Human Activity Recognition Based on the Hierarchical Feature Selection and Classification Framework

Yuhuang Zheng

Department of Physics, Guangdong University of Education, Guangzhou 510303, China

Correspondence should be addressed to Yuhuang Zheng; zhyhaa@126.com

Received 31 January 2015; Revised 17 April 2015; Accepted 19 May 2015

Academic Editor: Sos Agaian

Copyright © 2015 Yuhuang Zheng. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Human activity recognition via triaxial accelerometers can provide valuable information for evaluating functional abilities. In this paper, we present an accelerometer sensor-based approach for human activity recognition. Our proposed recognition method used a hierarchical scheme, where the recognition of ten activity classes was divided into five distinct classification problems. Every classifier used the Least Squares Support Vector Machine (LS-SVM) and Naive Bayes (NB) algorithm to distinguish different activity classes. The activity class was recognized based on the mean, variance, entropy of magnitude, and angle of triaxial accelerometer signal features. Our proposed activity recognition method recognized ten activities with an average accuracy of 95.6% using only a single triaxial accelerometer.

1. Introduction

Recently, activity recognition has become an emerging field of research and one of the challenges for pervasive computing. A typical application for activity recognition is in health care. Activity recognition is also an important research issue in building a pervasive and smart environment to provide personalized support.

Computer vision-based techniques and body-fixed accelerators are the main methodologies used for activity recognition. Computer vision-based techniques for activity recognition should be conducted in a well-controlled environment and be subject to the limitations of the environment. However, they may significantly fail in an environment with clutter and variable lighting [1–3]. Body-fixed accelerometers offer a practical and relatively low-cost method to measure human motion.

The existing literature demonstrates many studies on activity recognition that use accelerometers. However, there are three primary challenges in these studies.

(1) The large muscles of the body are controlled for walking, running, sitting, and other activities. The glutes are the primary muscles that drive lower-body movement because of their natural strength and leverage advantage on the legs. Lower-body movement includes activities such as running, jumping, and walking. Sleeping, sitting, standing, walking, running, and jumping must be recognized as typical physical activities. The activity recognition algorithm in Khan et al. [4] did not consider jumping. Running and jumping were excluded from the experiments in the research of Trabelsi et al. [5], Tang and Sazonov [6], Lee et al. [7], and Deng et al. [8]. Gupta and Dallas [9] did not report how to recognize standing and sleeping, and Tao et al. [10] did not describe tests for recognizing sitting and sleeping. Alshurafa et al. [11] studied only walking and running recognition. These studies were incomplete in recognizing typical physical activities [12].

(2) Some studies [6, 9, 13, 14] required the combination of multiple sensors to increase recognition performance. However, a user is less likely to wear a more complex operating system at all times. People may not feel comfortable wearing multiple sensors. Nevertheless, the multisensor systems do not have an enormous advantage over the single-sensor system on the recognition accuracy if the single-sensor system uses a higher sampling rate, suitable features, a more sophisticated classifier, and the correct sensor position, which has the best performance for recognizing activities. A single sensor mounted at the right position can also obtain good recognition performance [15–17].
(3) A series of lectures [18–20] have been given on the topic of recognizing so-called ADL (activities of daily living), which is not physical-activity recognition. “Activities of daily living” is a term used in healthcare to refer to daily self-care activities, such as cooking and hair drying, within an individual’s place of residence or in outdoor environments. Physical activity included any body movement that works the muscles and requires more energy than resting, and it simply implies a movement of the body that uses energy, such as running or walking [21–23]. Physical-activity recognition is discussed in this paper.

Many researchers have used particular devices to collect the raw accelerometer data for a set of movements and various activity recognition algorithms including Artificial Neural Networks (ANN) [4, 7, 13], k-Nearest Neighbor (KNN) [8, 10, 11, 19], Support Vector Machines (SVM) [6, 14, 18], and Hidden Markov Model (HMM) [5, 20]. In our study, we addressed the activity recognition algorithm using SVM for three reasons.

(1) SVM and ANN have been widely used in human activity recognition, although they do not include a set of rules understandable by humans [24]. As two different algorithms, SVM and ANN share the same concept of using the linear learning model for pattern recognition. The difference is mainly on how nonlinear data are classified. Consequently, SVM models have preferable prediction performances to ANN models. SVMs have been demonstrated to have superior classification accuracies to neural classifiers in many experiments. The generalization performance of neural classifiers considers the structure size, and the selection of an appropriate structure relies on cross validation [25]. The performance of SVMs depends on the selection of kernel function type and parameters, but this dependence is less effective [26].

(2) KNN does not perform well when the size of dataset increases, and it is suitable for small datasets. SVM is a complicated classifier; here, we implement the leaner kernel function. We conclude that the accuracy and other performance criteria do not significantly depend on the dataset size, but they depend on the number of training cycles among all factors. The number of training cycles is the best classifier for activity recognition [27].

(3) When a continuous HMM approach to activities is used, the length of the event sequence that gives the best predictions uses sequential data. A HMM is used to model the sequential information in multiaspect target signatures. The parameter-learning task in HMMs is to determine the best set of state transition and emission probabilities given an output sequence or a set of such sequences. The task is usually to derive the maximum likelihood estimate of the parameters of the HMM for the set of output sequences. Typical physical activities are nonsequential, and it is not easy to use HMM to recognize a single physical activity [28].

The traditional SVM [29] is formulated for binary nonlinear classification problems. How to effectively extend the SVM for multiclassification remains a hot topic. The Least Squares Support Vector Machine (LS-SVM) is an advanced version of the standard SVM, and LS-SVM defines a different cost function from the classical SVM and changes its inequation restriction to an equation restriction. Recently, there have been relatively few studies that use LS-SVM to recognize activities using a triaxial accelerometer. Nasiri et al. [30] addressed the Energy-Based Least Square Twin Support Vector Machine (ELS-TSVM) algorithm, which is an extended LS-SVM classifier that performs classification using two nonparallel hyper planes instead of a single hyper plane, which is used in the conventional SVM. ELS-TSVM was used to recognize activities using computer vision instead of a triaxial accelerometer. Altun et al. [31] compared the performances of the least squares method (LSM) and the SVM but did not include the LS-SVM. The LS-SVM for multiclassification is decomposed into multiple binary classification tasks. The LS-SVM for multiclassification reduces the computational complexity by using a small number of classifiers and effectively eliminates the unclassifiable regions that possibly affect the classification performance of this algorithm [32–34].

In this paper, we aimed to overcome the limitations of the existing physical-activity recognition system and intended to develop a new method that could recognize a set of typical physical activities using only a single triaxial accelerometer. This method consisted of three parts, six features for activity recognition, the hierarchical recognition scheme, and the activity estimator based on the LS-SVM and NB algorithms. This method could recognize ten physical activities with a high recognition rate.

The remainder of the paper is organized as follows. Section 2 describes the experimental dataset and hierarchical classification framework in this paper. Section 3 involves feature extraction to improve the classification accuracy using feature data over raw sensor data. Section 4 focuses on an activity estimator for multiclassification to estimate the human activity from the feature data. The experimental results and conclusion are presented in Sections 5 and 6, respectively.

2. Hierarchical Classification Framework

2.1. Activities Dataset. For this work, the used dataset was the University of Southern California Human Activity Dataset (USC-HAD). The USC-HAD was specifically designed to include the most basic and common human activities in daily life from a large and diverse group of human subjects. The activities in the dataset were applicable to many scenarios. The activity data were captured using a high-performance inertial sensing device, which is MotionNode [35]. MotionNode integrates a 3-axis accelerometer, a 3-axis gyroscope, and a 3-axis magnetometer, and the measurement range for each axis of the accelerometer and gyroscope is ±6 g and ±500 dps, respectively. MotionNode was firmly attached onto the participant’s right front hip. The sampling rates of this dataset for both accelerometer and gyroscope were set to 100 Hz. The dataset included 10 activities: walking (forward, left, and right), walking (upstairs, downstairs), jumping, running, standing, sitting, and sleeping [36–38].

The main goal of this paper was to identify ten activities, which were divided into four groups: 2D walking (walking forward, left, and right), 3D walking (walking upstairs,
Table 1: Classified states and activities recognized in this study.

<table>
<thead>
<tr>
<th>State</th>
<th>Activities</th>
<th>Act_Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Walking 2D</td>
<td>Walking forward</td>
<td>WF</td>
</tr>
<tr>
<td></td>
<td>Walking left</td>
<td>WL</td>
</tr>
<tr>
<td></td>
<td>Walking right</td>
<td>WR</td>
</tr>
<tr>
<td>Walking 3D</td>
<td>Walking upstairs</td>
<td>WU</td>
</tr>
<tr>
<td></td>
<td>Walking downstairs</td>
<td>WD</td>
</tr>
<tr>
<td>Plane motion</td>
<td>Jumping</td>
<td>JU</td>
</tr>
<tr>
<td></td>
<td>Running</td>
<td>RU</td>
</tr>
<tr>
<td>Static activity</td>
<td>Standing</td>
<td>ST</td>
</tr>
<tr>
<td></td>
<td>Sitting</td>
<td>SI</td>
</tr>
<tr>
<td></td>
<td>Sleeping</td>
<td>SL</td>
</tr>
</tbody>
</table>

Table 2: A preliminary investigation of \( n \) selection.

<table>
<thead>
<tr>
<th>( n )-class classifier</th>
<th>Number of classifiers</th>
<th>Average accuracy rate of each classifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>9</td>
<td>(\geq 90)</td>
</tr>
<tr>
<td>3</td>
<td>6</td>
<td>(\geq 90)</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>(\geq 90)</td>
</tr>
<tr>
<td>5</td>
<td>3</td>
<td>(\approx 80)</td>
</tr>
</tbody>
</table>

downsstairs), plane motion (jumping, running), and static activities (standing, sitting, and sleeping). The division was performed using a single triaxial accelerometer. The activities are listed in Table 1.

2.2. Hierarchical Classification Framework. To achieve higher scalability than the single-layer framework, a multilayer classification framework was presented. In the first layer, because the walking-related activities (walking forward, walking left, walking right, walking upstairs, and walking downstairs), jumping, running, and static activities were differentiated from one another, we classified the activities into two subsets (walking and all static activities) and two activities (jumping and running) based on feature selection. In the second layer, the walking-related activities subset included plane motion and 3D motion. In this layer, the static activity subset could be classified by standing, sitting, and sleeping. In the third layer, all detailed activities of 2D and 3D walking were recognized [39, 40].

Figure 1 illustrates the structure of the hierarchical classification framework. The yellow boxes represent the activity set, and the green boxes represent the ten types of activities to recognize. Now, the problem of recognizing ten activity classes was broken down to \( n \) distinct classification problems, and the red boxes represent the classifiers. A preliminary investigation of \( n \) selection is reported in Table 2. The four-class classifier was the best selection in this hierarchical classification framework because of the small number of classifiers and high average accuracy rate of each classifier. The four-class classifier was used in this paper.

In the hierarchical classification framework of the four-class classifier, classifier 1, at the top layer, distinguishes walking-related activities, jumping, running, and static activities. Walking-related activities include walking forward, walking left, walking right, walking upstairs, and walking downstairs. Static activities include standing, sitting, and sleeping [37]. Classifier 2, at the second layer, distinguishes plane motions and 3D motions. Classifier 3 recognizes activities from plane motion, and classifier 4 distinguishes walking upstairs and downstairs from 3D motions. Finally, classifier 5 focuses on recognizing different static activities.

### 3. Feature Design and Selection

Recent related work in feature selection was performed in a filter-based approach using Relief-F and a wrapper-based approach using a variant of sequential forward floating search. Because different features were on different scales, all features were normalized to obtain the best results for KNN or Naive Bayes classifiers, which were used for error estimation and ensure equal weight to all potential features [1–6, 8–10, 13, 18, 24, 29].

In our approach, according to the elementary mechanics of walking, running, jumping, and sleeping, we used the means and variances of magnitudes and angles as the activity features and the magnitudes and angles that were produced by a triaxial acceleration vector. The reasons for this approach are as follows. First, according to [41–43], the muscles produce different forces when people walk, run, jump, and sleep. Normally, the forces increase in the order of sleeping, walking, running, and jumping. Based on Newton’s second law, the resultant accelerations of these activities also increase in that order. Second, as in [44], a model of persistent 2D random walks can be represented by drawing turning angles. Detailed features are described below. Third, Shannon entropy in the time domain can measure the acceleration signal uncertainty and describe the information-related properties for an accurate representation of a given acceleration signal.

The triaxial acceleration vector \(\hat{\boldsymbol{A}}(t)\) is

\[
\hat{\boldsymbol{A}}(t) = x_a(t) \hat{\boldsymbol{e}}_x + y_a(t) \hat{\boldsymbol{e}}_y + z_a(t) \hat{\boldsymbol{e}}_z, \quad t = 1, \ldots, n, \quad (1)
\]

where \(x_a(t)\), \(y_a(t)\), and \(z_a(t)\) represent the \( t \) acceleration sample of the \( x \), \( y \), and \( z \) axes. This feature is independent of the orientation of the sensing device and measures the instantaneous intensity of human movements at index \( t \).

We computed the mean, variance, and entropy of magnitude and of the angle of over the window and used them as six features: \( M_{\text{mag}}, V_{\text{mag}}, M_{\text{ang}}, V_{\text{ang}}, \) and \( E_{\text{ang}} \), where \( T \) is the window length, \( \theta \) is the angle between vectors \(\hat{\boldsymbol{A}}(t - 1)\) and \(\hat{\boldsymbol{A}}(t)\), as shown in the following. Let \(i = 1, 2, \ldots, n/T;\) then

\[
M_{\text{mag}} = \left[ M_{\text{mag}}(1), M_{\text{mag}}(2), \ldots \right],
\]

where \( M_{\text{mag}}(i) = \frac{1}{T} \sum_{t=(i-1)T+1}^{iT} |\hat{\boldsymbol{A}}(t)| \),
To explore the performance and correlation among these six features, a series of scatter plots in a 2D feature space is shown in Figure 2. The horizontal and vertical axes represent two different features. The points in different colors represent different activities. In Figure 2(a), the relationship between $M_{mag}$ and $V_{mag}$ is described, and the running, jumping, walking, and static activities are clustered. In Figure 2(b), the straight line between 2D walking (forward, left, and right) and 3D walking (upstairs and downstairs) implies that $M_{ang}$ is an available feature. Figure 2(c) illustrates that the $E_{mag}$ and $M_{mag}$ features successfully partition the triaxial acceleration data samples from walking forward, walking left,
Figure 2: Scatter plots in the 2D feature space ($T = 50$).
and walking right into three isolated clusters, where each cluster contains data samples roughly from one single activity class. Figure 2(d) demonstrates the discrimination power of the $E_{\text{mag}}$ and $M_{\text{mag}}$ features to differentiate walking upstairs and walking downstairs. Figure 2(e) shows that the triaxial acceleration signal can be classified into standing, sitting, and sleeping based on the $E_{\text{mag}}$ and $M_{\text{mag}}$ features.

In this study, we used $M_{\text{mag}}$, $V_{\text{mag}}$, $E_{\text{mag}}$, $M_{\text{ang}}$, $V_{\text{ang}}$, and $E_{\text{ang}}$ as the best features for the classifiers in each layer [45].

4. Activity Estimation for Multiclassification

We presented an activity estimator for multiclassification to estimate the human activity from the feature data. Each activity estimator for the multiclassification included one LS-SVM classifier and a maximum Activity frequency estimator (Figure 3).

We used the LS-SVM [34] method to cluster the feature data. After loading the testing data into Matlab, we built an activity-recognizing model from the data. After the parameters of the model were calculated, we estimated the activity by inputting some test feature data [46]. The function trainlssvm() was used to train the support features of an LS-SVM for classification, and the function simlssvm() was used to evaluate the LS-SVM for some test feature data.

Because $M_{\text{mag}}$, $V_{\text{mag}}$, $E_{\text{mag}}$, $M_{\text{ang}}$, $V_{\text{ang}}$, and $E_{\text{ang}}$ have $(n/T)$ elements, the LS-SVM for the multiclassifier outputs an activity set, which includes $n/T$ elements of Act_Label. The activity set may have different Act_Labels, and we must estimate the Act_Label maximum likelihood in this activity set. We used the Naive Bayes algorithm to compute all Act_Label likelihoods and obtained the human activity using the maximum Act_Label likelihood. The following described how to mathematically compute the maximum Act_Label likelihood:

$$\text{Activity} = \{\text{Act}_1, \text{Act}_2, \ldots\}$$

$$= \text{LS-SVM} \left( M_{\text{mag}}, V_{\text{mag}}, E_{\text{mag}}, M_{\text{ang}}, V_{\text{ang}}, E_{\text{ang}} \right),$$

Figure 4 shows the activity estimator working process, which includes the training stage and testing stage (online activity recognition). In the training stage, the labeled data of triaxial acceleration were normalized and the statistical features were extracted from those synthesized-acceleration data. Then, the multiclassification estimator was used to build the classification model. In the testing stage, unlabeled raw data of the triaxial accelerometer were processed with the method that was used in the training stage. These synthesized data were classified using the multiclassification estimator, and the recognized result was obtained [47, 48].

5. Experiment

The activity recognition dataset was the USC Human Activity Dataset. The activity dataset included ten activities and collected data from 14 subjects. To capture the day-to-day activity variations, each subject was asked to perform 5 trials for each activity on different days at various indoor and outdoor locations. Although the duration of each trial varies for different activities, it was sufficiently long to capture all information for each performed activity [37]. In this section, we estimated the performances of the five activity classifiers in this activity recognition scheme. Table 3 shows the results
of five activity recognition classifiers. These activity classifiers had over 95% accuracy [24] and were acceptable.

The results of these folds are summarized in Table 4. The average recognition accuracy of 95.6% indicates that our proposed human activity recognition scheme can achieve high recognition rates for a specific subject. Because 2D walking and 3D walking are similar, the recognition accuracy of the five walking activities is low. We will attempt to obtain higher recognition accuracy using an adequate amount of training data in future research.

We compared the accuracy rate and running time for common multiclassification methods. All algorithms were run on a computer with CPU i7-2670QM 2.2 G, 8 G ram, and Matlab 2013a. The LS-SVM performed notably well in the tests. The average running time for the hierarchical classification framework with the LS-SVM recognizing activities was 0.021 seconds, which was less than the ANN (Artificial Neural Network), DT (Decision Tree), and KNN (k-Nearest Neighbor) algorithms. We performed the ANN, DT, and KNN classifier tests with the built-in functions of Matlab. The LS-SVM method was also better than ANN, DT, and KNN in terms of the average recognition accuracy rate for the ten activities. Table 5 shows the results.

6. Conclusion and Future Work

This paper aims to provide an accurate and robust human activity recognition scheme. The scheme used triaxial acceleration data, a hierarchical recognition scheme, and activity classifiers based on the LS-SVM and the NB algorithm. The mean, variance, entropy of magnitude, and angle of triaxial acceleration data were used as the features of the activity classifiers. The scheme effectively recognized a typical set of daily physical activities with an average accuracy of 95.6%. It could distinguish walking (forward, left, right, upstairs, and downstairs), running, jumping, standing, sitting, and sleeping activities using only a single triaxial accelerometer. The experimental results of the hierarchical recognition scheme show significant potential in its ability to accurately differentiate activities using triaxial acceleration data. Although the scheme remains to be tested with USC-HAD datasets, the core of this scheme is independent of the features of other activity datasets; therefore, it is applicable to any dataset.

The novelty of the proposed human activity recognition scheme is the introduction of the LS-SVM method as the classifier algorithm. The LS-SVM is an advanced version of the standard SVM, and there are recently relatively few studies using LS-SVM to recognize activities with only one triaxial accelerometer. The human activity recognition scheme with LS-SVM classifiers simplifies the construction of the hierarchical classification framework and has a lower running time than other common multiclassification algorithms. Accuracy is the basic element that must be considered when any activity recognition system is implemented, and this recognition scheme has a high success rate, for which it can recognize ten different types of activities with an average accuracy of 95%.

The next stage of our research has two parts. First, the algorithms are improved to recognize these activities, and the user will not have to worry about placing the sensors at the correct positions to correctly detect the activities. Second, an unsupervised approach for automatic activity recognition is considered. An unsupervised learning framework of human activity recognition will automatically cluster a large amount of unlabeled acceleration data into discrete groups of activity, which implies that the human activity recognition can be naturally performed.

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.
Table 5: Accuracy rates and running times of the classification methods.

<table>
<thead>
<tr>
<th>Method</th>
<th>WF</th>
<th>WL</th>
<th>WR</th>
<th>WU</th>
<th>WD</th>
<th>JU</th>
<th>RU</th>
<th>ST</th>
<th>SI</th>
<th>SL</th>
<th>Average rate (%)</th>
<th>Running time (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ANN</td>
<td>96.1</td>
<td>91.4</td>
<td>90.2</td>
<td>90.5</td>
<td>85.4</td>
<td>77.5</td>
<td>98.6</td>
<td>96.7</td>
<td>95.2</td>
<td>99.1</td>
<td>92.1</td>
<td>0.085</td>
</tr>
<tr>
<td>DT</td>
<td>93.9</td>
<td>94.6</td>
<td>91.7</td>
<td>91.2</td>
<td>90.8</td>
<td>84.9</td>
<td>94.1</td>
<td>95.7</td>
<td>97.2</td>
<td>94.4</td>
<td>92.9</td>
<td>0.411</td>
</tr>
<tr>
<td>KNN</td>
<td>93.5</td>
<td>92.1</td>
<td>90.1</td>
<td>88.2</td>
<td>86.7</td>
<td>88.6</td>
<td>93.8</td>
<td>96.1</td>
<td>95.7</td>
<td>93.8</td>
<td>91.9</td>
<td>0.183</td>
</tr>
<tr>
<td>LS-SVM</td>
<td>95.7</td>
<td>92.9</td>
<td>91.4</td>
<td>92.9</td>
<td>94.3</td>
<td>97.1</td>
<td>97.1</td>
<td>98.6</td>
<td>97.1</td>
<td>98.6</td>
<td>95.6</td>
<td>0.021</td>
</tr>
</tbody>
</table>

Acknowledgments

This work was partially supported by Appropriative Researching Fund for Professors and Doctors, Guangdong University of Education, under Grant IIARF04, and Guangdong Provincial Department of Education under Grants 2013LYM_0063 and 2014GJK161.

References


Submit your manuscripts at
http://www.hindawi.com