Using accelerometers to improve position-based activity recognition

Dmitry Gimon, Hristijan Gjoreski, Boštjan Kaluža, Matjaž Gams
Jožef Stefan Institute, Dept. of Intelligent Systems, Jamova 39, 1000 Ljubljana, Slovenia.

Email: dmitry.guimon@gmail.com, hristijan_g@yahoo.com, {bostjan.kaluza, matjaz.gams}@ijs.si

Abstract. This paper will present the results of the research conducted on wireless accelerometers in fall detection and activity recognition. This research is a part of the Confidence project, whose goal is to provide a health monitoring system for the elderly. Normally, position-based body tags are used to detect postures and activities. This paper reports the results of using accelerometers as both a supplement to and substitute of the position tags. It introduces a combined approach based on machine-learning and wave analysis. Preliminary results indicate an important increase in activity recognition accuracy when using both acceleration and position tags.

1 INTRODUCTION

The research presented in this paper is a part of the Confidence project [1], which aims to create a remote care system to detect health problems of the elderly by monitoring their posture and activities. Currently, wearable position tags are in use for behavior analysis in the project. These tags provide three-dimensional coordinates in space and time. Using a real time stream of input data, filtering and machine learning [2], the system classifies activities and detects alarming situations, such as a slip or a fall. However, position tags are problematic for several reasons. First, they are expensive. Second, the data from position tags are noisy and require a set of filters, which cause a delay of the system. The goal of our research was to investigate the possibility and advantages of using a relatively cheap and precise accelerometer tag system together with the position tag system. Two areas of behavior analysis, fall detection and activity recognition, were explored.

Accelerometers are wireless wearable devices that provide three-axes projection of the acceleration vector with respect to the tag body. When the accelerometer does not move, the measured acceleration is equal to gravity and the angle to the horizon can be calculated. Otherwise, the obtained value is the sum of the acceleration vector and gravity. Accelerometers used in our research provided precise inertial data approximately seven times per second. Let us call a set of acceleration projections and the time of the measurement \( (a_{x,i}, a_{y,i}, a_{z,i}, t_i) \) a snapshot \( i \). In our experiments we used up to four accelerometers attached to chest, belt and ankles. The most significant results were achieved by analyzing chest tag data only.

Since activity recognition is a typical classification problem, the goal of the system therefore is to classify every snapshot coming in from the sensors. In the Confidence project, the following activities are distinguished: walking/standing, sitting, lying, the process of sitting down, the process of lying down and falling [2]. In our research we did not use machine learning classification for fall detection because the very short duration of this activity could be confused with sitting or lying, and the importance of being able to correctly detect falls requires a more accurate classification.

2 FALL DETECTION

In the Confidence project falls are defined as lying on the floor for a certain time [4]. While the analysis allows one to detect the actual moment of a fall by analyzing the length of the acceleration vector, it does not provide information about the location of the person. A graph of the length of the acceleration vector in time shows a low value (actual fall) followed by a high value that is the impact of the surface (such as a floor, bed, etc.). We decided to measure the second component because a person can sit or lie down fast, while an impact without control over the body can show abnormal activity better.

Formalization of the inertial fall detection was the following. If \( |A| = \sqrt{a_x^2 + a_y^2 + a_z^2} > A_{\text{max}} \), then a fall is detected, where the threshold value \( A_{\text{max}} \) was obtained empirically.

In our research, 103 tests of 4 people falling were performed in order to compare the results of using position and acceleration tags. Using inertial data allows the system to react faster while the accuracy of the fall detection is relatively low (76% of falls were detected including falls on the bed, compared to 100% detection with position tags). More promising approach is to evaluate position-based fall detection with inertial data to define the pace of going down.

3 MOVING FILTER

An important step in classifying activities is the ability to detect whether the person is moving. It helps to distinguish between static activities (e.g. lying, sitting) and dynamic ones (e.g. standing up, going down). The moving detection used in our approach uses the data sent by the chest
accelerometer only, and it is very similar to the one introduced in [3]. The change of differences in the acceleration vector length from one obtained measurement to another is used to define moving detection as well as an attribute for machine learning. Mathematical formalization of this principle is represented as:

\[
M = \sum_{i=0}^{n-2} |a_{i+1}^{\text{diff}} - a_i^{\text{diff}}|, \quad \text{where} \quad (1)
\]

\[
a_i^{\text{diff}} = \sqrt{a_{x,i}^2 + a_{y,i}^2 + a_{z,i}^2} - \sqrt{a_{x,i-1}^2 + a_{y,i-1}^2 + a_{z,i-1}^2}
\]

Empirically, \(n=15\) was chosen. The time interval is used to preserve the accuracy of the calculation if the event data is lost. If \(M>P\), where \(P\)-is a threshold value, then the tag is moving; otherwise it is considered as static.

4 ACTIVITY CLASSIFICATION USING MACHINE LEARNING

The activity recognition described in this paper is a process that can be explained in several phases shown in Fig. 1.

![Fig1. Steps of activity recognition.](image)

First, the sensors send the raw data to the system. The software then analyzes the raw data and extracts 22 attributes (features). There are 20 attributes based on acceleration data of the chest sensor, including lengths of the acceleration vectors, directions of the acceleration vectors (angles), and changes in the difference of the acceleration vector length. The direction of the acceleration vector was calculated as an angle between the acceleration vector and one of the axes (here, the \(y\)-axis). We also calculated traditional statistical attributes using the given data such as: mean values, standard deviation, and root mean square. The binary attribute for moving was used in the classification procedure as well.

Besides inertial data, 2 attributes of the position data were used. The first attribute is the value of the \(z\)-coordinate, which is the height of the sensor. This made activities occurring at different heights easily distinguishable. The other attribute is the difference between the current and previous values of the \(z\)-coordinate. This attribute is helpful in recognizing the direction and the speed of the tag’s movement along the vertical axis. Additional low pass filtering for the position attributes is used because of the position sensors’ inaccuracy (about 10-15 cm).

After the process of feature extraction is complete, all calculated attributes are collected together in one feature vector. This vector is analyzed with two modules: the classification model and fall detector. If a fall is not detected, then the machine learning classification model is used to recognize activities. For machine learning, we used a built-in freeware Weka toolkit [5].

We used a longer training sequence of snapshots (approximately 7000 instances) recorded during a person’s performed various activities to create the classification model and then test it with a shorter one (2000 instances). We tried to obtain an equivalent distribution of the activities in both sequences. However, some activities are very short compared to others (e.g. standing up and lying) and their occurrence is relatively rare.

<table>
<thead>
<tr>
<th>Acceleration</th>
<th>Acceleration and position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random forest</td>
<td>SMO-PolyKernel</td>
</tr>
<tr>
<td>Standing</td>
<td>97.28%</td>
</tr>
<tr>
<td>Going down</td>
<td>92.20%</td>
</tr>
<tr>
<td>Sitting</td>
<td>96.79%</td>
</tr>
<tr>
<td>Standing up</td>
<td>95.89%</td>
</tr>
<tr>
<td>Lying</td>
<td>98.55%</td>
</tr>
<tr>
<td>Sitting on the ground</td>
<td>92.30%</td>
</tr>
<tr>
<td>On all fours</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

Table 1. Accuracy of activity recognition.

Our results are presented in Table 1. The accuracy of the classifiers for each activity together with the overall accuracy are shown in this table. The best overall accuracy was archived with the Weka implementation of Support Vector Machine called Sequential Minimal Optimization (SMO) and Random Forest (94.07 % and 93.13 % respectively). The results as shown in Table 1 indicate the high importance of using position attributes to recognize certain activities. As expected, sitting and standing could not be classified correctly without the position data, because both acceleration vector length (absolute value of gravity) and the angles of gravity (orientation of the chest) are the same for these activities.

The position data allowed us to increase the accuracy of classification for the short activities such as sitting down or standing up. But even with all the attributes the accuracy is inadmissible. One cause of this problem is the use of filtered data to calculate the majority of the attributes. While use of filtered data is acceptable for stable activities (e.g standing), short activities are not recognizable. Another reason behind the problem is that inaccurate manual-labeled data were used to train the algorithm that is critical for short activity recognition.
5 WAVE ANALYSIS

Analyzing the graphs of acceleration length, we found that there are similarities in the shape of the areas responsible for the short activities. Segmentation of the graph to comparable areas for the same activity (ignoring of graph fluctuations) is necessary to use these similarities in the recognition process. We called such parts of the graph waves. The key idea of this approach is to compare not the actual values of the function but to use characteristics of its trends. To simplify formalization of the rules we decided to define a wave for discrete graphs fluctuating around 0 as a maximum set of successive points with the same sign. In Fig.2, a wave partition of a graph is shown.

Formally to define a wave of the discrete graph $f(x)$ given as a set of points $F$, we call points $b_i = (x_i^b, y_i^b), b_i \in F$ where $y_i^b \cdot y_{i-1}^b < 0$ as borders. The wave $W_i$ is a set of points $p_k = (x_k^i, y_k^i), p_k \in F$, where $x_k^i \leq x_{i+1}^i$. Each wave has length $l_i = x_{i+1}^i - x_i^i$ and the peak value $w_i$. These are parameters we use in machine learning.

$$w_i = y_{max}^i, \text{ where } |y_{max}^i| = \max |y_k^i|, \quad (2)$$

$$\forall x_k^i \leq x_{i+1}^i$$

To use the suggested approach for analyzing the acceleration length parameter fluctuating around gravity length we subtract it ($g$).

$$a_n(t) = |a(t)| - |g| \quad (3)$$

When the tag is not moving, the function $a_n(t)$ is 0. Fig. 3 shows the graph of the function $a_n(t)$ for the part of the recording with the marked waves.

To use waves with machine learning for activity classification, let us consider that one wave is a part of the only one activity of the person. However, each activity could consist of several waves. For example, the process of sitting down usually consists of three significant waves. Therefore, we relabeled training data so that all the points of one single wave were labeled as one activity. Since any activity is a sequence of waves to classify each point, we used not only the attributes of the wave it is a part of but also the attributes of the previous and the next wave. For the real-time recognition this causes a delay lasting one wave length that is usually less than a second.

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>&lt;-- classified as</th>
</tr>
</thead>
<tbody>
<tr>
<td>270</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>a=standing</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>b=going down</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>166</td>
<td>0</td>
<td>c=sitting</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>8</td>
<td>5</td>
<td>d=standing up</td>
</tr>
</tbody>
</table>

a. Confusion matrix of a classifier without wave attributes.

<table>
<thead>
<tr>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>&lt;-- classified as</th>
</tr>
</thead>
<tbody>
<tr>
<td>271</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>a=standing</td>
</tr>
<tr>
<td>2</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>b=going down</td>
</tr>
<tr>
<td>5</td>
<td>2</td>
<td>170</td>
<td>0</td>
<td>c=sitting</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>1</td>
<td>12</td>
<td>d=standing up</td>
</tr>
</tbody>
</table>

b. Confusion matrix of a classifier using wave attributes.

Table 2. Confusion matrix of activity recognition.

This hypothesis was tested on the recording of a person wearing one chest accelerometer only. The data was then partitioned and labeled. The Random Forest algorithm was used to detect the effect of using wave attributes together with some attributes listed in the previous section (e.g. moving attribute). Table 2 shows the comparison between
the confusion matrices of the classifier when the wave attributes were not used (a) and when the wave attributes were used (b).

The results show that the accuracy of short activity classification increased from 41% to 87% using the wave approach in this example.

6 CONCLUSION AND FUTURE WORK

The paper addressed the question of using acceleration-based system as an addition to or substitution for position-based system. The results indicate that even one accelerometer can increase the accuracy of activity recognition while using solely accelerometers does not give the acceptable accuracy in fall detection and activity recognition for the Confidence project.

Using a partition of the graph with wave analysis to improve machine learning is a promising field of study. The significant results were archived for the short activities. Further research might explore using this approach for statistical analysis of long activities, such as walking.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Community's Framework Programme FP7/2007–2013 under grant agreement N 214986. Operation was partially financed by the European Union, European Social Fund. The authors would like to thank Božidara Cvetković, Violeta Mirchevska, Erik Dovgan and Michelle Lin for their inestimable help.

Reference