CONFIDENCE

UBIQUITOUS CARE SYSTEM TO SUPPORT INDEPENDENT LIVING

Small or medium-scale focused research project

DELIVERABLE **D3.3-2**

Report on outdoor reconstruction and interpretation techniques

Contract number:	214986
Project acronym:	CONFIDENCE
Project title:	UBIQUITOUS CARE SYSTEM TO SUPPORT INDEPENDENT LIVING

Deliverable number:	**D3.3-2**
Nature:	Report
Dissemination level:	PU – Public
Report date:	2009-07-28

Author(s):	Mitja Luštrek, Boštjan Kaluža, Matjaž Gams, Naiara Arrue, Igone Vélez
0Partners contributed:	JSI, CEIT
Contact:	Jožef Stefan Institute, Jamova cesta 39, 1000 Ljubljana, Slovenia +386 1 4773380, mitja.lustrek@ijs.si

The CONFIDENCE project was funded by the European Commission under the 7th Framework Programme (FP7) –theme 3 “Information & Communication Technologies”.

ICT 1-7.1 ICT & Ageing

Coordinator: CEIT

CONFIDENCE

UBIQUITOUS CARE SYSTEM TO SUPPORT INDEPENDENT LIVING

Small or medium-scale focused research project

Reviewed by:	**Approved by:**	
Andrés García Alonso External Reviewer CEIT	Igone Vélez CONFIDENCE Project Coordinator CEIT	Miguel Ormazabal Head of CONFIDENCE Project Management Office CEIT

INDEX

EXECUTIVE SUMMARY

Considering that outdoor Confidence localization system will provide not only location data, but also data from accelerometers, gyroscopes and magnetic field sensors, more related work on fall detection and activity recognition becomes relevant than it used to be for the indoor version. Fall detection with accelerometers and gyroscopes can reach very high accuracy even with simple threshold-based algorithms. Therefore we believe these sensors will be valuable for Confidence and we will try to exploit them through machine learning and threshold rules. It should be noted, though, that experiments with these sensors reported in the literature were performed on types of falls and activities of daily living likely to be distinguished by acceleration and velocity, whereas in Confidence we hope to tackle more difficult falls as well.

The first step to be performed by the reconstruction and interpretation system is preprocessing, whose goal is to remove the noise in the data from the localization system. It removes constant small errors with Kalman filter and short-term large errors with median filter. More problematic are long-term large errors, which we plan to remove by a filter that enforces anatomic constraints, i.e., it recognizes when the user's perceived posture is anatomically unlikely. We also intend to improve localization by fusion with other sensors.

The next step is activity recognition, which will mostly be carried out by a machine-learned classifier. Its attributes will be distances between body tags, tag velocities, angles between body parts, accelerations, angular velocities and orientations of tags with respect to the Earth's magnetic field. Short intervals of time will be classified into activities and several ways to merge attributes derived from consecutive sensor readings in those intervals into attribute vectors are considered. We also plan to use expert rules and possibly other methods. The advantage of expert rules is that an expert understands what the general properties of an activity are and can ignore properties specific to an execution of the activity. Furthermore, an expert can imagine variations of the activity. We will pay special attention to user adaptation. We hope to use one of semi-supervised machine learning methods, which improve classification by combining labelled and unlabelled data.

After activity recognition, fall detection is performed. This step uses the information about falling, lying and other postures in which the user may end up after a fall that were recognized during the previous step in order to decide whether a fall has occurred. For this purpose we plan to use expert rules whose parameters will be tuned automatically.

The final task of the reconstruction and interpretation system is general disability/disease detection. It consists of two parts: analyzing the movement of the user's body parts with respect to each other (micro movement) and the user's movement from one location in their environment to another (macro movement). The user's micro movement will be analyzed through statistics characterizing their gait and speed of movement. An outlier detection method will be applied to them, which will learn normal behaviour and detect behaviours sufficiently differing from it. The analysis of the user's macro movement will first require identifying the locations the user frequents. Then it will track the frequencies of visiting these locations and of various activities performed there and raise a warning should they be unusual.

1. INTRODUCTION

The reconstruction and interpretation system receives coordinates of body tags from the localization system, reconstructs the user's activity from them and interprets it as normal or abnormal. Its purpose is to detect falls and other health problems manifesting in movement. If such an event is detected, it raises an alarm or a warning.

The Confidence system may work indoors and outdoors. Deliverable D3.1 [28] described indoor reconstruction and interpretation techniques. This deliverable describes proposed outdoor techniques. The outdoor techniques are similar, but there are three important differences. Firstly, it has been decided in Deliverable D2.3 [3] that outdoor localisation system will provide the distances between the tags instead of their absolute coordinates, which will make reconstruction more challenging. Secondly, instead of ultra-wideband localization hardware, which was planned at the time Deliverable D3.1 was written, Confidence will use FMCW localization hardware. Thirdly, to compensate for the unavailability of the absolute coordinates of the tags, additional sensors, such as accelerometers, gyroscopes and magnetic field sensors are planned to be employed.

2. RELATED WORK

Related work on fall detection and activity recognition can be broken down by the choice of hardware (sensors and possibly tags). We group it into the following categories:

- Accelerometers, which measure linear acceleration.
- Gyroscopes, which measure angular velocity.
- Cameras (not described here, since they will not be used in Confidence).
- Cameras and visible tags, which measure tag coordinates (they will not be used in Confidence, but the data they provide is similar to what Confidence localization system provides).

2.1. Accelerometers

The basic functionality of accelerometers is explained in Deliverable D2.3 [3], which also explains how accelerometers can be used to enhance localization. This subsection describes how accelerometers are directly used for fall detection and activity recognition.

The most common approach to fall detection is to use tri-axial accelerometers with threshold-based algorithms, examples being **Bourke et al.** [5] and **Kangas et al.** [19]. Such algorithms detect a fall simply when the measured accelerations reach a threshold value. They may reach accuracies of up to 98 % [19], but it should be noted that this was measured on fast falls from a standing position, which are well suited to detection with accelerometers. The algorithms were also tested on some activities of daily living that do not seem particularly likely to produce a false alarm. There are even several sensors with build-in hardware fall detection [1][11][30].

Willis [43] used pressure transducers besides accelerometers. He developed a more complex fall detection algorithm based on dynamic belief network models, which can be used to model and produce conclusions about the state of complex temporal environments.

Zhang et al. [47] detected the more obvious falls using One-Class Support Vector Machine machine learning algorithm whose features were accelerations, changes in acceleration etc. For the dubious cases, they used Kernel Fisher Discriminant to reduce the dimensionality of the problem and *k*-Nearest Neighbour algorithm to classify them as falls or non-falls. They reported the accuracy of 96.7 %. Embedding the accelerometer in a cell phone reduced the accuracy a bit [46].

Researchers using accelerometers for fall detection give a lot of attention to the optimal position of the sensor on the body [5][19]. A head-worn accelerometer provides excellent impact detection sensitivity, but it is problematic regarding usability and acceptance. A better option is a waist-worn accelerometer. The wrist does not appear to be an optimal position for fall detection. This is in line with our findings that shoulder, which is closest to the head, is the most suitable tag location for fall detection (and activity recognition) [27]. Some researchers made a step further and used accelerometers for trying to recognize the impact and the posture after the fall [20].

Accelerometers can also be used for activity recognition. **Tapia et al.** [39] used five tri-axial accelerometers to distinguish between 30 physical activities of various intensities. They reported the accuracy of 94.9 % with person-dependent training and 56.3 % with person-independent training. This was achieved with C4.5 decision trees using various time- and frequency-domain features.

Accelerations are apparently valuable information for fall detection and activity recognition. Deriving accelerations from the velocities computed from the changes in tag locations is not feasible in Confidence because of the low localization accuracy and sampling rate. However, as explained in Deliverable D2.3, the outdoor version of the Confidence system will also include accelerometers. Thus, accelerations can be used outdoors. We will first add them to the other attributes for machine learning. We may also use a special classifier for accelerations or we may use a threshold-based algorithm. We believe some falls are difficult to recognize from accelerations alone, but in Confidence we should be able to tackle them as well by using data from other sensors.

2.2. Gyroscopes

The basic functionality of accelerometers is explained in Deliverable D2.3 [3]. Again, that deliverable also explains how gyroscopes can be used to enhance localization, whereas this subsection describes how they can be used for fall detection.

Gyroscopes are not as commonly used for fall detection and activity recognition as accelerometers. **Bourke and Lyons** [4] measured the pitch and roll angular velocities with a bi-axial gyroscope mounted on the torso. They observed the peaks in angular velocity, angular acceleration and torso angle change. They introduced a threshold-based algorithm to distinguish between falls and activities of daily living, which proved 100 % successful in the test setting. This was again measured on falls from a standing position, which are in our experience the easiest to recognize, although they are admittedly also the most common in the elderly. The algorithms were also tested on mundane activities of daily living unlikely to produce a false alarm.

Gyroscopes are also planned for the outdoor version of the Confidence system. Angular velocities may be used in a similar fashion as accelerations: they may be added to the other attributes for machine learning, they may get their own classifier or we may use a threshold-based algorithm.

2.3. Cameras and visible tags

Tracking visible tags with cameras (normal or infrared) gives results similar to those obtained with radio-based localization. The main difference is that the radio-based localization typically has a lower accuracy.

Sukthankar and Sycara [38] 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 machine learning algorithm whose features were the tag coordinates belonging to two postures separated by 1/3 second, reduced in number to 20 using Principal Component Analysis.

Qian et al. [35] used 41 body tags sampled with 120 Hz to distinguish between 21 dance gestures, reporting the accuracy of 99.3 %. The gestures were represented with Gaussian mixture models of joint angles. The high accuracy can in part be attributed to the high quality of input data (little noise, many tags and high sampling frequency), the strictly defined gestures and the fact that testing was always done on the same dancer as training.

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. **Lakany** [22] used them to distinguish between health problems such as hemiplegia and diplegia with the accuracy of 92.5 %. This was achieved with Self-Organizing Maps whose features were wavelet-transformed gait characteristics such as walking speed and stride length.

Since methods using cameras and visible tags produce tag coordinates, they can certainly be applied to Confidence reconstruction and interpretation. This is particularly true indoors, where such methods are already used. Outdoor localization system will not provide absolute tag coordinates, but similar methods can be applied to distances between tags. Potential improvements inspired by related work are representing motion with Gaussian mixture models and wavelet-transforming attributes for machine learning. However, considering that Confidence localization is less accurate than localization with cameras and visible tags, it is possible that these improvements will not prove successful.

3. PROPOSED METHODS FOR THE OUTDOOR RECONSTRUCTION AND INTERPRETATION SYSTEM

The architecture of the outdoor reconstruction and interpretation system will follow the architecture of the indoor system. It consists of modules performing the following main tasks:

- Preprocessing to reduce noise in the data from the localization system.
- Activity recognition to recognize the user's current activity. Recognizing activities is a prerequisite for the next two tasks.
- Fall detection.
- General disability/disease detection, which aims to identify changes in the user's behavior that may indicate a health problem.

A detailed description of the reconstruction and interpretation architecture can be found in Deliverable D3.4 [26].

3.1. Preprocessing

The exact nature of FMCW noise is yet to be studied. However, given measurement with the Ubisense real-time location system [41], which is also radio-based (although it uses ultra-wideband), we anticipate the following types of noise:

- Constant small errors – the perceived tag location fluctuates slightly around the true location. These errors can be characterised with the variance of the error of the localisation subsystem.
- Short-term large errors – the perceived tag location moves far away from the true location for a moment. These errors can be considered as outliers of the localisation subsystem.
- Long-term large errors – the perceived tag location is far away from the true location for a longer period of time. These errors could be characterised as a mean error of the localisation subsystem. This mean error depends on the spatial configuration of the environment and on the placement of the tags and sensors, so it is variable on time.

The first type of noise can be successfully reduced with Kalman filter, which is already used in the indoor reconstruction and interpretation system and is described in Deliverable D3.1 [28]. Kalman filter is also applied in the localization system, so including it in the preprocessing step of the reconstruction and interpretation system may be superfluous. However, here it will be used to smooth tag trajectories after applying other filters, so we may observe an additional benefit.

The second type of noise can be reduced with median filter, which is also used in the indoor system. This filter simply sets each coordinate of each tag to the median of the coordinates within a sliding window. For example, the x coordinate of the neck tag is set to the median of the ten preceding and ten succeeding x coordinates of the neck tag.

Long-term large errors are more difficult to reduce. Two methods are considered. The first one takes into account anatomic constraints and the second one uses machine learning to infer the relations between noise, the properties of the environment and the location of the user. The latter needs predictable environment, which is only possible indoors, so the former seems better suited for outdoors.

3.1.1. Filter enforcing anatomic constraints

A way to detect errors in perceived tag locations if to verify whether the user's perceived posture violates anatomic constraints. For example, if the user's leg appears to be 2 m long, this is most likely because of noise in the data from the localization system.

To detect a violation of anatomic constraints, the minimum and maximum distance between each pair of tags is determined first. These distances are computed from the recordings captured with the

Smart infrared motion capture system [37]. The Smart system has the accuracy of ~1 mm, so its measurements are very reliable. Since we made 815 recordings of various behaviours with it (as described in Deliverable D3.2 [27]), a wide range of possible postures was captured. The distances can be multiplied by a safety factor to account for postures and constitutions not seen before.

Detecting a violation of anatomic constraints is easy, but determining which tags violate them is more difficult. To determine this, first each single tag is assumed to be erroneous. If exactly one such assumption explains all the violations, the error of the tag in question is confirmed. If more than one such assumption explains all the violations, the situation is ambiguous and the erroneous tag cannot be determined. Second, each two tags are assumed to be erroneous. Again, if exactly one such assumption explains all the violations, the error of the two tags in question is confirmed. In principle the procedure could continue for groups of more tags, but since more than five tags are unlikely to be used, going beyond two erroneous tags probably does not make sense.

Finally, the error needs to be corrected. This is accomplished by first computing the average move $\boldsymbol{ma}^t$ of the non-erroneous tags since the last correct measurement of all the tag coordinates. This move is described by Equation (1), where the upper index t indicates the current time, t_0 is the last time all the tag coordinates were measured correctly and B is the number of body tags.

$$ma^t = a^t - a^{t0}; a^t = \frac{\sum_{i=1,\text{tag } i \text{ not erroneous}}^{B} (x_i^t, y_i^t, z_i^t)}{B - \text{the number of erroneous tags}} \qquad (1)$$

Then, the current coordinates of all the erroneous tags are computed as the last non-erroneous coordinates shifted by $\boldsymbol{m}$, as described by Equation (2).

$$e^t = e^{t0} + ma^t \qquad (2)$$

3.1.2. Sensor fusion

Each sensor type used in Confidence is prone to errors, but these errors are independent of each other, so the overall accuracy can be improved by combining the data from different sensors. Deliverable D2.3 [3] already addresses this issue. Here we propose a straightforward approach to using accelerometers and possibly gyroscopes to improve localization accuracy. The approach takes advantage of the fact that averaging several location readings at a given location improves accuracy. If accelerometers and possibly gyroscopes are used to compute relative moves $\boldsymbol{m}^{t+1}_{i\ \mathrm{acc}}$, ..., $\boldsymbol{m}^{t+M}_{i\ \mathrm{acc}}$ from a starting location $\boldsymbol{p}^t_i$ of tag i at time t, these moves can be subtracted from their respective locations, resulting in several readings for a single location. Averaging them results in a more accurate location $\boldsymbol{p}^t_{i\ \mathrm{avg}}$. Then the moves can be added back to $\boldsymbol{p}^t_{i\ \mathrm{avg}}$, yielding more accurate locations $\boldsymbol{p}^{t+1}_{i\ \mathrm{avg}}$, ..., $\boldsymbol{p}^{t+M}_{i\ \mathrm{avg}}$. The procedure is given in Equations (3).

$$\begin{aligned} m_{i\,\mathrm{acc}}^{t+j} &= p_i^{t+j} - p_i^t; i = 1 \ldots B, j = 1 \ldots M \\ p_{i\,\mathrm{avg}}^t &= \frac{\sum_{j=1}^{M} p_i^{t+j} \quad m_{i\,\mathrm{acc}}^{t+j}}{M} \\ p_{i\,\mathrm{avg}}^{t+i} &= p_{i\,\mathrm{avg}}^t + m_{i\,\mathrm{acc}}^{t+j} \end{aligned} \qquad (3)$$

3.2. Activity recognition

The following activities will be recognized:

- Walking/standing
- Sitting (on a chair or something similar)
- Sitting on the ground
- Lying
- On all fours

- Going down (sitting down, lying down or another normal way of lowering the body)
- Standing up
- Falling

Activity recognition first requires attributes that are computed from the raw sensor data. This attributes are then joined into attribute vectors for machine learning or used by some other activity recognition method.

3.2.1. Attributes

In Deliverable D3.1 [28] we assumed there will be twelve tags attached to the user's body: on shoulders, elbows, wrists, hips, knees and ankles. However, after initial experiments [27] and discussions with the stakeholders [2] it was decided that five tags will probably be sufficient. The placement that seems best at the moment is neck (N), belt (B), both ankles (A_L and A_R) and the dominant wrist (W_R for right wrist). Such a placement leads to a simplified kinematic model shown in Figure 1.

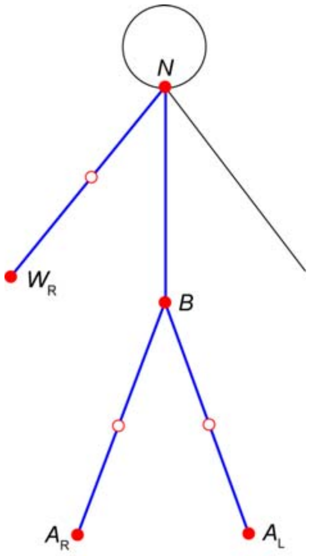

Figure 1. Kinematic model with five tags.

The kinematic model in Figure 1 contains too few tags to define a body coordinate system. Therefore only reference and angle attributes can be computed from tag locations. These two proved best in the past experiments anyway [27]. Note that absolute coordinates are not available outdoor. In addition, non-location attributes obtained from other sensors are also available. B is the number of body tags and N is the number of sets of coordinates in the time interval under observation.

Reference attributes:

- d^t_{ij} ... distance between the body tags i and j at time t; $i = 1 \ldots B$, $j = i + 1 \ldots B$, $t = 1 \ldots T$
- v^t_i … absolute velocity of a body tag (estimated with Kalman filter)
- $(\varphi^t_i, \theta^t_i)$ … angles of movement of a body tag with respect to the z and x axes (estimated with Kalman filter)

Angle attributes:

- $\alpha^t_{\text{LL-T}}$, $\alpha^t_{\text{LR-T}}$ ... the angles of the left and right leg with respect to the torso at time t ($A_R - B - N$, $A_L - B - N$)
- $\alpha^t_{\text{LL-LR}}$ ... the angle between the left and right leg ($A_R - B - A_L$)
- $\alpha^t_{\text{AR-T}}$ ... the angle between the right arm and the torso ($W_R - N - B$)
- α^t_{ER} ... the estimated angle of the right elbow (from the distance between the wrist and the neck)

- α^t_{KL}, α^t_{KR} ... the estimated angles of the left and right knee

Non-location attributes:

- a^t_i ... absolute acceleration of the body tag *i* at time *t* (obtained from accelerometers)
- (a^t_{ix}, a^t_{iy}, a^t_{iz}) ... accelerations in three directions of a body tag (obtained from accelerometers)
- (ω^t_{ix}, ω^t_{iy}, ω^t_{iz}) ... angular velocities in three directions of a body tag (obtained from gyroscopes)
- (Φ^t_{ix}, Φ^t_{iy}, Φ^t_{iz}) ... orientations with respect to the Earth's magnetic field in three directions of a body tag (obtained from magnetic field sensors)

Angle computations

The angles $\alpha^t_{\mathrm{LL-T}}$, $\alpha^t_{\mathrm{LR-T}}$, $\alpha^t_{\mathrm{LL-LR}}$ and $\alpha^t_{\mathrm{AR-T}}$ are computed as follows. Let the angle of interest α be at the point *C* and the lines between which the angle is to be computed between lines *XC* and *YC* (e.g., for $\alpha^t_{\mathrm{LL-LR}}$, $C = B$, $X = A_{\mathrm{L}}$ and $Y = A_{\mathrm{R}}$). The known quantities are the distances $|XC|$, $|YC|$ and $|XY|$. Using the law of cosines, we obtain α with Equation (4).

$$\alpha = \cos^{-1}\frac{|XC|^2 + |YC|^2 - |XY|^2}{2|XC|\,|YC|} \tag{4}$$

The angles α^t_{ER}, α^t_{KL} and α^t_{KR} are estimated as follows. Let the angle of interest α lie at the middle of the points *X* and *Y* (e.g., the right elbow angle lies at the middle of $X = W_{\mathrm{R}}$ and $Y = N$). The known quantities are the current distance $|XY|$ and the maximum distance $|XY|_{\mathrm{max}}$. We obtain α with Equation (5).

$$\alpha = 2\sin^{-1}\frac{|XY|}{|XY|_{\mathrm{max}}} \tag{5}$$

3.2.2. Attribute vectors and machine learning

In order to assign an activity to the time interval under observation, the sets of attributes 1 ... *T* captured in consecutive sensor readings are joined into attribute vectors. A new attribute vector is obtained after every reading (thus overlapping with the previous one). The class of the attribute vector is the activity described by the vector. We consider several methods of joining:

- Concatenation merely concatenates the sets of attributes belonging to each moment in time within the interval of observation. Despite its simplicity, we used this method successfully used in the past [27].
- Average and the slope of a linear approximation of the values of each attribute over all the moments in time within the interval of observation. In recent experiments with the data from the Ubisense system this method proved superior to concatenation.
- Standard deviation, minimum and maximum of the values of each attribute over all the moments in time within the interval of observation. We have not used this method so far, because it seems inappropriate for location data, since it would likely capture more noise than real information. However, we plan to try it on the data from other sensors, which are expected to be less noisy.

An attribute vector may contain attributes joined by one or multiple of the proposed methods. Furthermore, not all attributes need to be joined by the same method.

Activity recognition may be enhanced by appending *A* previous activities as recognized by the classifier to the attribute vector. The danger of appending previous activities is that the machine learning algorithm may learn that the current activity is always the same as the previous one, since this will often be the case. The problem may be circumvented by having two classifiers C_A and C_0. C_A's attribute vector contains *A* previous activities as recognized by C_0. C_0's attribute vector does not contain any previous activities. This way even if C_A gives a lot of weight to previous activities, the previous activities as recognized by C_0 will change, as C_0 is not burdened with C_A's inertia.

The final attribute vectors will be tested with various machine learning algorithms from publicly available toolkits [10][44] and those developed by the project partners. The primary algorithm selection criterion will be the performance (in terms of accuracy, precision, recall) on the recordings of activities of interest. We expect algorithms that cope well with many attributes (such as Support Vector Machine [9]) to perform particularly well, as well as ensemble methods [31] (such as boosting and bagging). Various attribute selection techniques will also be tried. However, we will consider the explicability of the results, too, in which decision trees and rules excel. These algorithms – even if they are not used in the final Confidence software – will help us understand which attributes are important. Such understanding will be useful to design new attributes for machine learning and expert rules.

3.2.3. Other methods

Although we expect machine learning as described in the previous sub-subsection to be the main technique for activity recognition, we will also investigate other methods.

Attribute transformations

Various transformations may be applied to the attributes described in Sub-subsection 3.2.1. Two promising options are the Fourier transform that computes attributes in the frequency domain, and the related but more advanced wavelet transform [14].

Expert rules

These are hand-crafted rules to recognize activities. They may use all the attributes described in Sub-subsection 3.2.1, both alone and joined by methods described in Sub-subsection 3.2.2. The advantage of expert rules over machine-learned classifiers is that an expert understands what the general properties of an activity are and can ignore properties specific to an execution of the activity. Furthermore, an expert can imagine variations of the activity. Machine learning algorithms, on the other hand, can only learn from observed executions of the activity.

Movement primitives

Several methods to represent movement with simple features called movement primitives have been developed. These methods were inspired by the movement of living organisms, which is thought to consist of primitive ‘building blocks’ as well. They are typically used in robotics to model the movement and to control robots. For the purpose of Confidence, only the modelling part is interesting.

Movement is typically represented by a mixture of nonlinear differential equations [16][17][33][36]. Each particular type of movement is characterized by its set of weights. The weights are learned from an observed movement by some form of regression. These weights are the data representation we are seeking.

Other approaches to modelling movement with movement primitives are possible: they may be selected manually [12][29] or derived automatically with dimensionality reduction techniques [18].

Final decision

If multiple methods for activity recognition prove feasible, it may make sense to implement all of them and then reach the final decision by combining their outputs. There are at least three options to do that:

- The methods may vote on the final decision. Their votes can be weighted by the degree of certainty in their decision.
- Another machine-learned classifier may be used. Its attributes are the outputs of the single methods and the attributes they use to reach their decision. Its class may be either the activity or the choice of method.
- Expert rules may be designed to choose between the single methods depending on circumstances.

3.2.4. User adaptation

It is not feasible to (fully) train a machine-learned classifier on the user who actually uses a particular Confidence system. Given the large variation in users, a generic system trained on a range of persons may not perform well on some users with unusual physique or behaviour. Therefore it would be useful to be able to adapt to a particular user. An additional argument for user adaptation is that the behaviour of a user may easily change over time, so even if the system were trained on that user, we could observe the so-called concept drift, which would result in degradation of performance over time.

Height adaptation

The simplest adaptation is probably to measure the height of the user and divide all the location measurements with it. Unfortunately this method has brought little benefit in testing so far.

Supervised learning

The user could be instructed to perform the activities to be recognized, either by the supervision of the Confidence system itself or the technician installing it. These activities are then added to the training set for machine learning (with a large weight) and the classifier is re-trained. However, this may not be acceptable to the users, not to mention that properly labelling training data is difficult, particularly for activities such as going down and standing up. Furthermore, obtaining training data for falls in such a way is not feasible.

Semi-supervised learning

Semi-supervised learning methods [48] attempt to combine labelled and unlabeled data in a way that results in a better classifier than can be built from the labelled data alone. This can be accomplished in a variety of ways.

Generative models assume that the domain can be modelled as an identifiable mixture distribution. Given enough unlabeled data, mixture components can be identified. Then ideally only one labelled example per component is needed to determine the mixture distribution. A straightforward way to adapt this method to Confidence activity recognition is to cluster the unlabeled data of a particular user and then use generic labelled data to assign the class to each cluster. However, the classes in the unlabeled data would probably be unevenly represented (falls would likely be altogether missing) and the method in general seems somewhat unpredictable.

Self-training might be a safer approach. This method uses a classifier trained on the labelled data to classify the unlabeled data. The most reliably classified examples are added to the labelled data and the classifier is re-trained. Unfortunately our initial experiments with self-training were largely unsuccessful.

Co-training is somewhat similar to self-training. It splits attributes in two parts, each of which should be sufficient to train a good classifier. Two classifiers are then trained on the labelled data and used to label the unlabeled data. The examples about which each classifier is most confident are added to the labelled data of the other classifier. The process is then repeated. Since we have devised many attributes for Confidence activity recognition, there may be enough of them for co-training.

Graph-based methods use a graph in which the nodes are labelled and unlabeled examples and the edges reflect the similarity between them. The goal of these methods is to learn a classifier that correctly classifies the labelled examples and uses the similarity between examples to infer the class of the unlabeled examples. The classifier has the form of a function that is close to the label on the labelled examples and is smooth, which means that its value for similar nodes is similar. A variety of algorithms can be used to learn the function, an overview of which is provided in [48]. For Confidence activity recognition, a graph-based method may be applied to a graph consisting of labelled and unlabeled examples connected to their nearest neighbours weighted by the similarity expressed as the distance between attribute vectors.

A concept related to semi-supervised learning is transfer learning. Transfer learning attempts to transfer knowledge learned on one domain (generic activity recognition) to another, related domain (activity recognition for a particular user). This approach has been used for activity recognition [8][21], but due to the differences in setting it is not directly applicable to Confidence. However, the area warrants further study.

3.3. Fall detection

A fall has occurred if the user was first falling and then lying or sitting in a location not meant for lying/sitting or if they were on all fours for an extended period of time. These activities are recognized during activity recognition. Falling itself is quite difficult to recognize in the current indoor reconstruction system because it is distinguished from going down normally mainly by speed, which is difficult to measure reliably from tag coordinates alone. Because of that we expect that accelerometers and gyroscopes will improve the recognition of falling. Lying is quite distinctive, so recognizing it is usually not so difficult.

Fall detection methods in the literature (described in Section 2) mostly rely on measuring accelerations and velocities during falling, so in Confidence terms they are activity recognition of falling. Machine-learning-based methods from the literature are close to machine-learning-based activity recognition in Confidence. Threshold-based algorithms can and probably will be represented with expert rules in Confidence. Recognizing both falling and lying is also known – Tunstall fall detector for example attempts to accomplish this [40].

This subsection describes how recognizing falling and lying is combined into detecting falls and raising an alarm. Based on the research carried out so far, expert rules seem to be best suited to this task. Below is the current set of rules both for indoors and outdoors. They are based on consultations with the partners with expertise on the behaviour of the elderly and on experimental observations.

- IF falling detected in the last 10 s AND lying / sitting at an inappropriate location for 60 % of the last 10 s AND immovable during the same time THEN alarm
- IF falling detected in the last 10 s AND lying / sitting / on all fours at an inappropriate location for 80% of the last 10 s THEN alarm
- IF lying/sitting at an inappropriate location for 90% of the last 10 s AND immovable during the same time THEN alarm
- IF lying/sitting at an inappropriate location for 70% of the last 20 seconds THEN alarm
- IF immovable for 100% of the last hour THEN alarm

These rules raise an alarm if a fall is detected explicitly and the user seems to be unable to get up or if the fall is not detected (as mentioned before, falling is an activity difficult to detect) and the user seems to be in trouble anyway. They assume that certain locations are marked for lying and possibly sitting. If no locations are marked for sitting, the ground is considered inappropriate and elevated surfaces appropriate.

The times and percentages are only provisional and have yet to be tuned. Since each rule only has two parameters, they can be tuned on training data with a simple grid search. Should we add more parameters, we will use a more sophisticated tuning method, possibly based on evolutionary computation. We expect to be able to detect quite reliably whether the user is moving with accelerometers and gyroscopes. The new outdoor sensors may of course cause further changes in the rules.

3.4. General disability/disease detection

Besides fall detection, the goal of the Confidence project is also to detect whether the user has developed some sort of disability, has fallen ill or is otherwise unwell. This is accomplished by monitoring a number of statistics of daily living. The statistics are divided into two groups. The first group describes the user’s movement at a micro scale, mainly of their body parts with respect to each other. The second group describes the user's movement on a macro scale, from one location in their environment to another.

General disability/disease detection will be based on an outlier detection method, such as the Local Outlier Factor algorithm [7]. This algorithm can recognize abnormal behaviour based on examples of normal behaviour 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 behaviour of a Confidence user by only observing them behave normally. This means that after some time it adapts to each user without needing examples of that user behaving abnormally, which can be difficult to obtain. The degree of abnormality required to raise a warning can be adapted based on the user's indication of false alarms.

3.4.1. Micro statistics

The selection of micro statistics [25] was based on medical literature [6][15][32][34][42][45] and experiments with the Ubisense system, which is currently used for testing of the indoor reconstruction and interpretation system. Considering that both Ubisense and FMCW localization hardware are radio-based and have comparable accuracy, we expect most of them will work outdoors with FMCW localization as well. Possible shortcomings of FMCW localization and the absence of absolute coordinates will hopefully be compensated by the additional sensors. It should be noted, though, that some of the statistics, particularly those included in walking signature, are on the border of what can be detected.

Walking signature consists of the following attributes:

- 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).

Walking speed is the speed of walking in a straight line.

Turning consists of the speed while turning, of the angular velocity and of the radius of the turn.

Posture transitions consist of the times for the following posture transitions:

- Sitting → standing
- Standing → sitting
- Lying → standing
- Standing → lying

The general speed of movement measures the speed of each body tag and the average speed of all the tags.

Attribute aggregation

Micro statistical attributes are aggregated over various periods of time. Aggregation means primarily averaging, but standard deviation and other quantities may also be computed. Outlier detection is applied to these aggregated attributes. We consider two approaches to aggregation.

The **straightforward approach** simply aggregates the attributes over various periods of time (the shortest and the longest periods may be omitted, at least for some attributes):

- 1 minute
- 10 minutes
- 1 hour
- 1 day
- 1 week
- 1 month
- 1 year

The **weighted approach** first aggregates the attributes over two-hour periods each hour (so that there are one-hour overlaps between the periods). A weighted average of these two-hour aggregates is then computed over one month, in which greater weight is given to the more recent past.

3.4.2. Macro statistics

Macro statistics deal with the time spent moving / being still performing various activities at various locations. Movement is considered the first attribute, activity the second and location the third. The activities are those described in Subsection 3.1.2. Locations of interest are identified by the fact that the user spends a lot of time there. This can be accomplished by clustering. We consider two approaches to the analysis of macro movement.

In the **simple approach**, each attribute may be first considered separately (e.g., being still, standing, in the garden). Then pairs of attributes may be combined into compound attributes: movement + activity, movement + location and activity + location (e.g., moving sitting, being still at a neighbour’s, walking in a shop). Finally, all three attributes may be combined (e.g., lying still at home). For each attribute, we compute the overall duration over various periods, the average duration of a single episode, the number of episodes and the standard deviation of episode duration.

In the **complex approach**, we have a graph of locations and transitions between them. Based on the user’s behaviour, we compute the probability of going from one location to another. We may also observe the spatial trajectory of the user when making such a transition. At each location, we compute a probability distribution of activities. From such a graph equipped with activity probabilities we compute the probability of the user’s current behaviour and raise a warning if the behaviour is improbable based on their past behaviour.

4. CONCLUSION

The reconstruction and interpretation system has four main tasks. The first task it to reduce noise in the data from the localization system. This is essential because all the other tasks depend on the quality of input data. The second task is to recognize the user's current activity. Recognizing activities is a prerequisite for the next two tasks. The third task is fall detection, which is the main goal of the Confidence project. The final task is general disability/disease detection, which aims to identify changes in the user's behaviour that may indicate a health problem.

Quite a lot of research has been done on fall detection and activity recognition. Little of it was based on location data, though, so until now it has not been used much for indoor reconstruction and interpretation, which relies on location data alone. However, in the outdoor as well as in the enhanced indoor localization system, sensors such as accelerometers, gyroscopes and magnetic field sensors will be added as described in Deliverable D2.3 [3]. This will make it possible to take a greater advantage of related work.

In preprocessing, we plan to address three types of noise. Constant small errors will be corrected with Kalman filter. Short-term large errors will be corrected with median filter. These two have already been implemented and perform adequately. The most problematic are long-term large errors, which are very difficult to recognize automatically. We hope to do so by a filter that enforces anatomic constraints, i.e., it recognizes when the user's perceived posture is anatomically unlikely. We also intend to improve localization by fusion with other sensors.

Activity recognition will mostly be carried out by a machine-learned classifier. It will work similarly to the classifier used indoors, except that it will take advantage of the data from the additional sensors. We also plan to use expert rules and possibly other methods. We will pay special attention to user adaptation. We hope to use one of semi-supervised machine learning methods, which improve classification by combining labelled and unlabelled data. It is, however, difficult to predict how well they will work. Initial experiments with self-learning, which is one of these methods, were not particularly promising. In addition, while unlabelled data usually improves classification, which in terms of Confidence means improved performance with time, the danger of unlabeled data leading the classifier astray cannot be ignored.

Fall detection will mostly rely on expert rules. Since falling, lying and other postures in which the user may end up after a fall are recognized during activity recognition, little additional effort is needed to raise an alarm due to a fall.

General disability/disease detection will consist of two parts. Firstly, the user's micro movement will be analyzed through statistics characterizing their gait and speed of movement. An outlier detection method will be applied to them, which will recognize deviations from normal behaviour. This means that general disability/disease detection will be tailored to each particular user. An additional advantage of outlier detection methods is that they do not require examples of abnormal behaviour, which are difficult to obtain, especially for each particular user. Secondly, we will also analyze the user's macro movement, observing at which location they are and what activities they are performing there.

REFERENCES

[1] AlertOne Services, Inc. iLife Fall Detection Sensor. http://www.falldetection.com/.

[2] Antomarini, M., Cesaroni, F., Piangerelli, E., Sdogati, C., Larsson, K., Bergström, E., González, N., Kämäräinen, A., Palonen, E. M., Rosqvist, E., Vélez, I., Cortés, A., Von der Grün, T., Pietrzyk, M., Gams, M., Luštrek, M., Val, I., Quemada, C., Gomes, S., Gonnet, S., Tamburriello, F., Amedei, G., and Karlsson, P. (2009). D1.3 Report on corrected requirements. Confidence project.

[3] Arrue, N., Vélez, I., Pietrzyk, M. M., and Gruen, T. van den (2009). D2.3 Report on outdoor-location techniques. Confidence project.

[4] Bourke, A. K., and Lyons, G. M. (2006). A threshold-based fall-detection algorithm using a bi-axial gyroscope sensor. Medical Engineering & Physics 30 (1), 84–90.

[5] Bourke, A. K., Scanaill, C. N., Culhane, K. M., O'Brien, J. V., and Lyons, G. M. (2006). An optimum accelerometer configuration and simple algorithm for accurately detecting falls. Proceedings of 24th IASTED International Conference on Biomedical Engineering, 156-160.

[6] Brach, J. S., Berlin, J. E., VanSwearingen, J. M., Newman, A. B., and Studenski, S. A. (2005). Too much or too little step width variability is associated with a fall history in older persons who walk at or near normal gait speed. Journal of NeuroEngineering and Rehabilitation 2 (21).

[7] Breunig, M. M., Kriegel, H.-P., Ng, R. T., and Sander, J. (2000) LOF: Identifying density-based local outliers. Proceedings of SIGMOD'00, 93–104.

[8] Chieu, H. L., Lee, W. S., and Kaelbling, L. P. (2006). Activity recognition from physiological data using conditional random fields. Proceedings of SMA Symposium 2006.

[9] Cristianini, N., and Shawe-Taylor, J. (2000). An introduction to Support Vector Machines and other kernel-based learning methods. Cambridge University Press.

[10] Demšar, J., Zupan, B., Leban, G. (2004). Orange: From experimental machine learning to interactive data mining. White Paper (www.ailab.si/orange), University of Ljubljana, Faculty of Computer and Information Science.

[11] Doughty, K., Lewis, R., and McIntosh, A (2000). The design of a practical and reliable fall detector for community and institutional telecare. Journal of Telemedicine and Telecare 6, 150–154.

[12] Fihl, P., Holte, M. B., Moeslund, T. B., and Reng, L. (2006). Action recognition using motion primitives. Proceedings of 15th Danish Conference on Pattern Recognition and Image Analysis.

[13] González, N., Larsson, K., Snellman, F., Bergström, E., and Cesaroni F. (2008). D6.1 Ethics manual. Confidence projact.

[14] Graps, A. (1995). An introduction to wavelets. IEEE Computational Science and Engineering 2 (2).

[15] Heiden, T. L., Sanderson, D. J., Inglis, J. T., and Siegmund, G. P. (2006). Adaptations to normal human gait on potentially slippery surfaces: The effects of awareness and prior slip experience. Gait & Posture 24, 237–246.

[16] Ijspeert, A. J., Nakanishi, J., and Schaal, S. (2001). Trajectory formation for imitation with nonlinear dynamical systems. Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems.

[17] Ijspeert, A. J., Nakanishi, J., and Schaal, S. (2002). Movement imitation with nonlinear dynamical systems in humanoid robots. Proceedings of International Conference on Robotics and Automation.

[18] Jenkins, O. C., and Matarić, M. J. (2002). Deriving action and behavior primitives from human motion data. In Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems.

[19] Kangas, M., Konttila, A., Lindgren, P., Winblad, I., and Jamsa, T. (2008). Comparison of low-complexity fall detection algorithms for body attached accelerometers. Gait & Posture 28 (2), 285–291.

[20] Kangas, M., Konttila, A., Winblad, I., and Jamsa, T. (2007). Determination of simple thresholds for accelerometry-based parameters for fall detection. In proceedings of 29th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society, 1367–1370.

[21] Kasteren, T. L. M. van, Englebienne, G., and Kröse, B. J. A. (2008). Recognizing activities in multiple contexts using transfer learning. Proceedings of AAAI Fall 2008 Symposium: AI in Eldercare.

[22] Lakany, H. (2008) Extracting a diagnostic gait signature. Pattern recognition 41, 1627–1637.

[23] Luštrek, M. (2009). R3.10-2 Ubisense experiments. Confidence project.

[24] Luštrek, M., and Gams, M. (2008). R3.1-2 Scenarios to detect as suggested by a medical doctor. Confidence project.

[25] Luštrek, M., Gams, M., et al. (2008). R3.4-7 Scenario Details. Confidence project.

[26] Luštrek, M., Kaluža, B., and Mirčevska, V. (2009). D3.4 Indoor reconstruction and interpretation system prototype. Confidence project.

[27] Luštrek, M., Kaluža, B., Dovgan, E., Gams, M., and Vélez, I. (2009). D3.2 Report on performance of indoor reconstruction and interpretation systems, indoor reconstruction and Interpretation system prototypes specification and test plan. Confidence project.

[28] Luštrek, M., Nemec, B., Kaluža, B., Piltaver, R., Gams, M, Vélez, I., Larsson, K., and González, N. (2008). D3.1 Report on indoor reconstruction and interpretation techniques. Confidence project.

[29] Nakaoka, S., Nakazawa, A., Yokoi, K., and Ikeuchi, K. (2004). Leg motion primitives for a humanoid robot to imitate human dances. Journal of Three Dimensional Images 3 (1), 73–78.

[30] Noury, N., Barralon, P., Virone, G., Boissy, P., Hamel, M., and Rumeau, P. (2003). A smart sensor based on rules and its evaluation in daily routines. In proceedings of 25th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society, volume 4, 3286–3289.

[31] Okun, O., and Valentini, G. (2008). Supervised and unsupervised ensemble methods. Springer.

[32] Paróczai, R., Bejek, Z., Illyés, Á., Kocsis, L., and Kiss, R. M. (2006). Gait parameters of healthy, elderly people. Facta Universitatis 4 (1), 49–58.

[33] Perk, B. E., and Slotine, J. J. E. (2006). Motion primitives for robotic flight control.

[34] Podsiadlo, D., and Richardson, S. (1991). The Timed 'Up & Go': A test of basic functional mobility for frail elderly persons. Journal of the American Geriatrics Society 39, 142–148.

[35] Qian, G., Guo, F., Ingalls, T., Olson, L., James, J., and Rikakis, T. (2004) A gesture-driven multimodal interactive dance system. Proceedings of ICME '04, 1579–1582.

[36] Schaal, S. (2003). Dynamic movement primitives – a framework for motor control in humans and humanoid robotics. Proceedings of The International Symposium on Adaptive Motion of Animals and Machines

[37] Smart motion capture system. http://www.emotion3d.com/smart/smart.html.

[38] Sukthankar, G., and Sycara, K. (2005). A cost minimization approach to human behavior recognition. Proceedings of AAMAS 2005, 1067–1074.

[39] Tapia, E. M., Intille, S. S., Haskell, W., Larson, K., Wright, J., King, A., and Friedman, R. (2007). Real-time recognition of physical activities and their intensities using wireless accelerometers and a heart rate monitor. Proeedings of ISWC 2007, 37–40.

[40] Tunstall. http://www.tunstallhealth.com.

[41] Ubisense. http://www.ubisense.net.

[42] Williams, H. G., and Greene, L. S. (1990). Williams-Greene est of physical/motor function. Laboratory report, University of South Carolina, Columbia, Department of Exercise Science, Motor Development / Motor Control Laboratory.

[43] Willis, D. J. (2000). Ambulation monitoring and fall detection system using Dynamic Belief Networks. PhD thesis, Monash University, School of Computer Science and Software Engineering.

[44] Witten, I. H., and Frank, E. (2005). Data Mining: Practical machine learning tools and techniques, 2nd Edition. Morgan Kaufmann, San Francisco, USA.

[45] Woledge, R. C., Birtles, D. B., and Newham, D. J. (2005). The variable component of lateral body sway during walking in young and older humans. Journal of Gerontology: Medical Sciences 60A (11), 1463–1468.

[46] Zhang, T., Wang, J., Liu, P., and Hou, J. (2006). Fall detection by embedding an accelerometer in cellphone and using KFD algorithm. IJCSNS International Journal of Computer Science and Network Security 6 (10).

[47] Zhang, T., Wang, J., Xu, L., and Liu, P. (2006). Fall detection by wearable sensor and one-class SVM algorithm. Lecture Notes in Control and Information Science 345, 858–863.

[48] Zhu, X. (2008). Semi-supervised learning literature survey. Technical report TR 1530, University of Wisconsin – Madison.