DETECTION OF SENSOR WEARING POSITIONS FOR ACCELEROMETRY-BASED DAILY ACTIVITY ASSESSMENT

Bin Yin† and Annelies H. C. Goris‡
Philips Research Laboratories, High Tech Campus 34, 5656 AE Eindhoven, The Netherlands
†Group Biomedical Sensor Systems, email: bin.yin@philips.com
‡Group Care and Health Applications, email: annelies.goris@philips.com

ABSTRACT
In accelerometry-based activity monitoring, to correctly assess the level and recognize the type of physical activity performed by the monitored subject, it is essential to have the knowledge of the sensor wearing position. For consumer healthcare and lifestyle applications, such as with an activity monitor (AM), to restrict users with a fixed sensor position is not friendly and unrealistic as well. Therefore, it raises a need of developing a sensor position detecting method that allows a flexible sensor wearing manner and is able to extract the sensor location with very limited user-device interactions. In this paper, two scenarios for achieving this goal are investigated. They are based on comparing the body position dependent features that are extracted from the measured acceleration data with those in an established feature database. The experiments using naturalistically collected data show the effectiveness of both scenarios, and the one that employs user-specific learning appears more promising for a practical use.

KEY WORDS
Accelerometer, physical activity level, sensor position, pattern classification.

1. Introduction

Accelerometry, compared to other methods such as video recording, electromyography and questionnaires, provides a tool suitable for objective, reliable, long term and low cost monitoring of free-living subjects with very limited restrictions on their daily lives [1-4]. A movement registration unit based on a single body-fixed triaxial accelerometer (TA) has shown a great potential in assessing energy expenditure (EE) and physical activity levels (PALs) as well as in identifying and classifying activities of the monitored subjects [5-8].

Depending on the target population and activities being performed, correlation between the accelerometer readout and EE measured with the doubly labeled water (DLW) method varies. The examples of $R > 0.8$ have been reported in literature [1,9-11] in which Plasqui et al recorded $R > 0.90$ for 15 healthy twin pairs in investigating genetic variation of physical activity.

Bouten et al [12] found that during walking the effect of the place of the accelerometer on the body did not influence the accuracy for the prediction of EE and PALs (i.e. total EE divided by resting EE), and it is believed that this holds for other human common activities as well. However, because of the kinematic differences among various body parts during movement, the place of the accelerometer on the body does have influence on the (analytic) relation between the measured accelerations and the PAL. An example of studying the relation between the one-day activity counts and the PA level (obtained with the DLW method) is shown in Figure 1, where 20 subjects were involved each wearing three TAs (the directlife™ activity pod, New Wellness Solutions, Philips) attached to the chest, waist and wrist, respectively. The resultant curves from linear regression, using only the data from the valid cases, show significant differences that need to be taken into account in interpreting the TA readouts.

Analogously, in activity recognition where basically the same TA readouts as in the PAL prediction are used, selecting a proper classifier, defining relevant features, and optimizing the structure and parameters of the selected classifier are therefore very much dependent on the sensor position as well, especially when only a single TA is adopted. Thus, for obtaining desired recognition performance it is essential to have knowledge of to which part of the body a TA is attached.

Progress in miniaturization, design and wireless communications is making consumerized activity monitoring
products possible. Depending on the factors like weather, clothing, types of activities, a user should have freedom to change the wearing position of the device as he/she wants. This asks for an automatic sensor position detection (ASPD) method that avoids the need for the user to report to the device every time the sensor placement changes.

[13] proposed an activity monitor that can estimate EE taking into account its wearing position. The user is required to, under instructions, conduct predefined standardized activities for a certain duration, such as walking or standing, or provide at what time which standard activity has been performed, to allow the AM extract position information by comparing with an established feature database. This method needs considerably amount of user intervention and may lead to incorrect interpretation of the activity data if the information is wrongly provided.

In this paper, two possible scenarios are investigated that aim to reduce user-device interaction as much as possible, thus achieving an automated sensor position detection. In the first scenario, a general feature database is used and the identification of sensor wearing positions is realized without the need of any user intervention, whereas in the second scenario a user-specific feature database is gradually built up online, during which very limited user action is needed, and automation takes place as soon as a satisfactory identification accuracy is obtained. The experiments show good results for both scenarios, and the one based on a user-specific learning concept appears more promising for a practical use.

2. Detection Scenarios

For an AM device with an ASPD function, any predefined activity to be performed under instructions must be avoided. Features should be extracted from daily activities that are conducted in a completely unsupervised and naturalistic environment. In the proposed scenarios, walking is opted as an example of such activities. Walking is an activity that occurs frequently in the daily life of a mobile person. From the pattern recognition point of view, it has rather distinctive features and thus can be easily recognized from other activities even without the prior knowledge of the position of a sensor.

Based on the acceleration data measured during walking, two scenarios for the sensor position detection are proposed, which are schematically depicted in Figure 2 and Figure 3 respectively. As in a typical pattern classification scheme, features dependent on sensor positions are extracted from the recognized walking data, which serve as inputs to a position detector. The position detector has been trained and validated and is able to provide a classified sensor position based on its inputs. In the first scenario, the training and evaluation of the detector is done completely off-line before the actual detection stage, using a feature data set that consists of records extracted from walking data of various subjects. Then, the established detector is applied generically.

In the second scenario, the detector is made user specific. For this purpose, two modes are required, namely, the training mode and the detection mode. At the beginning of the use, the training mode takes place. The feature extractor analyzes the walking waveform segments collected from a specific user and recognized by the activity recognizer, and derives features that are predefined and sensor position dependent. At the training mode, a user interface is enabled. It gets the sensor position label from the user manual input which is essential to obtain the corresponding class of the extracted features, and hence to be able to train the classifier. The label, together with the extracted features, constitutes a record and is stored in a feature database. The number of records grows as more and more walking events occur and are recognized. Using this feature database, the sensor position detector is trained and validated. Once the classification accuracy reaches a satisfactory level and remains stable, the training mode ends and is succeeded by the detection mode. In this mode, the sensor position is detected in the same way as in the first scenario except that the detection is performed only for a specific user.

In the following two sessions, both scenarios will be investigated with a number of experiments.

3. Data Collection

3.1 Instrumentation

The system used to register body movements is a modified version of ”the tri-axial accelerometer for movement registration” (Tracmor, Philips Research, Eindhoven NL) involved in previous studies for the assessment of physi-
cal activity [1,6,10,14]. The system integrates a sensor and a data logger in one device. The device is 8 × 3.5 × 1 cm in size and weighs 34.85 gram (battery included). During measurement acceleration data are stored locally in the memory of the device and thus there are no tangling wires around the body. Therefore, it has an easy wear-ability and imposes no restrictions of movements for subjects.

The sensor is able to detect accelerations in three perpendicular directions (aligned with the shape of its housing) with a dynamic range of 5G (G = 9.8 m/s²). Considering the major frequency band for common human movements is between 0 and 20 Hz [1,15,16], a sampling frequency of 40 Hz is selected. After being download via USB the data are processed off-line with MATLAB® in a personal computer.

Three sensor positions were considered in the experiments, which were chest, waist and thigh (see Figure 4). For the chest position, the device was attached on a cord that was worn by the subject around the neck and placed under the clothing; the second sensor was worn at waist with the help of an elastic belt; for the thigh position, the device was put in the side-pocket of the subject’s trousers. To allow flexibility in the way of wearing the device, which is important for consumer lifestyle products, the subject was not asked to strictly tune the orientation of each device with respect to the anatomical axis, in particular for the pocket position where fixing the sensor orientation appears difficult anyway.

![Figure 4. Sensor wearing positions being investigated](image)

### 3.2 Subjects

8 healthy subjects were involved in the experiment, 5 male and 3 female. They were requested to wear three sensors at chest, waist and thigh for about 24 hours, typically starting around noon on the first day and finishing about the same time the next day excluding the sleeping hours, so that their daily physical activities could be continuously recorded. Subjects behaved as they normally did in the daily life, so the activities were performed in a completely unsupervised and naturalistic condition. For the purpose of validation, subjects were asked to annotate as exactly as possible their (major) activities. Informed consent formula was signed by each subject before the experiments. Subject characteristics at the time of experiments are shown in Table 1.

#### Table 1. Subject characteristics (Mean±SD)

<table>
<thead>
<tr>
<th>N</th>
<th>Age (y)</th>
<th>Height (m)</th>
<th>Weight (kg)</th>
<th>BMI (kg/m²)</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>33.5±7.3</td>
<td>1.75±0.12</td>
<td>68.9±14.7</td>
<td>22.3±2.1</td>
</tr>
</tbody>
</table>

### 4. Data Processing and Classification Results

#### 4.1 Data Processing

The acquired acceleration signals are normalized and also low-pass filtered to remove high frequency disturbances. Based on the annotation, a number of acceleration waveform segments corresponding to walking was picked up manually from the 24-hour data of each subject. The automated recognition of walking without the knowledge of the sensor placement will be separately discussed in future publication.

Every walking segment consists of 3 arrays of samples, i.e., the readout signals {x_k}, {y_k} and {z_k} from 3 axes of the accelerometer where k represents the sampling moment. At each moment k, the magnitude of the vector \( \vec{v}_k = (x_k, y_k, z_k) \) can be obtained by \( ||\vec{v}_k|| = \sqrt{x_k^2 + y_k^2 + z_k^2} \). Both \( \{x_k\}, \{y_k\}, \{z_k\} \) and \( ||\vec{v}_k|| \) were used in feature extraction.

For each walking segment, typically having a duration of a couple of minutes, features were first calculated on the basis of a window size of 256 samples, i.e., 256/40 = 6.4 seconds, and then averaged over the whole segment to get segment-based features. Together with the sensor position information, the features extracted from one walking segment constituted one record for this segment. All records were stored in databases for the use in classifier training and evaluation. Being measured simultaneously at three body locations, every walking segment generates three records.

#### 4.2 Feature Databases

A user-specific database contains only the features extracted from the walking data of a specific subject. The conjunction of all user-specific databases forms a general feature database. In the study presented here, a general database and a user-specific database (from a specific subject Y) are used, denoted by \( \mathcal{F} \) and \( \mathcal{F}_Y \) respectively. For \( \mathcal{F} \), six subjects each provided two walking segments resulting in \( 2 \times 3 \times 6 = 36 \) records, one subject provided three walking segments resulting in \( 3 \times 3 = 9 \) records, and one subject provided one walking segment resulting in 3 records, so in total \( 36 + 9 + 3 = 48 \) records were obtained. For \( \mathcal{F}_Y \), eight walking segments were picked up from the data of subject Y, giving \( 8 \times 3 = 24 \) records from this specific person.

With information gain evaluation on the databases \( \mathcal{F} \) and \( \mathcal{F}_Y \), eleven most relevant features (among twenty features being generated) were selected. These features can
be divided into three clusters. The first cluster includes five features reflecting the degree of involvement of independent axes in the movement; in the second cluster are another five features that estimate the accelerations in the different planes of the earth coordinate system; in the third cluster, one feature in frequency domain is calculated. For getting the features in the first cluster, a covariance matrix \( R_{xyz} \) among the readouts of the three axes is obtained whose eigenvalues are computed and sorted in a descending order. The features in the first cluster are

- magnitudes of the three eigenvalues of \( R_{xyz} \);
- ratio of the magnitudes of the first and third eigenvalues;
- ratio of the magnitudes of the second and third eigenvalues.

The principle behind this calculation is the so-called principal component analysis (PCA). For the features in the second cluster, the inertial acceleration vector \( \vec{a}_k^{(verti)} \) along the earth graviation direction \( \vec{g} \) and the inertial acceleration vector \( \vec{a}_k^{(horiz)} \) in the horizontal plane were estimated, the horizontal plane being defined as the plane that is perpendicular to the earth graviation direction \( \vec{g} \). The features in the second cluster are

- standard deviation of \( ||\vec{v}_i|| \), indicating the total acceleration power;
- standard deviation of \( \vec{a}_k^{(horiz)} \), indicating the acceleration power in the horizontal plane;
- mean of the magnitude of \( \vec{a}_k^{(horiz)} \);
- peak-to-peak value of \( \vec{a}_k^{(horiz)} \);
- ratio of the standard deviations of \( \vec{a}_k^{(verti)} \) and \( \vec{a}_k^{(horiz)} \).

The feature in the third cluster is

- frequency domain entropy that is computed on \( ||\vec{v}_i|| \) in the bandwidth of \( 0.16Hz - 5Hz \) (walking frequency usually under \( 2.5Hz \) after being transformed into frequency domain with FFT.

In consumer lifestyle applications where sensors could be loosely attached to the body (e.g., being put in the pocket or hanging around the neck), the orientations of the sensors with respect to the earth coordinate system differ from subject to subject and vary over time, so in our study this information was not made use of as prior knowledge.

### 4.3 Scenario 1: Classification on \( F \)

As described in Section 2, the first scenario aims at a completely automatic sensor position detector. In this case, since there will be no interface for a user to provide sensor position information during the actual usage, and therefore training is no longer possible there, the classifier in the detector needs to be trained and evaluated prior to the use and made as generally as possible applicable for different users. In this experiment, classifiers were trained and evaluated with the features from the general database \( F \), and their recognition accuracies are expected to reflect the performance of a practical detector in Scenario 1. Three types of classifiers were tested, namely, Bayes net, Multi-layer perceptron (MLP) and Decision tree [17]. For each classifier, three evaluation settings were considered. They are described as follows:

- setting 1: the classifier is cross validated (10 fold) on all the records in \( F \);
- setting 2: records are randomly split into two parts, one (consisting of 66% of the total records) serving as a training set and the classifier being tested on the rest;
- setting 3: a leave-one-subject-out cross validation is conducted for every subject and an averaged classification accuracy is given.

The evaluation results (true positive rates) are listed in Table 2. In Table 3, as an example, a confusion matrix from the decision tree classifier for Setting 1 is also shown. Compared to the other two, despite the similar low accuracies in the discrimination of the chest or waist position, Decision Tree yielded the best overall results.

**Table 3. Confusion matrix of the decision tree (Setting 1)**

<table>
<thead>
<tr>
<th>Classified</th>
<th>Chest</th>
<th>Waist</th>
<th>Thigh</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td>14</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Waist</td>
<td>6</td>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>Thigh</td>
<td>2</td>
<td>0</td>
<td>14</td>
</tr>
</tbody>
</table>

### 4.4 Scenario 2: Classification on \( F_Y \)

To avoid trading the generality with the accuracy of the detector, Scenario 2 is proposed. In this scenario, an interface is available during the training mode for receiving sensor position information from the user. In this mode, the classifier gets trained and evaluated with the data from a specific user only. As soon as the classifier reaches a pre-defined classification accuracy and remains stable, the detector switches to the detecting mode where complete automation may be achieved. Therefore, the scenario aims at a personalized sensor position detector that is expected to outperform the general detector in Scenario 1. Similarly, three types of classifiers, i.e., Bayes net, Multi-layer perceptron (MLP) and Decision tree were considered, each of which was evaluated for Setting 1 and Setting 2. The results are shown in Table 4. Clearly, much lower classification error rates occurred.

To investigate how many walking records are required in the training mode for a user to reach a satisfactory detection performance, the classification accuracy is tested as the number of records in the database increases. The decision tree is chosen for this evaluation, and the result is plotted in Figure 5. The horizontal axis represents the number of 3-tuple records (including one from the chest, one from the waist and one from the thigh) in the training set, and the vertical axis gives the detection accuracy on an unknown set. At least about \( 4 \times 3 = 12 \) records are needed for training. At \( 7 \times 3 = 21 \) error-free classifier results (in this experiment).
### Table 2. Accuracy (%) of the classifiers for three evaluation settings

<table>
<thead>
<tr>
<th>Position</th>
<th>Bayes Net</th>
<th>MLP</th>
<th>Decision Tree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Setting 1</td>
<td>Setting 2</td>
<td>Setting 3</td>
</tr>
<tr>
<td>Chest</td>
<td>62.5</td>
<td>83.3</td>
<td>64.6</td>
</tr>
<tr>
<td>Waist</td>
<td>81.3</td>
<td>60.0</td>
<td>75.0</td>
</tr>
<tr>
<td>Thigh</td>
<td>93.8</td>
<td>83.3</td>
<td>93.8</td>
</tr>
<tr>
<td>Overall</td>
<td>79.2</td>
<td>76.5</td>
<td>77.8</td>
</tr>
</tbody>
</table>

### Table 4. Accuracy (%) of the classifiers on the database $\mathcal{F}_Y$

<table>
<thead>
<tr>
<th>Position</th>
<th>Bayes Net</th>
<th>MLP</th>
<th>Decision Tree</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Setting 1</td>
<td>Setting 2</td>
<td>Setting 1</td>
</tr>
<tr>
<td>Chest</td>
<td>88.9</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Waist</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Thigh</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
<tr>
<td>Overall</td>
<td>96.0</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

![Figure 5. Classification accuracy versus the number of training records](image)

#### 5. Discussion

In Scenario 1, the overall classification accuracy was between 75% and 80%. The decision tree performed better than the other two, which was well supported by the studies reported in the literature. The thigh position got clearly higher classification rate than the other two, simply due to that it moves more during walking as the lower limbs make strides. The chest and waist positions were relatively poorly classified. This can be also seen from the confusion matrix, where 8 out of 10 errors occurred in recognizing the two positions. It may be explained by the fact that because of the rigidity of the human torso the accelerations during walking generated from these two parts are expectedly very similar, and the limited difference can be easily ruined by other factors. For instance, depending on subject characteristics, such as height and gait pattern, the accelerations measured differ from subject to subject. This subject-to-subject variation can blur the chest-to-waist difference. Moreover, the manner of wearing the device also influences. It was noticed in the experiments that the cord used around the subject’s neck might be too long in some cases, especially for people of small height, resulting in a not well-defined chest position.

With personalized training and evaluation in Scenario 2, the classification accuracy improved significantly, showing no errors at all in most cases for this particular subject. The decision tree, with a classification rate of 100%, again outperformed the rest two. Classification of the thigh position was flawless in all cases. It can be still seen that the misclassifications were prone to occur upon detecting the chest and waist positions. However, as above analyzed, since the subject-to-subject variation became absent, the chest-to-waist difference got better pronounced, leading to a much more accurate detection result.

The curves in Figure 5 show that, as the number of training records increases, the classification accuracy converges quickly to 100%. In this example, in total $7 \times 3 = 21$ records are needed before it reaches this level. In practical use, more records should be required because the acquisition of data corresponding to each sensor position may not grow equally, causing insufficient statistics of the training set. This may be solved by asking the user regularly alternating the wearing positions during the training session. Moreover, in reality regular re-training is advised since the same wearing position may drift from its original place because of the change in clothing, body shape and etc. For instance, depending on the pocket style, the thigh position can differ from trousers to trousers.

In the detection mode, since walking is expected to occur frequently in the daily life for a mobile person, a majority voting technique upon classifying the sensor location
may be employed. For instance, the decisions made for the past \( n \) walking segments are checked, and the one appearing most often is taken as a final decision. This technique helps to further improve the robustness of the detection in a practical and usually noisy environment.

6. Conclusion

In this paper, two scenarios are investigated that aim at a generic and a personalized sensor position detector in accelerometry-based activity monitoring respectively. The experiments with activity data measured in unsupervised and naturalistic environments demonstrate the effectiveness of both scenarios. Because of the absence of subject-to-subject variation, the user-specific detector achieves a remarkable accuracy of \( > 95\% \), showing no error with a decision tree classifier for the particular experiment in the paper. Hence, the proposed personalized detector shows a potential in improving the flexibility of use especially for consumerized activity monitoring products.

Acknowledgment

The authors gratefully acknowledge Ton Rademaker and Henry van Vugt from New Wellness Solutions, Philips, and Alberto Bonomi from Philips Research Eindhoven for the inspiring and valuable discussions. The authors also thank Jan van Herk from Philips Research Eindhoven for the instrumental support.

References


