Behavior Analysis Based on Coordinates of Body Tags

Mitja Luštrek [1], Boštjan Kaluža [1], Erik Dovgan [1], Bogdan Pogorelc [2, 1], Matjaž Gams [1, 2]

[1] Jožef Stefan Institute, Dept. of Intelligent Systems, Jamova 39, 1000 Ljubljana, Slovenia
[2] Špica International d. o. o., Pot k sejmišču 33, 1231 Ljubljana, Slovenia
{mitja.lustrek, bostjan.kaluza, erik.dovgan, bogdan.pogorelc, matjaz.gams}@ijs.si

Abstract. This paper describes fall detection, activity recognition and the detection of anomalous gait in the Confidence project. The project aims to prolong the independence of the elderly by detecting falls and other types of behavior indicating a health problem. The behavior will be analyzed based on the coordinates of tags worn on the body. The coordinates will be detected with radio sensors. We describe two Confidence modules. The first one classifies the user's activity into one of six classes, including falling. The second one detects walking anomalies, such as limping, dizziness and hemiplegia. The walking analysis can automatically adapt to each person by using only the examples of normal walking of that person. Both modules employ machine learning: the paper focuses on the features they use and the effect of tag placement and sensor noise on the classification accuracy. Four tags were enough for activity recognition accuracy of over 93 % at moderate sensor noise, while six were needed to detect walking anomalies with the accuracy of over 90 %.

Keywords: Activity recognition, fall detection, gait, machine learning

1 Introduction

The population of developed countries is aging at an alarming rate, threatening to overwhelm the society's capacity for taking care of its elderly members. New technical solutions are being sought worldwide to ensure that the elderly can live longer independently with minimal support of the working-age population. This is also the primary goal of the Confidence project [3] discussed in this paper.

The Confidence project aims to develop a ubiquitous care system for monitoring users in order to detect health problems. Such problems can be immediate (fall), short-term (limping due to injury, dizziness) or long-term (hemiplegia, Parkinson's disease, age-related deterioration of movement).

The user of the Confidence system will wear a number of radio tags placed on the body. The coordinates of the tags will be acquired by sensors situated in the apartment and a portable device carried outside. This will make it possible to reconstruct the user's posture and movement and analyze his/her behavior. Radio technology is a departure from the more common video surveillance. It was chosen for being thought a lesser threat to privacy by the users – in interviews carried out in the Confidence project the elderly accepted wearing tags even during activities such as bathing.

This paper describes two modules of the Confidence system. The first module recognizes the user's activity as one of the following: walking/standing, sitting, lying, the process of sitting down, the process of lying down and falling. Activity recognition is needed for further analyses specific to each activity. In addition, recognizing falls is important in itself.

The second Confidence module analyzes walking. It computes a general-purpose walking signature intended to detect changes in the user's gait. This signature is used to recognize abnormal walking based on the knowledge of normal walking alone. This is an advantage since obtaining examples of abnormal walking of a particular person can be difficult. Other Confidence modules are not discussed in the paper.

We also present a classifier for recognizing a few of the most common and critical health problems of the elderly that manifest in walking: Parkinson's disease, hemiplegia, pain in the leg and pain in the back.

The objective of our research was twofold. First, to find out if the coordinates of radio tags worn on the body are suitable for health-related behavior analysis. And second, to investigate the classification accuracy achievable using various numbers and placements of tags on the user's body and various amounts of noise in tag coordinates. Both the findings regarding noise and tag placement can affect hardware selection and further development and applications of care systems for the elderly.

2 Data Acquisition

We used 370 recordings of 5 persons performing the activities of interest:
- 45 recordings of falling.
- 30 recordings of lying down.
- 30 recordings of sitting down.
- 85 recordings of walking normally (30 of them with a burden).
- 80 recordings of walking limping (25 due to pain in the leg, 25 due to pain in the back and 30 of unspecified type).
- 50 recordings of walking dizzily.
- 25 recordings of walking with hemiplegia (the result of stroke).
- 25 recordings of walking with Parkinson's disease.

Due to the unavailability of persons with the diseases, those recordings were made under the supervision of a physician by healthy volunteers imitating patients. The physician demonstrated the behaviors and provided guidance during recording.

The recordings consisted of the coordinates of 12 tags worn on shoulders, elbows, wrists, hips, knees and ankles, sampled with 10 Hz. Tag coordinates were acquired with Smart infrared motion capture system. The Smart system adds negligible noise to the coordinates (under 1 mm), which allowed us to control the total amount of noise by adding varying degree of Gaussian noise to the raw coordinates. It also supports an unlimited number of tags, so we could explore various tag placements. The closest approximation to the hardware planned for the Confidence project is Ubisense real-time location system, which was used to determine the amount of noise to be added to the coordinates. The standard deviation of the noise measured in the Ubisense system was 4.36 cm horizontally and 5.44 vertically, to which we refer as Ubisense noise.

3   Activity Recognition

The first step in our analysis of user behavior is to classify the user's activity into walking/standing, sitting, lying, the process of sitting down, the process of lying down or falling. This is accomplished by training a classifier that can recognize the activity from a one-second interval of the user's behavior (other durations were tried, but one second proved most suitable). The feature vector for machine learning is a concatenation of the features belonging to the ten snapshots of the user's posture in that interval. Six feature sets described in the following subsection were considered. We tested multiple machine learning algorithms [8], with Support Vector Machine (SVM) offering the highest classification accuracy.

3.1   Features

Reference Features. The reference coordinate system is immobile with respect to the environment. The reference features consist of the z coordinates and the velocities of all the tags in each of the ten snapshots of the user's posture within the one-second interval to be classified. The x and y coordinates were omitted because the location where an activity takes place is not important. Additional features are the absolute distances and the distances in the z direction between all pairs of tags.

Body Features. The body coordinate system described in our previous work [8] is affixed to the user's body. It enables the observation of the x and y coordinates of the user's body parts, since these coordinates no longer depend on the location in the environment. We considered four variants of the body coordinate system differing in two characteristics. First, the coordinate system may be either fully affixed to the body or it may use reference z coordinates. Second, it may be computed for each snapshot in the one-second interval separately or it may be computed for the first snapshot in the interval and the coordinates in the remaining snapshots expressed in the coordinate system of the first snapshot. The main features are the x, y and z coordinates, the velocities, and the angles of movement of all the tags.

Angle Features. These are the angles between adjacent body parts: the shoulder, elbow, hip and knee angles and the angle between the lower and upper torso.

3.2   Machine Learning Experiments

The data for machine learning were prepared as follows. The recordings described in Section 2 were first labeled with the six activities of interest. Then a sliding window passed over each recording, splitting it into overlapping one-second intervals (one interval starting every one-tenth of a second). Afterwards, the features described were extracted from these intervals. This resulted in 5,760 feature vectors consisting of 240–2,700 features each (depending on the combination of features used). An activity was assigned to each feature vector. Finally, these vectors were used as training data for the SVM learning algorithm. The algorithm was implemented in Weka [12] and used default settings. The results were obtained with ten-fold cross validation.

Machine learning experiments proceeded in two steps. In the first step we compared the classification accuracy of the six single feature sets: reference features, the four types of body features and angle features. In the second step we discarded the less promising feature sets and compared the remaining sets in all reasonable combinations at three levels of noise. The results of the second step are shown in Table 1; the classification accuracy of the best feature set combination is in bold.

Table 1. Classification accuracy of feature set combinations.

Features \ Noise	None	Ubisense	Ubisense × 2
Reference + body with body z	96.7	94.9	93.4
Reference + body with reference z	96.9	95.4	93.5
Reference + angle	**97.7**	**96.5**	**94.7**
Body with body z + angle	95.6	91.3	87.8
Body with reference z + angle	95.5	92.5	89.6
All (body z)	96.9	95.0	93.7
All (reference z)	96.9	95.5	93.8

Table 1 indicates that the reference + angle features are the best feature set combination. It is what we used in all the following experiments. The angle features alone performed rather poorly, but they seem to complement the reference features well. The more difficult to compute body features are apparently not worth the effort.

3.3 Tag Placement and Noise Level

Even though interviews carried out in the Confidence indicated that the users would accept many tags if the benefit was clear, wearing the full complement of 12 tags may be annoying. Therefore we investigated ways to reduce the number of tags and studied the interplay between tag placement and noise level.

The experimental results were obtained by leave-one-person-out method, which means that the recordings of all the persons but one were used for training and the recordings of the remaining person for testing. The intention was not to overfit the classifiers to the specific persons in the training recordings, so the results show the expected classification accuracy on a previously unseen person. This is the setting for the Confidence system, which should work on a new user without a training session.

The classification accuracies for activity recognition were compared for all $2^{12} - 1 = 4095$ combinations of 1 to 12 tags and all noise levels from none to Ubisense × 2 in increments of Ubisense × 0.2. The best tag placement for each number of tags is shown in Fig. 1. The accuracy of the best tag placement for each number of tags and each noise level is shown in Fig. 2. The highest classification accuracy of over 93 % is achieved with four to eight tags at low noise levels. One would expect higher numbers of tags to always perform better than lower numbers, but this turned out not to be the case. The reason is probably that more tags yield more features, which make overfitting to the persons in the training recordings more likely. Since we test with the leave-one-person-out method, such overfitting is punished with a lower accuracy.
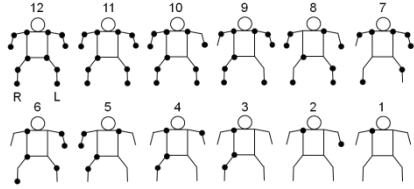
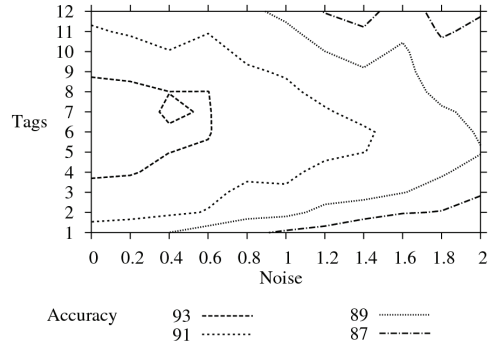**Fig. 1.** Best tag placement for each number of tags for activity recognition.

| Accuracy | 93 ---------- | 89 ·········· |
| | 91 ·········· | 87 —·—·— |

Fig. 2. Classification accuracy with respect to the number of tags and noise level for activity recognition.

Fall Detection. We used a simple rule that recognized a fall when at least three classifications of falling were followed by at least one classification of lying. The accuracy of fall detection was mostly independent of noise. It rarely exceeded 95 %, but it was above 94 % for more than six tags and above 93 % for more than three tags.

4 Analyses of Walking

In the following two subsections we present a classifier for the detection of specific health problems and a Confidence module for the detection of abnormal walking.

4.1 Detection of Specific Health Problems

The specific health problems for detection were suggested by a medical expert based on the incidence in the elderly aged 65+, medical significance and the feasibility of recognition from movement. Four health problems were chosen: Parkinson's disease, hemiplegia, pain in the leg and pain in the back. A physician usually diagnoses such health problems while observing a patient's gait. For the computer to do the same, the relevant gait characteristics must be transformed into computable features [4].

Features. The features for identification of the four health problems were designed with the help of a medical expert. They assume the person is affected on the right side of the body; if he/she were affected on the left side, the sides would be reversed:

- Absolute difference of average distances right elbow – right hip and right wrist – left hip.
- Average angle of the right elbow.
- Quotient between maximal angle of the left and maximal angle of the right knee.
- Difference between maximal and minimal angle of the right knee.
- Difference between maximal and minimal height of the left shoulder.
- Difference between maximal and minimal height of the right shoulder.

- Quotient between {difference between maximal and minimal height of the left and maximal and minimal height of the right ankle}
- Absolute difference of {difference of maximal and minimal speed of the left and difference of maximal and minimal speed of the right ankle}
- Absolute difference of average distances of right shoulder – right elbow and left shoulder – right wrist.
- Average speed of the right wrist.
- Frequency of angle of the right elbow passing average angle of the right elbow.
- Average angle between the vector {right shoulder – right hip} and the vector {right shoulder – right wrist}
- Absolute difference of average heights of the right and the left shoulder.

Machine Learning, Tag Placement and Noise Levels. The machine learning task was to classify walking into five classes: four types of walking with the chosen health problems and the fifth without health problems as a reference. The classifier was trained on the recordings described in Section 2, which were labeled with the type of walking. For each recording the feature vector consisted of the 13 features averaged over the recording. These vectors were used as training data for several machine learning algorithms, of which the SVM learning algorithm achieved the best performance. The algorithm was implemented in Weka [12] and used default settings. Testing was performed with the leave-one-person-out method.

The classification accuracy with respect to the tag placement and noise level was computed. First, various numbers and positions of tags were tested. We started with all 12 tags and then removed them in the order that achieved the best performance. The best tag placement for each number of tags is shown in Fig. 3. Noise level was varied from none to Ubisense × 2 in increments of Ubisense × 0.2. Fig. 4 shows the classification accuracy for each tag placement and noise level. At Ubisense noise, the classification accuracy of 95 % is just out of reach, and to exceed 90 %, at least six tags are needed. In the upper left corner of the graph there is an area with an extremely high accuracy. It requires more tags and lower noise than expected in the Confidence system, but it may be interesting in a clinical setting.
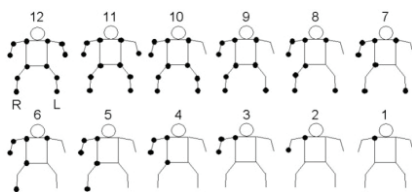

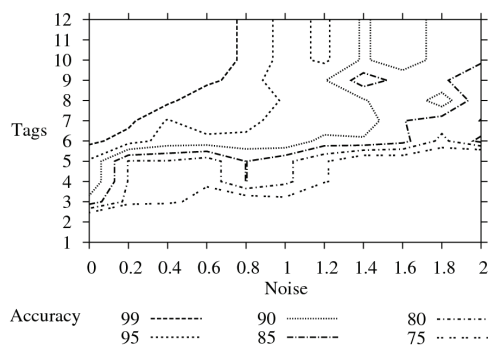**Fig. 3.** Best tag placement for each number of tags for detection of specific health problems.

Fig. 4. Classification accuracy with respect to the number of tags and noise level for detection of specific health problems.

4.2 Walking Signature

Gait is an important indicator of general health condition, particularly in the elderly, and a large body of medical literature is devoted to its study [4, 5, 9]. Walking signature consists of a number of features characterizing the way a person walks. It can be used to detect changes in a person's gait that may be related to a health problem. Unlike the features from the previous subsection, they are not geared towards any specific health problems. Most features refer to a pair of steps, so to compute them we had to develop an algorithm for step detection.

Step Detection. We detect steps by observing the x and y coordinates of the user's ankles (the signal-to-noise ratio in the z coordinates is too low). For each snapshot of the user's posture, the distance in the xy plane an ankle has travelled from the previous snapshot is computed first. The snapshots are then sorted by this distance. The snapshots in the group with the lowest 30 % of distances are considered standing still. Each period of standing still is refined by moving its boundaries to the first and last snapshot with an above-average distance for the group.

Features. The features were adapted from medical literature [5, 9]. Each feature refers to two steps, one with each leg. Wherever applicable, the features are computed for each leg separately and the difference in the values for both legs is also included:
- Support (foot on the ground), swing (foot off the ground) and step (support + swing) times.
- Double support time (both feet on the ground).
- Step length and width.
- Maximal distance of the foot from the ground.
- Ankle, knee and hip angles upon touching the ground.
- Knee angle when the ankle of the leg on the ground is directly below the hip, and knee angle of the opposite leg at that time.
- Minimal and maximal knee and hip angles, the angle of the torso with respect to the ground, and the range for each.
- Hip and shoulder sway (the difference between the extreme left and right deviation from the line of walking).

Machine Learning, Tag Placement and Noise Levels. Since the walking signature was not intended for the recognition of specific health problems, but rather to detect any type of abnormal walking, we used the Local Outlier Factor (LOF) algorithm [2]. This algorithm can recognize abnormal walking based on examples of normal walking alone. It computes a degree of 'outlierness' or abnormality of each example. If the degree exceeds a certain bound for a given example, the example is considered abnormal. The algorithm can thus recognize abnormal walking of a Confidence user by only observing him/her walk normally. Thus it can adapt to each user without needing examples of that user walking abnormally, which can be difficult to obtain.

The training of the LOF algorithm was carried out on the recordings described in Section 2, which were labeled with the type of walking. The step detection algorithm first extracted pairs of steps, after which the walking signature was computed for each pair. This resulted in around 534 feature vectors (depending on how many steps were

detected), consisting of up to 58 features (depending on tag placement). The experimental results were obtained with the leave-one-person-out method.

We again studied the classification accuracy with respect to tag placement and noise level. Four tag placements were considered. Ankles tags were always used, since they are needed for step recognition. They were first used alone, then knee tags were added (four tags in all), then hip tags (six) and finally shoulder tags (eight). Noise level was varied from none to Ubisense × 2 in increments of Ubisense × 0.2. The classification accuracy with respect to the number of tags and noise level is shown in Fig. 5. At Ubisense noise, the accuracy with eight tags is above 95 %, with six tags above 90 %, with four tags around 80 % and with two tags still above 75 %.
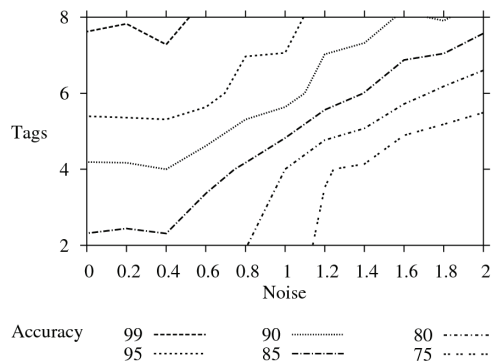

Fig. 5. Classification accuracy with respect to the number of tags and noise level for recognizing normal and abnormal walking with the walking signature.

5 Related Work

Related work on fall detection and activity recognition can be broken down by the choice of hardware (sensors and possibly tags): accelerometers (measure linear acceleration), gyroscopes (measure angular velocity), cameras (not discussed here) and cameras + visible tags (measure tag coordinates). It should be noted that the hardware used in the experiments described in this paper actually belongs to the last category. However, since we added noise to the results, we did not unfairly take advantage of its main strength, which is accuracy.

Fall detection with accelerometers is quite common, particularly using simple threshold algorithms [6]. With a more advanced approach using the One-Class SVM learning algorithm, the accuracy of 96.7 % was reported [13]. A fall detector using a gyroscope attached to the torso achieved the accuracy of 100 % [1]. In both cases falls and the activities from which falls were being distinguished were performed by the same persons in training and in testing, which may account for the high accuracies. Our person-independent testing resulted in accuracies around 94 %.

Accelerometers can also be used for activity recognition. Five tri-axial accelerometers distinguished 30 physical activities of various intensities with the accuracy of 94.9 % with person-dependent training and 56.3 % with person-independent training [11]. We used person-independent training, which resulted in

accuracy above 90 %, although the number of activities in our experiments was admittedly lower.

The related work described so far had objectives similar to ours, but the data it used were significantly different due to the sensors employed. As a consequence, the methodology was different as well, particularly when using video as the source of the data. The approaches belonging to the cameras + visible tags category, however, use cameras to locate tags and thus – like us – work with tag coordinates. The work most similar to ours [10] used 43 body tags sampled with 30 Hz to distinguish between seven activities related to military operations, reporting the accuracy of 76.9 %. This was achieved with the SVM learning algorithm whose features were the tag coordinates belonging to two postures separated by 1/3 second. Our accuracies exceeded 90 % despite more noise and fewer tags, so apparently our features are better suited to activity recognition from tag coordinates.

Motion capture systems consisting of cameras and visible tags are also used for medical research. They commonly provide data for human experts to evaluate, but they can also be used automatically [7]. In distinguishing between health problems such as hemiplegia and diplegia, the accuracy of 92.5 % was reported. Our accuracies were comparable despite more noise and fewer tags (and probably also lower sampling frequency – this is not reported in the related paper).

6 Conclusion

We have investigated the feasibility of using the coordinates of radio tags worn on the body for fall detection, activity recognition and the detection of health problems. The performance of fall detection with person-independent accuracy of around 94 % seems to be comparable to the competitive approaches. The accuracy of activity recognition (over 90 %) often exceeds the alternatives, although admittedly the recognized activities were quite basic. More complex activities will be investigated in the future. Finally, the detection of health problems, which is rarely addressed outside of clinical setting in this form, is quite promising (accuracy 85–95 %). Radio tags and sensors combined with the methods presented in this paper can tackle all these tasks in a single package. They are a viable alternative to inertial and other sensors that can serve the same purpose. At the moment the greatest barrier to the acceptance of such an approach is the price and maturity of the available hardware. However, we are hopeful that this problem will be solved before long.

We have studied the impact of tag placement and noise level on the accuracy of fall detection, activity recognition and the detection of health problems. In general more noise resulted in lower accuracy, as expected. The number of tags sometimes also behaved as expected, i.e., fewer tags resulted in lower accuracy. In activity recognition, however, a moderate number of tags performed best, probably because too many tags caused overfitting to the persons in the training recordings. These results can be used as guidance in further development of the Confidence system and potentially in other projects in the area of ambient assisted living.

Last but not least, the paper describes and compares various features for machine learning. The relatively straightforward features in the reference coordinate system

and angles combined with the SVM learning algorithm proved best for activity recognition. For the detection of specific health problems, specific features turned out to be needed. For recognizing abnormal walking, the walking signature consisting of general gait features was sufficient.

References

1. Bourke, A.K., and Lyons, G.M.: A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor. Medical Engineering & Physics 30, 1, 84--90 (2006)
2. Breunig, M.M., Kriegel, H.-P., Ng, R.T., Sander, J.: LOF: Identifying density-based local outliers. In: 2000 ACM SIGMOD International Conference on Management of Data, pp. 93--104 (2000)
3. Confidence: Ubiquitous Care System to Support Independent Living, http://www.confidence-eu.org
4. Craik R, and Oatis C.: Gait Analysis: Theory and Application. Mosby-Year Book (1995)
5. Heiden, T.L., Sanderson, D.J., Inglis, J.T., and Siegmund, G.P.: Adaptations to normal human gait on potentially slippery surfaces: The effects of awareness and prior slip experience. Gait & Posture 24, 237--246 (2006)
6. Kangas, M., Konttila, A., Lindgren, P., Winblad, P., and Jamsa, T.: Comparison of low-complexity fall detection algorithms for body attached accelerometers. Gait & Posture 28, 2, 285--291 (2008)
7. Lakany, H.: Extracting a diagnostic gait signature. Pattern recognition 41, 1627--1637 (2008)
8. Luštrek, M., and Kaluža, B.: Fall detection and activity recognition with machine learning. Informatica 33, 2, 205--212 (2009)
9. Paróczai, R., Bejek, Z., Illyés, Á., Kocsis, L., and Kiss, R.M.: Gait parameters of healthy, elderly people. Facta Universitatis 4, 1, 49--58 (2006)
10. Sukthankar, G., and Sycara, K.: A cost minimization approach to human behavior recognition. In: The Fourth International Joint Conference on Autonomous Agents and Multi-Agent Systems (AAMAS), pp. 1067--1074 (2005)
11. Tapia, E.M., Intille, S.S., Haskell, W., Larson, K., Wright, J., King, A., and Friedman, R.: Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. In: The 6th International Semantic Web Conference, pp. 37--40 (2007)
12. Witten, I.H., and Frank, E.: Data Mining: Practical Machine Learning Tools and Techniques (2nd edition). Morgan Kaufmann (2005)
13. Zhang, T., Wang, J., Liu, P., and Hou, J.: Fall detection by wearable sensor and One-Class SVM algorithm. In: LNCS, vol. 345, pp. 858--863 (2006)