscope,” in *Proc. 5th Int. PervasiveHealth*, Dublin, Ireland, May 2011, pp. 370–373.
[30] Q. T. Huynh, U. D. Nguyen, L. B. Irazabal, N. Ghassemian, and B. Q. Tran, “Optimization of an accelerometer and gyroscope-based fall detection algorithm,” *J. Sensors*, vol. 2015, Mar. 2015, Art. no. 452078.
[31] S.-W. Lee and K. Mase, “Activity and location recognition using wearable sensors,” *IEEE Pervasive Comput.*, vol. 1, no. 3, pp. 24–32, Jul./Sep. 2002.
[32] X. Yun, J. Calusdian, E. R. Bachmann, and R. B. McGhee, “Estimation of human foot motion during normal walking using inertial and magnetic sensor measurements,” *IEEE Trans. Instrum. Meas.*, vol. 61, no. 7, pp. 2059–2072, Jul. 2012.
[33] Y. Kwon, K. Kang, and C. Bae, “Unsupervised learning for human activity recognition using smartphone sensors,” *Expert Syst. Appl.*, vol. 41, no. 14, pp. 6067–6074, Oct. 2014.
[34] G. Bieber, J. Voskamp, and B. Urban, “Activity recognition for everyday life on mobile phones,” in *Universal Access in Human-Computer Interaction. Intelligent and Ubiquitous Interaction Environments*, vol. 5615. San Diego, CA, USA: Springer, pp. 289–296, 2009.
[35] J.-H. Chiang, P.-C. Yang, and H. Tu, “Pattern analysis in daily physical activity data for personal health management,” *Pervasive Mobile Comput.*, vol. 13, pp. 13–25, Aug. 2014.
[36] L. Zhang, J. Liu, H. Jiang, and Y. Guan, “SensTrack: Energy-efficient location tracking with smartphone sensors,” *IEEE Sensors J.*, vol. 13, no. 10, pp. 3775–3784, Oct. 2013.
[37] P. Gupta and T. Dallas, “Feature selection and activity recognition system using a single triaxial accelerometer,” *IEEE Trans. Biomed. Eng.*, vol. 61, no. 6, pp. 1780–1786, Jun. 2014.
[38] N. Bicocchi, M. Mamei, and F. Zambonelli, “Detecting activities from body-worn accelerometers via instance-based algorithms,” *Pervasive Mobile Comput.*, vol. 6, no. 4, pp. 482–495, Aug. 2010.
[39] A. Wang, N. An, G. Chen, L. Li, and G. Alterovitz, “Accelerating wrapper-based feature selection with K-nearest-neighbor,” *Knowl.-Based Syst.*, vol. 83, pp. 81–91, Jul. 2015.
[40] L. Yu and H. Liu, “Efficient feature selection via analysis of relevance and redundancy,” *J. Mach. Learn. Res.*, vol. 5, no. 10, pp. 1205–1224, Dec. 2004.
[41] X. Hong and C. D. Nugent, “Segmenting sensor data for activity monitoring in smart environments,” *Pers. Ubiquitous Comput.*, vol. 17, no. 3, pp. 545–559, Mar. 2013.
[42] O. Banos, J.-M. Galvez, M. Damas, H. Pomares, and I. Rojas, “Window size impact in human activity recognition,” *Sensors*, vol. 14, no. 4, pp. 6474–6499, Apr. 2014.
[43] D. Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes-Ortiz, “A public domain dataset for human activity recognition using smartphones,” in *Proc. 21st ESANN*, Bruges, Belgium, 2013, pp. 437–442.
[44] M. Arif, M. Bilal, A. Kattan, and S. I. Ahamed, “Better physical activity classification using smartphone acceleration sensor,” *J. Med. Syst.*, vol. 38, no. 95, pp. 1–10, Jul. 2014.
[45] M. Robnik-Šikonja and I. Kononenko, “Theoretical and empirical analysis of ReliefF and RReliefF,” *Mach. Learn.*, vol. 53, nos. 1–2, pp. 23–69, Oct. 2003.
[46] M. Ringnér, “What is principal component analysis?” *Nature Biotechnol.*, vol. 26, no. 3, pp. 303–304, 2008.
[47] A. Wang, N. An, G. Chen, L. Li, and G. Alterovitz, “Predicting hypertension without measurement: A non-invasive, questionnaire-based approach,” *Expert Syst. Appl.*, vol. 42, no. 21, pp. 7601–7609, Nov. 2015.

Aiguo Wang received the B.S. degree from the Hefei University of Technology, China, in 2006.

He is currently pursuing the Ph.D. degree with the School of Computer and Information, Hefei University of Technology. He is also an Assistant with School of Computer and Information Engineering, Chuzhou University, Anhui, China. His current research interests include data mining and its applications in healthcare, assisted living systems, and activity recognition.

Guilin Chen received the B.S. degree from Anhui Normal University, China, in 1985, and the M.S. degree from the Hefei University of Technology in 2007. He is currently a Professor with the School of Computer and Information Engineering, Chuzhou University, Anhui, China. His current research interests include cloud computing, wireless networks, healthcare, and Internet of Things.

Jing Yang received the B.S. and Ph.D. degrees from the Hefei University of Technology, China, in 2004 and 2013, respectively.

She is currently an Associate Professor with the School of Computer and Information Engineering, Hefei University of Technology, China. Her current research interests include artificial intelligence, data mining, and bioinformatics.

Shenghui Zhao received the M.S. degree from the Hefei University of Technology, China, in 2003, and the Ph.D. degree from Southeast University, China, in 2013.

She is currently a Professor with the School of Computer and Information Engineering, Chuzhou University, Anhui, China. Her current research interests include trusted computing, wireless networks, healthcare, and Internet of Things.

Chih-Yung Chang (M'05) received the Ph.D. degree in computer science and information engineering from National Central University, Zhongli, Taiwan, in 1995.

He is currently a Full Professor with the Department of Computer Science and Information Engineering, Tamkang University, Taipei, Taiwan. His current research interests include Internet of Things, wireless sensor networks, ad hoc wireless networks, and long-term evolution broadband technologies. He served as an Associate Guest Editor of several SCI-indexed journals, including the *International Journal of Ad Hoc and Ubiquitous Computing* from 2011 to 2014, the *International Journal of Distributed Sensor Networks* from 2012 to 2014, *IET Communications* in 2011, *Telecommunication Systems* in 2010, the *Journal of Information Science and Engineering* in 2008, and the *Journal of Internet Technology* from 2004 to 2008.

Dr. Chang was an Area Chair of the IEEE International Conference on Advanced Information and Applications in 2005, the Taiwan Academic Network Conference in 2000 and 2010, the IEEE International Symposium on Wireless IP in 2005, IFIP International Conference on Embedded and Ubiquitous Computing in 2005, the IEEE International Conference on Information Technology: Research and Education in 2005, and the IEEE AINA in 2008, the Program Co-Chair of the IEEE International Workshop on Multimedia Network Systems and Applications in 2005, the Mobile and Ubiquitous Technologies Enhanced Learning Conference in 2006, the Workshop on Wireless, Ad Hoc, and Sensor Networks in 2007, the ACM International Workshop on Sensor, Ad Hoc, and Mesh Networks in 2008, the IEEE International Workshop on Ad Hoc and Ubiquitous Computing in 2008, and the International Conference on University Basic Computers Education and E-Learning in 2010 and 2011, Workshop Co-Chair of the International Workshop on Mobile Systems, E-Commerce and Agent Technology in 2003 and 2004, the IEEE International Workshop on Information Networking and Applications in 2005, the International Computer Symposium in 2008, the National Computer Symposium in 2009, and the IEEE International Workshop on Vehicular Communications, Networks, and Applications in 2009, and the Publication Chair of MSEAT 2005 and the International Conference on Sharable Content Object Reference Model 2006.