24th International Conference on Knowledge-Based and Intelligent Information & Engineering Systems

Optimal Feature Set for Smartphone-based Activity Recognition

Maryam Banitalebi Dehkordi[a,*], Abolfazl Zaraki[a], Rossitza Setchi[a]

[a]Centre for Artificial Intelligence, Robotics and Human-Machine Systems (IROHMS), Cardiff University, Cardiff CF24 3AA, UK

Abstract

Human activity recognition using wearable and mobile devices is used for decades to monitor humans' daily behaviours. In recent years as smartphones being widely integrated into our daily lives, the use of smartphone's built-in sensors in human activity recognition has been receiving more attention, in which smartphone accelerometer plays the main role. However, in comparison to the standard machine, when developing human activity recognition using a smartphone, the limitations such as processing capability and energy consumption should be taken into consideration, and therefore, a trade-off between performance and computational complexity should be considered. In this paper, we shed light on the importance of feature selection and its impact on simplifying the activity classification process, which enhances the computational complexity of the system. The novelty of this work is related to identifying the most efficient features for the detection of each individual activity uniquely. In an experimental study with human users and using different smartphones, we investigated how to achieve an optimal feature set, using which the system complexity can be decreased while the activity recognition accuracy remains high. For that, in the considered scenario, we instructed the participants to perform different activities, including static, dynamic, going up and down the stairs, and walking fast and slow while freely holding a smartphone in their hands. To evaluate the obtained optimal feature set implementing two major classification algorithms, the decision tree and the Bayesian network, we investigated activity recognition accuracy for different activities. We further evaluated the optimal feature set by comparing the performance of the activity recognition system using the optimal feature set and three feature sets taken from the state-of-the-art. The experimental results demonstrated that replacing a large number of conventional features with an optimal feature set has only a negligible impact on the overall activity recognition system performance while it can significantly decrease the system's complexity, which is essential for smartphone-based systems.

Keywords: Activity recognition, feature extraction, feature selection, inertial sensor, smartphone;

1. Introduction

Human activity recognition (HAR) methods are used for decades to monitor humans' daily behaviours. Nowadays, HAR plays a significant role in different areas such as elderly care, smart building, assistance for people with cognitive disorders, ubiquitous computing, context-aware data acquisition for clinical purposes, assistive technologies, healthcare services, and applications (e.g., human activity monitoring for health assessment), rehabilitation, and

* Maryam Banitalebi Dehkordi. Tel.: +44-292-087-4278
  E-mail address: banitalebidehkordim@cardiff.ac.uk

sports applications [19, 10]. In general, HAR is a system that receives the collected sensory information associated with human body motions while performing different activities as input. It afterwards manages and analyses the data flow, and as output returns the activity class. Therefore, the three common phases for HAR systems can be considered as data acquisition, feature extraction, and activity classification. The most widely used data acquisition techniques for HAR are employing inertial sensors [4], vision-based motion capture sensors, wearable sensors [7], etc. Nowadays, smartphones are widely available in our lives and have played the prominent role in providing solutions for the HAR problem. Smartphones are easy to use and enable continuous and long-term human activity monitoring. Therefore, in this work, we consider acquiring data from built-in smartphone sensors.

Acquiring the data, they are further processed to extract low-level feature descriptors which can represent the performed activities. Researchers have been implementing several classification methods to distinguish the human activities based on the acquired data, such as deep learning, neural network, decision trees, Hidden Markov Model (HMM), k-nearest neighbours, Support Vector Machine (SVM), naïve Bayes, Gaussian mixture models, etc. [1, 3]. In the context of HAR, several complex systems with high accuracy are developed; However, when developing smartphone-based systems, due to the limitations of processing capability and energy consumption of smartphones compared to standard machines, a trade-off between performance and computational complexity must be considered. When developing a HAR system, the efficiency of the HAR system depends in addition to the type of the classifier, on the complexity of the feature extraction process, and the number of selected features. As reported in [13] increasing the number of features and including all features do not necessarily improve HAR accuracy. Therefore, in this work, we shed light on the importance of feature selection, in which by obtaining an optimal feature set (that can be employed instead of a group of features), the complexity of the HAR system is reduced significantly. To investigate it, in our scenario, we look for features, which are essential cues for distinguishing different activities. Feature selection is considered a popular problem in machine learning and data mining. Feature selection is a pre-processing technique that analysing the significant features gives an understanding of the HAR problem by identifying the discriminative features for that given problem. The goal of implementing feature selection is to find the most important and optimal features that can be resulted in a better classifier with higher accuracy, which can also decrease the system computational overload [8].

The number of features is important particularly when we develop a smartphone-based HAR system since the computation procedure affects the smartphone battery consumption as well as the real-time performance. In our previous work [5] using smartphone-based sensors, through feature evaluation and feature selection, we found an optimal feature that can solely detect static versus dynamic activities. In this paper, we introduce a few number of features called *optimal feature set* that are able to distinguish different activities with high accuracy. In an experimental study, we collected the labelled data of 10 participants and using two different smartphones, where users were able to freely hold a smartphone in their hands and were instructed to perform various activities, including static, dynamic, walking up and down the stairs, and walking fast and slow. Implementing the feature selection technique, we identified an optimal feature set and using classification algorithms including decision tree and Bayesian network, we investigated the trade-off between HAR complexity and overall performance. Finally, through an experimental evaluation, we demonstrated that by replacing the conventional feature set with an optimal set, the complexity and computational overload of the HAR system can be significantly decreased, while it has only a negligible impact on HAR accuracy. The obtained results show that the developed HAR system distinguished different activities with the weighted average accuracy of 92.44%, with some activities (i.e., dynamic activities) recognition rate reaching values as high as 99.7%.

2. Feature Analysis in Activity Recognition

Among humans Activities of Daily Living (ADL), we consider five commonly performed activities that humans usually do while carrying their smartphones. Figure 1 shows these activities, where dynamic activities indicate the motions in which human has a notable displacement in the global coordinates, and static activities indicate that the user has no/negligible displacement in the global coordinates. We assume that users have their smartphones freely in their hands while performing different pedestrians' activities. When using body-mounted sensors, or when the smartphone is in the user's pocket, the sensor has fewer degrees of freedom than when the user holds it in a freely moving hand, where the motion and consequently the associated signal behaviour become more complicated.

The considered static and dynamic activities are as follows: *Standing (ST)*: this state includes the moments that the user is almost still but not quite motionless with a freely moving hand. In addition, we considered the slight movements
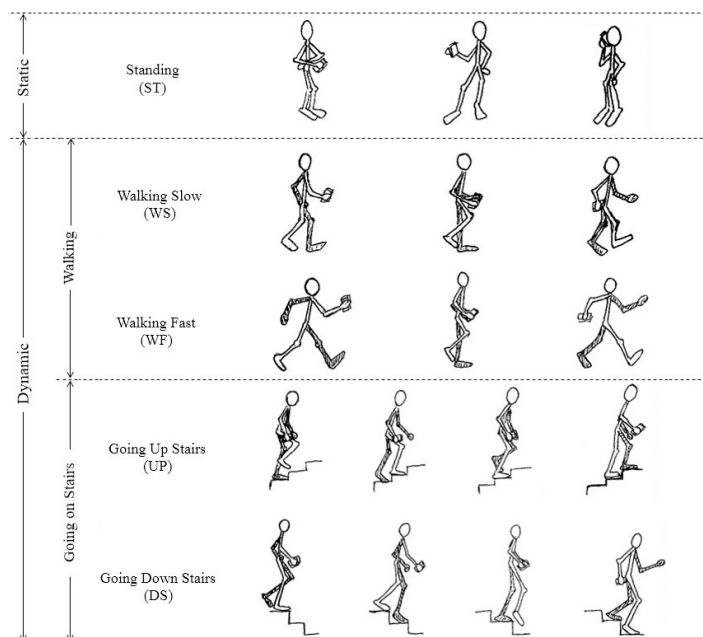
Fig. 1. The human body postures in different motion states.

of the user when normally stays about the point (e.g., when the user remains at the point and talks to someone) as a standing state. *Walking Fast (WF)*: during freely fast walking, the strides are longer, the foot hits the floor intensely, and the arms swing remarkably. *Walking Slow (WS)*: in freely slow-walking, the user's strides are shorter compare to fast walking; at every step, the foot hits the floor more gently, and the arms swing normally. *Going up the stairs (US)*: the user displacement is in both horizontal and vertical planes. We consider the user free motions on the stairs (with desired speed and intensity), where in general, the user's hand has more intensive motions compare to walking. *Going down the stairs (DS)*: it is considered as going down the stairs state, with free movements of the user during the activity.

To perform an in-depth evaluation of features, firstly, we consider the *conventional features* that are widely used in the state-of-the-art (details are reported in our previous work [5]). Afterwards, we precisely observe the behaviour of the signals associated with each activity and extract a new set of features. The main feature extraction strategy arises from the fact that signals relevant to different activities show different behaviours in the sense of sharpness, periodicity, peakedness, similarity to the known functions (e.g., sinusoidal and polynomial), etc. Following this notation, considering the signal characteristics in frequency and time domains, we extract new features employing signal processing methods such as curve fitting, binning, Discrete Fourier Transform (DFT), signal predictions, parametric and non-parametric features from zero-crossings, etc. For example, we fit the signals associated with *walking fast* activity to sinusoidal function and consider the fitting's parameters as features.

Assume $a^l = \left[a^l_x, a^l_y, a^l_z\right]$ is the measured acceleration in smartphone local coordinates, a_n is the norm of acceleration, $a = \left[a_x, a_y, a_z\right]$ is the acceleration in the global coordinates (where global coordinates are parallel to the main walls of a building and the vertical acceleration a_z is along gravity), and $a^h_n = \sqrt{a_x^2 + a_y^2}$ is the acceleration in the horizontal plane (x-y plane of global coordinates). We take into account the following features, and among them, we select the most efficient ones.

Cumulant, and Moments: The features such as variance, mean, kurtosis, and skewness have been previously taken into account. The difference in peakedness and sharpness of acceleration signals obtained from distinctive activities motivate us to generalize the idea, calculate cumulants, *central* and *standardized moments* up to 6th orders (3rd to 6th orders) for each acceleration signal (i.e., a^l, a, a_n). These features seem promising for distinguishing walking in the horizontal plane from going on the stairs.

Extra Zero-Crossing Features: Zero-crossings indicate a change of acceleration signs. We assume that despite the difference between the numbers of zero-crossing points, the intervals between them can provide distinctive features for

different activities. Therefore, in this work, new parametric and non-parametric features are extracted from the events of zero-crossings, zero-crossing up (signal passes from negative to positive), and down (signal passes from positive to negative). These features are *mean, variance, kurtosis, median* and *skewness* of the intervals between zero-crossing points. Besides, the histogram of the intervals is obtained (over seven bins) and is considered as an extra feature.

Residuals of Matched Filter: The acceleration signal exhibits a periodic and sinusoidal pattern for some activities (e.g., going up the stairs). To take advantage of this characteristic, the matched filter is employed. The acceleration is convolved with the sinusoidal signal, and the residual is used as a feature.

Signal Predictions: The coefficients of signal predictions using different methods are used as features. These methods are linear prediction, ARMA modelling (Pronys method and Steiglitz-McBride method), smoothing, Fourier, and polynomial functions. In addition, the residuals obtained from each method are further analysed, and their summation, mean, and variance are evaluated as features.

Curve Fitting Functions: Considering several curve fitting functions, we find out that acceleration signals obtained from each activity can be best fitted by one of the functions. In this work, several curve fitting functions including *sum of sines*, *Gaussian*, *polynomial*, *Fourier* and *power* functions of 2nd order, are taken into consideration and are fitted to different acceleration signals (a^l, a, a_n). The 2nd order linear polynomial model can be written as

$$f(x) = p_1 x^2 + p_2 x + p_3 \qquad (1)$$

where p_1, p_2 and p_3 are model coefficients, which are considered as features.

The general model of 2nd order power function is

$$f(x) = ax^b + c \qquad (2)$$

where a, b and c are coefficients that are considered as features.

The combination of two Gaussian functions can be written as

$$f(x) = a_1 e^{\frac{-(x-b_1)^2}{c_1}} + a_2 e^{\frac{-(x-b_2)^2}{c_2}} \qquad (3)$$

where a_i, b_i and c_i are coefficients which are considered as features.

The general model of 2nd order sum of sines fitting, can be written as

$$f(x) = a_1 sin(b_1 x + c_1) + a_2 sin(b_2 x + c_2) \qquad (4)$$

where a_i, b_i and c_i are coefficients which are considered as features.

The general model of 2nd order Fourier function for curve fitting is written as

$$f(x) = a_0 + a_1 cos(xw) + b_1 sin(xw) + a_2 cos(2xw) + b_2 sin(2xw) \qquad (5)$$

where a_0, a_1, a_2, b_1, b_2 and w are coefficients which are considered as features.

Implementing these curve fitting functions, the features are defined as fitting coefficients and goodness of fit parameters, which are the Sum of Squared Errors of prediction (SSE), R-Square ($R^2 = 1 - \frac{residual\ sum\ of\ squares}{total\ sum\ of\ squares}$), Adjusted R^2 and Root Mean Square Error (RMSE).

3. Data Collection and Preprocessing

We collected the data of 10 participants of both genders in their twenties, with different height and body constitution while holding a smartphone (either Samsung Galaxy or LG Nexus) and performing different required activities (e.g., walking, and going on the stairs), during a free walk having a smartphone in their left/right hands. The participants were asked to walk inside/outside of multi-floor buildings at the Technical University of Munich for 10 minutes, and no further constraints were imposed so that they could follow any arbitrary path with their normal speed and pattern of walking. Similar to [11] we manually labelled the participant's current activity to record ground truth data, where using a graphical user interface of an app, the participant was able to unobtrusively select the current activity. We recorded the measurements from all of the smartphone embedded sensors, however similar to [15] we focus on the features extracted from triaxial accelerometers since accelerometers are available as built-in sensors in the majority of today's smart devices. Figure 2 illustrates the sample data of accelerometer recorded during standing and
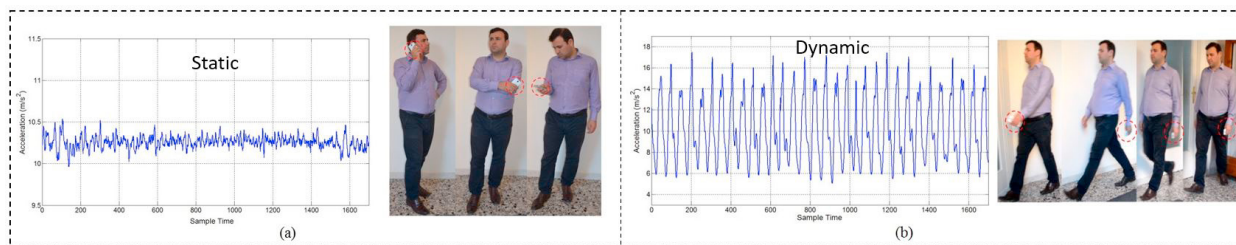
Fig. 2. Acceleration signals associated to different activities. (a) Standing (b) Walking Fast.
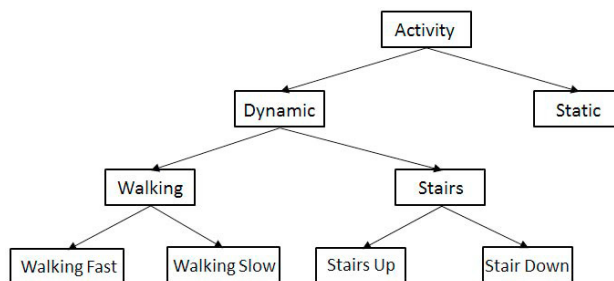

Fig. 3. Hierarchical classification for activity recognition.

walking fast activities. In this work, we focus on system efficiency and for that, one parameter to be considered is the sampling rate for recording the acceleration signal. Working at lower frequencies requires less computation and correspondingly less energy consumption, e.g., [4] reports the sampling rate of 150 Hz as a trade-off for detection of different phases of pedestrian walking. We recorded the acceleration signal with a sampling rate of 50 Hz, which is obtained experimentally. We down-sampled the acquired data down to 20 Hz; however, it resulted in difficulties distinguishing different activities.

The fixed-length intervals of acceleration signals are required for obtaining the statistical features. Considering the fact that step periods typically lie below 1 second, we determined the two-second intervals to support two periods of periodic activity. Therefore, we partitioned the recorded acceleration signals (e.g., a^l, a, and a_n) over two-second segments with 75% overlapping (to minimise the loss). The signal processing and feature extraction are implemented using Matlab programing language, which resulted in the extraction of above 600 features, including conventional features and the features introduced in section 2 (for more details see [5]). To identify the optimal set for HAR, all of the extracted labelled features should be evaluated using classifiers. For that reason, different classification algorithms are implemented, including decision trees and Bayesian networks. We employed the supervised attribute filter [6] as a pre-processing method to perform feature selection for classifying activities.

4. Experimental Evaluation and Feature Selection

In this work, the activities are detected using a hierarchical classifier [18, 14]. Figure 3 shows the different steps of the classifying procedure. As can be seen, firstly, dynamic activities are distinguished from static activities. Afterwards, the dynamic activities are sub-classified to detect walking versus going on the stairs, which respectively are sub-classified as walking fast and slow and going up and down the stairs.

The efficiency and performance of the HAR system are highly dependent on the proper feature selection. To apply feature selection techniques for detecting the most discriminative features, we applied a combined approach of forward-backwards feature selection (which searches through the space of feature subsets) using a hill-climbing algorithm [17, 9], which is resulted in the best performance in this work and on our feature set. In order to further evaluate the resulting feature sets and select the optimal feature set, implementing classifiers, we evaluated the performance of the HAR system using each feature set. Classifiers were trained using the recorded labelled data, where following [2] 10-fold cross-validation was performed. Besides, the system's accuracy was determined as the accuracy in recognising each activity and the weighted accuracy (that takes into account the number of instances that exist for each activity,

which is important when the number of instances is not equal for different activities). As a result of classification, the feature sets that distinguished activities with the best performance were selected as optimal feature sets for recognising different activities.

Amongst the classification algorithms, we used a decision tree as a popular classifier following [16] that used it for activity recognition based on a hand-held IMU, and naïve Bayes classifier following [3]. Decision trees are fast in reasoning and capable of processing both numerical and categorical data. Compare to other algorithms, decision trees are less computationally expensive due to the fact that they can be evaluated in $O(log\ n)$ for n attributes. In addition, it has the ability to deal with noisy or incomplete data. A decision tree has a tree structure where at each traversed node, one feature is investigated, and each tree's leaf corresponds to one class label. Therefore, the features at the top of the tree are required to be the most relevant ones where their ranking is usually performed based on information gain [12].

5. Result

As a result of the experimental feature evaluation presented in section 4, the set of the most efficient features for recognising each activity have been identified. The obtained features for different activities are detailed as follows.

5.1. Static vs. Dynamic Motions

In our previous work [5] we reported the top three features amongst conventional features that can recognise the static versus dynamic motions with high accuracy. In this work, considering the features introduced in 2 we identified 11 features out of all evaluated features. Using each of these features solely, the HAR system can recognise the static motions versus dynamic motions with high accuracy. Table 1 reports the feature names together with the corresponding weighted accuracy using 10-fold cross-validation. Among the features that resulted in the highest accuracy (99.2%), we selected the *range of acceleration along z axis* (RangeAccZ) as the first component of the optimal feature set while it can be obtained with less computation compared to other winner features. This feature is obtained as the difference of the maximum and minimum values of the acceleration signal along the z-axis over each segmentation. Therefore, the first selected feature can be written as

$$f_1 = [RangeAccZ] \tag{6}$$

Table 1. Selected features for distinguishing static and dynamic motions.

Feature	GausSSEAccY*	RangeAccZ	PwRMSEAccY/Z	PlyRMSEAccY/N	FrRMSEAccN	VarAccHrz	RangeAccVert
Acc.(%)	99.2	**99.2**	99.1	99.1	99.1	99.0	98.9

* The abbreviations are defined in Appendix A.

5.2. Walking vs. Stairs

Distinguishing walking versus going on the stairs activities using only accelerometer-based features is challenging, while the collected signals have slightly different behaviour. We identified that no feature could solely distinguish walking versus going on the stairs. Therefore, the feature selection process resulted in five different combinations of features. These feature sets are reported in Table 2. As can be seen, the feature set resulted in the best accuracy (91.3%)consists of four features: (i) mean of vertical acceleration (MeanAccVr); (ii) ratio of the mean of vertical over the mean of horizontal acceleration (RatVHAcc); (iii) coefficient b_1 obtained from Sum of sine fitting to acceleration along the x-axis (b1SSAccX); and (iv) coefficient w obtained from Fourier fitting to acceleration along the z-axis (wFAccZ). However, to identify the optimal feature set, we selected the set that returned the second high accuracy, as it can be obtained with less computation and yet results in an acceptable high accuracy. The features of the selected set are: (i) mean of vertical acceleration (MeanAccVr); (ii) variance of vertical acceleration (VarAccVr); (iii) 3rd central moment of acceleration along y (Mu3Y); (iv) 5th bin of zero-crossing histogram along z-axis (BinZ5ZC); (v) variance of acceleration along y (VarAccY). Therefore, the second selected feature set can be written as

$$f_2 = [MeanAccVr, VarAccVr, Mu3Y, BinZ5ZC, VarAccY] \tag{7}$$

5.3. Going Up vs. Going Down the Stairs

Feature selection for distinguishing going up versus going down the stairs resulted in six winner feature sets, which are shown in Table 3. As can be seen, the first feature set number resulted in the highest accuracy (92.7%); however,

Table 2. Selected features for distinguishing walking and going on the stairs.

No.	Feature Name	Acc.(%)
1	1. MeanAccVr, 2. RatVHAcc, 3. b1SSineAccX, 4. wFrAccZ	**91.3**
2	1. MeanAccVr, 2. VarAccVr, 3. Mu3Y, 4. BinZCZ5, 5. VarAccY	<u>88.9</u>
3	1. MeanAccVr, 2. VarAccVr, 3. BinZCZ5, 4. VarAccY	88.7
4	1. MeanAccVr, 2. VarAccVr, 3. BinZCZ5, 4. VarAccY, 5. BinAccZ5, 6. BinAccN1, 7. BinAccN4	87.5
5	1. BinAccZ1, 2. BinAccZ4, 3. BinAccZ5, 4. BinAccN2	81.7

considering the computational cost, we have selected the fourth feature set consists of the following features: (i) ratio of the mean of vertical over the mean of horizontal acceleration (RatVHAcc); (ii) skewness of acceleration along the x-axis (SkewAccX); (iii) the 3rd central moment of acceleration along the x-axis (Mu3X); (iv) number of zero-crossings of acceleration along the x-axis (NoZCAccX); (v) energy of acceleration along the x-axis (EngAccX); (vi) mean of acceleration along the x-axis (MeanAccX), and (vii) mean of acceleration norm (MeanAccN). Therefore, the third selected feature set can be obtained as

$$f_3 = [RatVHAcc, SkewAccX, Mu3X, NoZCAccX, EngAccX, MeanAccX, MeanAccN] \tag{8}$$

As shown in Table 3, the feature sets in the first three rows resulted in higher accuracies compared to the selected feature set. However, due to the existence of the curve fitting and frequency domain features, we have selected f_3, which can be computed with less complexity.

Table 3. Selected features for distinguishing going up and down the stairs.

No.	Feature Name	Acc(%)
1	1.RatVHAcc,2.SSineRSquAccY,3SkewAccX,4NoZCAcX,5a3PrAcVer,6BinZCY5,7BinZCY7,8BinZCZ5,9BinZCN6,10.DFTAccY6	92.7
2	1.RatVHAcc,2.SkewAccX,3.Mu3X,4NoZCAcX,5.EngAcX,6.MeanAcX,7.MeanAcN,8.DFTAccY6,9.DFTAccZ8,10.DFTAccN6	90.7
3	1.SSineRSquAccY, 2.NoZCAccX, 3.BinZCY7, 4.BinZCN6, 5.DFTAccY6	88.6
4	1.RatVHAcc, 2.SkewAccX, 3.Mu3X, 4.NoZCAccX, 5.EngAccX, 6.MeanAccX,7.MeanAccN	<u>87.2</u>
5	1.NoZCAccX, 2.EngAccX, 3.MeanAccX	85.6
6	1.NoZCAccX, 2.BinZCY7, 3.BinZCN6	85.0

5.4. Walking Fast vs Walking Slow

During walking fast motions, as shown in Figure 1, the user's hand swings more intensively compare to walking slow activity. These intensive swings can be collected using gyroscopes, and therefore employing the data that are solely collected by accelerometer imposes challenges on distinguishing these two activities.

Following the feature selection process, we obtained different feature sets for distinguishing walking fast versus slow, which are reported in Table 4. As shown, the first feature set resulted in the highest accuracy (90.1%); however, compared to other feature sets, it has a higher number of features and contains curve fitting and frequency domain features. Therefore, taking into consideration the computational cost of obtaining these features, we selected the feature set that resulted in the second high accuracy (recognition accuracy of 88%), which consists: (i) the range of maximum minus minimum value of the vertical acceleration (RangeAccVert); (ii) the 4th central moment of vertical acceleration (Mu4Vert); and (iii) 7th DFT coefficient of acceleration along the z-axis (DFTAccZ7). Therefore, the fourth selected feature set can be written as

$$f_4 = [RangeAccVert, Mu4Vert, DFTAccZ7] \tag{9}$$

Following this notation, the optimal feature set for activity recognition is obtained as follows

$$F = [f_1, f_2, f_3, f_4] \tag{10}$$

It should be noted that the optimal feature set F contains both conventional features and the features introduced in 2, where *Mu3Y*, *BinZ5ZC*, *RatVHAcc*, *Mu3X*, *RangeAccVert*, and *Mu4Vert* are among introduced features and the

Table 4. Selected features for distinguishing walking fast and walking slow.

No.	Feature Name	Acc(%)
1	1.RangeAccVert,2.SSineSSEAccZ,3SSineSSEAccN,4PwSSEAccZ,5PwRMSEAccZ,6.FrRMSEAccY,7.Mu4Vert,8.DFTAccZ7	**90.1**
2	1.RangeAccVert, 2.Mu4Vert, 3.DFTAccZ7	88.0
3	1.RangeAccVert	86.3
4	1.Mu4Vert	84.1

Table 5. Acceleration-based features taken from the state-of-the-art.

Name	Features
Group I [13]	1. Mean value of acceleration along x,y, z, norm of acceleration, 2.Variance of acceleration along x, y,z, and norm of acceleration, 3. Mean value and variance of the horizontal acceleration and the vertical acceleration minus gravity acceleration, 4.Variance and mean value of the dynamic acceleration, the dynamic acceleration in vertical and horizontal planes of the global coordinates, 5. Amplitude of 1st and 2nd dominant frequencies of the acceleration, 6. 1st and 2nd dominant frequencies of the acceleration, 7. Amplitude scale and difference of two dominant frequencies.
Group II [16]	1. Accelerometer energy, 2. Variance of acceleration, 3. Amplitude of 1st and 2nd dominant frequencies of the acceleration, 4. Dominant frequencies of the acceleration.
Conventional	1. DFT coefficients, 2. Amplitude of 1st and 2nd dominant frequencies of the acceleration, 3. 1st and 2nd dominant frequencies of the acceleration, 4. Spectral entropy spectral energy, 5. Variance of acceleration along x, y, z, and norm of acceleration, 6. Mean value of the horizontal acceleration and the vertical linear acceleration, 7. Mean value of acceleration along x, y, z, and norm of acceleration, 8. Signal magnitude area, 9. Percentile and interquartile range, 10. Accelerometer energy, 11. Binned distribution and cumulative histogram, 12. Peak counting, amplitude and time interval between peaks, 13. Zero-crossing rate, 14. Variance and mean value of the dynamic acceleration in the vertical and horizontal plane of global coordinates.

remainders are among the conventional features. Considering each activity's recognition accuracy using the feature set F, we computed the weighted average recognition rate of our HAR system in distinguishing different activities as 92.44%. The weighted average accuracy is computed considering the number of instances for each activity and the accuracy of distinguishing that activity.

6. Discussion

This section presents the further evaluation of the obtained optimal feature set by comparing the performance of the HAR system using the optimal set and using three different groups of features from the state-of-the-art, implementing a decision tree and naïve Bayes classifiers. We employed two accelerometer-based groups of features presented in [16] and [13] to provide practical evaluation and comparison. Furthermore, we used the combination of conventional features that are selected amongst the most commonly used acceleration-based features as reported in [5]. Table 5 shows the features of all three groups.

The results of the evaluation and comparison of the HAR performance for optimal feature sets, using a decision tree and naïve Bayes classifiers, are illustrated in Figure 4. As can be seen in Figure 4(a), in recognition of static versus dynamic activities, there is no statistical significance on the HAR system performance using these groups or the optimal feature set. The same results are obtained for recognition of walking versus going on the stairs, as illustrated in Figure 4(b). As shown, using naïve Bayes classifier, the optimal feature set resulted in even a higher HAR performance. Figure 4(c) shows the results of walking slow versus walking fast recognition. As shown, the lowest HAR performance using optimal feature set compared to other three groups is obtained in recognition of walking fast using decision tree; However, this optimal feature set resulted in the highest accuracy for walking fast using naïve Bayes classifier. Considering the trade-off between the accuracy and system efficiency and taking into account the negligible difference in recognition of walking slow, we prioritised using the optimal feature set instead of using a large number of computationally expensive features. Finally, Figure 4(d) shows the results of distinguishing going up versus going down the stairs, where recognition using the optimal feature set employing naïve Bayes resulted in the highest performance and employing decision tree resulted in performance slightly lower than the best accuracy.

In this study, we obtained an optimal feature set for activity recognition. As demonstrated in recognising all activities, there is no statistical significance in the HAR system performance using the optimal feature set and the best results obtained from the other three groups. In other words, using the optimal feature set has only a negligible impact on the overall HAR performance in recognition of all activities. This performance is obtained using one feature solely (e.g., for detecting static vs dynamic) or a small group consisting of a few features (e.g., three features for recognising
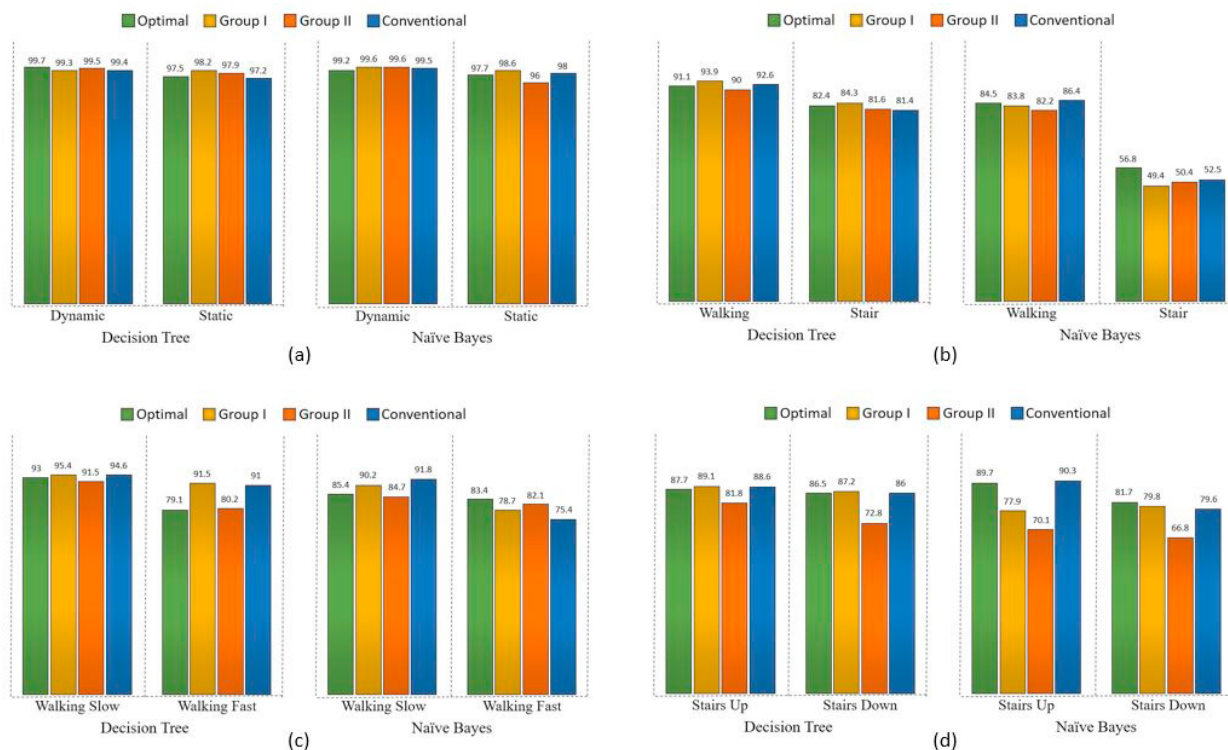
Fig. 4. Evaluation of features (in percentage %) for recognition of (a) static versus dynamic activities, (b) walking versus going on the stairs, (c) walking fast versus walking slow, (d) going up versus going down the stairs, using decision tree and naïve Bayes.

walking fast vs slow) instead of a group that contains a large number of features. Therefore, using an optimal feature set significantly reduced the complexity of the HAR system.

7. Conclusion

This paper focused on the significance of feature selection for human activity recognition and its impact on system complexity and performance, which is highly important in developing smartphone-based systems. Evaluating a trade-off between the activity recognition system performance and the relevant computational complexity, we acquired an optimal feature set that can distinguish different activities using only a few number of features. To provide a fair evaluation, in an experimental study, we recorded a data set with several users and different smartphones, where users were instructed to perform different activities including static, dynamic, walking fast and slow, and going up and down the stairs, while freely holding a smartphone in their hands. We implemented decision tree and Bayesian network classifiers to compare the activity recognition performance using the obtained optimal feature set to three groups of features obtained from the state-of-the-art. Through an experimental evaluation, we demonstrated that the conventional feature set can be successfully replaced with the proposed optimal feature set, where it can significantly decrease the complexity of the system with only a negligible impact on the overall activity recognition performance. In the future development, we aim to consider the performance of our current setup with other embedded sensors, for example, accelerometer and gyroscope of Apple Watch or Xsens 3D motion tracking system.

Appendix A. Abbreviation Definition

MeanAccX/MeanAccN: mean of acceleration along x-axis and mean of acceleration norm, respectively.
MeanAccVr/VarAccVr: Respectively, mean and variance of vertical acceleration.
VarAccHrz/VarAccY: variance of horizontal acceleration and variance of acceleration along y, respectively.
RangeAccVert: the range of the max. minus the min. value of the acceleration in the vertical plane over the segmentation.
RangeAccZ: the range of the maximum minus the minimum value of the acceleration along z-axis over the segmentation.
EngAccX: energy of acceleration along x-axis.
RatVHAcc: ratio of the mean of acceleration in vertical plane over the mean of acceleration in horizontal plane.
Mu3X/Y: 3rd central moment of acceleration along x-axis, and y-axis, respectively.
Mu4Vert: 4th central moment of acceleration in vertical plane.
SkewAccX: skewness of acceleration along x-axis.
PwSSEAccZ: Sum of Squared Errors (SSE) obtained from power function fittings to acceleration along z-axis.
PwRMSEAccY/Z: Root-Mean-Square Error (RMSE) obtained from power function fittings to acceleration along y and z-axis, respectively.
FrRMSEAccY/N: Root-Mean-Square Error (RMSE) obtained from Fourier function fittings to acceleration along y-axis, and acceleration norm, respectively.
wFrAccZ: coefficient w obtained from Fourier fitting to acceleration along z-axis.
PlyRMSEAccY/N: Root-Mean-Square Error (RMSE) obtained from linear polynomial function fittings to accel. along y-axis, and acceleration norm, respectively.
SSineRSquAccY: R^2 obtained from sum of sines fittings to acceleration along y-axis.
b1SSineAccX: coefficient b_1 obtained from Sum of sine fitting to acceleration along x-axis.
SSineSSEAccZ/N: Sum of Squared Errors (SSE) obtained from sum of sines fittings to acceleration along z-axis, and acceleration norm, respectively.
GausSSEAccY: Sum of Squared Errors (SSE) obtained from Gaussian fittings to acceleration along y-axis.
DFTAccN6/DFTAccY6: 6th DFT coefficient of acceleration norm, and acceleration y axis, respectively.
DFTAccZ7/8: Respectively, 7th, and 8th DFT coefficient of acceleration along z axis.
NoZCAccX: number of zero-crossing of acceleration along x-axis.
BinZCY5/7: Respectively, 5th, and 7th bin of zero-crossing histogram along y-axis.
BinZCZ5/BinZCN6: 5th bin of zero-crossing histogram along z-axis, and 6th bin of zero-crossing histogram of acceleration norm, respectively.
BinAccZ1/4/5: Respectively, 1st, 4th, and 5th bin of histogram of acceleration along z-axis.
BinAccN1/2/4: Respectively, 1st, 2nd, and 4th bin of histogram of acceleration norm.

References

[1] Akbari, A., Jafari, R., 2020. Personalizing activity recognition models with quantifying different types of uncertainty using wearable sensors. IEEE Transactions on Biomedical Engineering , 2530–2541.

[2] Banos, O., Galvez, J.M., Damas, M., Pomares, H., Rojas, I., 2014. Window size impact in human activity recognition. Sensors 14, 6474–6499.

[3] Capela, N.A., Lemaire, E.D., Baddour, N., 2015. Feature selection for wearable smartphone-based human activity recognition with able bodied, elderly, and stroke patients. PLOS ONE , 1–18.

[4] Dehkordi, M.B., Frisoli, A., Sotgiu, E., Loconsole, C., 2014. Pedestrian indoor navigation system using inertial measurement unit. Sensor Netw Data Commun 3, 1–9.

[5] Dehkordi, M.B., Zaraki, A., Setchi, R., 2020. Feature extraction and feature selection in smartphone-based activity recognition. Procedia Computer Science 176, 2655–2664.

[6] Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I., 2009. The weka data mining software: an update. ACM SIGKDD Explorations Newsletter 11, 10–18.

[7] Kerdjidj, O., Ramzan, N., Ghanem, K., Amira, A., Chouireb, F., 2020. Fall detection and human activity classification using wearable sensors and compressed sensing. Journal of Ambient Intelligence and Humanized Computing 11, 349–361.

[8] Khaire, U.M., Dhanalakshmi, R., 2019. Stability of feature selection algorithm: A review. Journal of King Saud University-Computer and Information Sciences , 1–14.

[9] Liu, H., Motoda, H., 2012. Feature selection for knowledge discovery and data mining. volume 454. Springer Science & Business Media.

[10] Loconsole, C., Dehkordi, M.B., Sotgiu, E., Fontana, M., Bergamasco, M., Frisoli, A., 2016. An imu and rfid-based navigation system providing vibrotactile feedback for visually impaired people, in: International Conference on Human Haptic Sensing and Touch Enabled Computer Applications, Springer. pp. 360–370.

[11] Moder, T., Hafner, P., Wisiol, K., Wieser, M., 2014. 3d indoor positioning with pedestrian dead reckoning and activity recognition based on bayes filtering, in: 2014 International Conference on Indoor Positioning and Indoor Navigation (IPIN), IEEE. pp. 717–720.

[12] Morales, J., Akopian, D., 2017. Physical activity recognition by smartphones, a survey. Biocybernetics and Biomedical Engineering 37, 388–400.

[13] Pei, L., Liu, J., Guinness, R., Chen, Y., Kuusniemi, H., Chen, R., 2012. Using ls-svm based motion recognition for smartphone indoor wireless positioning. Sensors 12, 6155–6175.

[14] Ronao, C.A., Cho, S.B., 2017. Recognizing human activities from smartphone sensors using hierarchical continuous hidden markov models. International Journal of Distributed Sensor Networks 13, 1–16.

[15] Su, H., Yang, C., Mdeihly, H., Rizzo, A., Ferrigno, G., De Momi, E., 2019. Neural network enhanced robot tool identification and calibration for bilateral teleoperation. IEEE Access 7, 122041–122051.

[16] Susi, M., Renaudin, V., Lachapelle, G., 2013. Motion mode recognition and step detection algorithms for mobile phone users. Sensors 13, 1539–1562.

[17] Wan, C., 2019. Feature selection paradigms, in: Hierarchical Feature Selection for Knowledge Discovery. Springer, pp. 17–23.

[18] Xu, W., Miao, Z., Zhang, X.P., Tian, Y., 2017. A hierarchical spatio-temporal model for human activity recognition. IEEE Transactions on Multimedia 19, 1494–1509.

[19] Zaraki, A., Giuliani, M., Dehkordi, M.B., Mazzei, D., D'ursi, A., De Rossi, D., 2014. An rgb-d based social behavior interpretation system for a humanoid social robot, in: 2014 Second RSI/ISM International Conference on Robotics and Mechatronics (ICRoM), IEEE. pp. 185–190.