Activity Recognition based on High-Level Reasoning
An Experimental Study Evaluating Proximity to Objects and Pose Information

Julia Richter, Christian Wiede, Enes Dayangac, Markus Heß and Gangolf Hirtz
Department of Electrical Engineering and Information Technology
Chemnitz University of Technology, Reichenhainer Str. 70, 09126 Chemnitz, Germany

Keywords: Video Analysis, 3-D Image Processing, Activity Recognition, Pose Estimation, High-Level Reasoning, Ambient Assisted Living.

Abstract: In the context of Ambient Assisted Living (AAL), the detection of daily activities is an active field of research. In this study, we present an algorithm for the performed Activities of Daily Living (ADLs) related to personal hygiene, which is based on the evaluation of a person's proximity to objects and pose information. To this end, we have employed a person detection algorithm that provides a person's position within a room. By fusing the obtained position with the objects' position, we were able to deduce whether the person was occupied with a certain object and to draw conclusions about the performed ADLs. One prerequisite for a reliable modelling of human activities is the knowledge about the accuracy of the person detection algorithm. We have, therefore, analysed the algorithm with regard to its accuracy under different, application-specific conditions. The results show that the considered algorithm ensures high accuracy for our AAL application and that it is even suitable for environments, in which objects are very close to each other. On the basis of these findings, tests with video sequences have been conducted in an AAL environment. This evaluation confirmed that the reasoning algorithm can reliably recognise activities related to personal hygiene.

1 INTRODUCTION

In recent years, a lot of studies have been set on the development of technical support systems for the maintenance of human care standards. The central aim of the AAL system we have developed together with our medical partners is to provide assistance for elderly people and their caregivers, so that the elderly can continue to live in their familiar environment instead of moving to a nursing home. Besides assisting elderly, our system supplies caregivers and relatives with information that helps to optimise the care process. In this study, we focus on elderly people at an early stage of dementia, firstly because they are the group most likely to forget their daily routines, and secondly because they are often incapable of communicating to their caregivers, which daily routines they have already performed. We aim, therefore, at providing caregivers with meta information about their patients' daily activities (ADLs). Based on this information, caregivers can be supported in planning the individual care process and in assessing their patients' need of care.

For this purpose, we detect ADLs by utilising so-called smart sensors. These sensors consist of stereo cameras with an internal processing unit that monitors ADLs. We thus attach high importance to privacy issues: the sensors, which are mounted at the ceiling of the living environment, make information about ADLs available – without releasing raw image data, as it is stored in a database in the form of meta data. This system also generates reminding messages for the elderly if they failed to perform certain activities. Furthermore, the contents of the database can be accessed by the caring personnel via a web interface. By this means, the caring personnel can obtain relevant information that has been inaccessible so far. It enables caregivers to interpret a patient's uncooperative behaviour in the morning when a disturbed sleeping pattern has been detected the previous night. Moreover, drastic changes in the daily activities can be noticed, so that caregivers can react promptly and adapt the caring process to the actual situation. These examples illustrate how the assistance and information system can contribute to a better understanding of the patient's behaviour and improve the caring process. With the presented approach, it will be possible for caregivers to respond appropriately to their patients' individual needs. Current assistance systems have not yet considered the above mentioned aspects.

Richter, J., Wiede, C., Dayangac, E., Heß, M. and Hirtz, G.
Activity Recognition based on High-Level Reasoning - An Experimental Study Evaluating Proximity to Objects and Pose Information.
DOI: 10.5220/0005658804150422
In *Proceedings of the 5th International Conference on Pattern Recognition Applications and Methods (ICPRAM 2016)*, pages 415-422
ISBN: 978-989-758-173-1

2 RELATED WORK

In order to identify special demands or detect emergencies, several approaches have been developed that monitor a person's daily activities, for example the use of motion detectors (Steen et al., 2013) or body-worn sensors (Scanaill et al., 2006). In the work of Pirsiavash et al. (Pirsiavash and Ramanan, 2012), ADLs are detected by a wearable camera that processes the first-person camera view. In our project, we have decided to employ non-wearable, optical sensors, i. e. wide-angle stereo cameras, since people with dementia are apt to put wearable sensors off and to forget them. Besides, more valuable information can be derived from image data than from motion detectors.

One indicator for ADLs is the room in which the person stays (Steen et al., 2013; Richter et al., 2014). In the context of an AAL project, Richter et al. have introduced a person detection algorithm that assigns a room to the person's current position. Based on the chronological order of the rooms a person entered, conclusions about the performed ADLs could be drawn. If the person was detected in the sleeping room at the beginning of the day and afterwards in the bathroom for a few minutes, it can be inferred that the person has attended to his or her personal hygiene. Conclusions about other ADLs, such as the preparation of food, which is usually done in the kitchen, or the leaving of the flat, can be drawn in a similar way.

However, for a more reliable ADL prediction, the analysis of the room alone is insufficient. For this reason, we have developed an algorithm that deduces the performed activities by evaluating the proximity of a person to certain objects in the room. Moreover, pose information obtained by a machine learning based approach (Richter et al., 2015) is used in order to draw conclusions about the activity. The presented study focuses on three objects in the bathroom, the sink, the shower and the toilet. This choice was mainly motivated by medical reasons. The frequency of toileting, for example, provides relevant information for diagnosing and treating incontinence.

In order to determine the proximity of a person to a certain object, the person's position in the room has to be known. Due to the large number of algorithms, a brief summary of the state-of-the-art person detection algorithms should suffice here. Harville et al. (Harville and Li, 2004) used a stereo camera and a person detection algorithm similar to the one already described (Richter et al., 2014). By using the generated depth map as an input, they estimated the virtual overhead view (plan view), which they used for person detection and tracking. Furthermore, they dealt with occlusions and 3-D noise by combining occupancy and height maps. Yous et al. (Yous et al., 2008) introduced the world z-map, which represents the z coordinate of the corresponding world point for every pixel. Their algorithm utilises the fact that the boundaries between close objects persist, whereas they merge on the plane view. In this approach, the height of the mounted camera in relation to the floor and the angle in relation to the wall are required for the determination of the world coordinates of a certain point.

Both these studies have revealed shortcomings of their stereo vision-based algorithms. Harville et al. (Harville and Li, 2004) determined a point-wise mean positional error from ground truth of 16 cm. Yous et al. (Yous et al., 2008) evaluated their algorithm by a video in which each detected person was marked by a cuboid. Furthermore, false positive and false negative rates were calculated. However, neither of the studies provided detailed numerical results about spatial accuracy. When applying a person detection algorithm, it is essential to know the grade of reliability of the determined position in order to draw conclusions about the performed activity. If the person is detected close to the toilet, for example, it is necessary to know how reliable this information is in order to avoid misinterpretations. If the error of the person detection algorithm reaches or exceeds the distance between the objects, the assignment to a certain object is unreliable.

In our project, we utilise the person detection algorithm described in (Richter et al., 2014) with the aim of refining the room assignment to the assignment of objects the person is probably occupied with. Since detailed spatial accuracy measurements have not yet been considered for this algorithm and since there is no standard method to be found in literature, we have developed a method to determine the accuracy in different conditions and scenarios (walking speed, walking direction, exposure time, distance from sensor). By means of this analysis, the accuracy of the algorithm described by Richter et al. (Richter et al., 2014) has been determined in order specify the occurring errors and that might lead to false conclusions with regard to the activities. We thus introduce a method, which can also be applied to other person detection algorithms that determine a person's position in the world.

The accuracy analysis presented in this study yields that persons can be localised in a very accurate way, so that the determined position can be relied on even if the specific objects are very close to each other. On the basis of this finding, we have developed a reasoning algorithm that recognises activities performed in a bathroom. In addition, we have evaluated the algorithm in our testing flat under realistic conditions by comparing the activity determined by the algorithm with labelled video sequences.

The algorithm for reasoning about ADLs is presented in section 3, whereas the employed and analysed person detection algorithm and the pose estimation algorithm are briefly summarized in section 3.1 and section 3.2 respectively. Section 4 presents the methods of analysis for the person detection and the reasoning algorithm. The results of both analyses are presented and discussed in sections 5.1 and 5.2. In section 6, conclusions about the accuracy of both the person detection and the reasoning algorithm are drawn. Besides, an outline regarding the detection of further ADLs by employing the presented approach is given.

3 METHODS AND IMPLEMENTATION

3.1 Location Data Acquisition from Person Detection Algorithm

The employed person detection algorithm (Richter et al., 2014) is able to locate several persons in 3-D coordinates by processing the 3-D point cloud derived from stereo data. In principal, this algorithm works for any sensor providing a 3-D point cloud. In our study, high-quality wide field of view lenses with a focal length of 3.5 mm and small radial distortion are employed. The image resolution is 1360×1024 pixels. The stereo sensor has a baseline distance of 150 mm. Firstly, the algorithm derives foreground hypotheses from the world z-map by applying a mixture of Gaussian segmentation method (Zivkovic, 2004), while the learning rate is set to 0.01 in our study. All points belonging to the foreground are projected onto the floor plane that was previously defined during extrinsic camera calibration. In this floor plane view, blobs of a certain size are detected. The maximum z value in a blob defines the height, the size of the blob defines the width as well as the length of the cuboid that finally characterises the detected person. The center of the cuboid denotes the person's location in the scene. In Figure 1, a detected person of an example recording is shown. At this point, the x and y components of the center point, i.e. the projection of the location onto the floor $p_s = (x_s, y_s)$, are used for the proximity determination, whereas the z component is not relevant at this point.

During the accuracy analysis, the person is always moving and therefore constantly detected as foreground. In our practical AAL application, however, when a person is resting for a longer time, previously detected foreground pixels will become background. In order to localise persons even in such cases, we additionally detect persons by employing a head-shoulder detection algorithm (Dayangac et al., 2015).

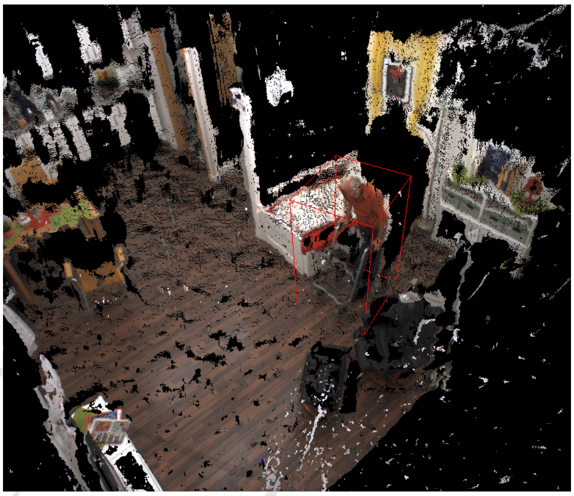

Figure 1: Example point cloud of a proband in our testing flat. The detected person is characterised via a cuboid in 3-D data.

3.2 Pose Information

Information about the general pose is obtained by using the algorithm described by (Richter et al., 2015). In this approach, a classifier is trained using a Support Vector Machine by assigning the points belonging to the person's point cloud according to their z component to vertically aligned bins. Depending on the calculated feature vector, the classifier predicts whether the person is standing, sitting or lying. In our study, we evaluate whether a person is sitting or standing.

3.3 Reasoning About ADLs

For reasoning about ADLs, it is required to determine whether the detected person is close to a certain object. Therefore, the stereo sensors that are distributed in the test flat have been calibrated in such a way that they all share the same world coordinate system. The objects' positions (objects' centres) with respect to the origin of the coordinate system as well as the expansions of relevant and fixed object, such as bed, refrigerator, shower, basis and toilet, have been stored in a look-up-table (LUT). Consequently, if there are n objects, the LUT contains n entries with the according object centres and their expansions. These expansions serve as thresholds $thresh_n$ for the proximity determination.

In order to determine whether a person is close to an object, the distance $dist_n$ between the person's position $p_s = (x_s, y_s)$ and all the objects' positions $p_{o,n} = (x_{o,n}, y_{o,n})$ is compared with the according expansion values $thres_n$ in the LUT, whereas $n = \{1, 2, ..., N\}$, Hereby, N denotes the number of objects and n the index of a specific object. If the distance between the person's center and the object's center is smaller than the defined threshold value in the LUT, the person is considered to be interacting with the object. In this case, the boolean variable $close$ is 1, otherwise it is 0. The following two equations are applied n times if there are n objects in a room. In the presented project, three objects, i. e. sink, shower and toilet, are present in the bathroom of our testing flat.

$$dist_n = \|p_{o,n} - p_s\| \ , \tag{1}$$

$$close = \begin{cases} 1, & \text{if } dist_n < thresh_n \\ 0, & \text{otherwise.} \end{cases} \tag{2}$$

At this point, it should be stated that the bathroom is relatively small (approximately 2.00 m $\times$ 1.80 m) compared to other rooms in our testing flat. This results in a more challenging assignment, because the objects are very close to each other.

This procedure, in combination with pose information (standing and sitting), allows us to reason about the following ADLs:

- Activities that are typical performed when standing in front of a sink, such as washing hands, combing, teeth brushing, etc., if the person is close to the sink and standing.
- Using the toilet if the person is close to the toilet and sitting.
- Taking a shower if the person is close to the shower and standing.

If the person is detected in the shower and standing, then the algorithm outputs the activity "Taking a shower". If the person is detected close to the sink, then it is concluded that the person performs activities that are typically for standing in front of a sink, like "washing hands, combing, teeth brushing, etc.". Similarly, if the person's center is near the toilet and the person is detected to be sitting, we reason that the person probably uses the toilet. If none of the previously mentioned scenarios occur, we assume that the person is doing another action.

Figure 2 shows a scenario in the bathroom of our testing flat, where a person is attending to his or her personal hygiene.

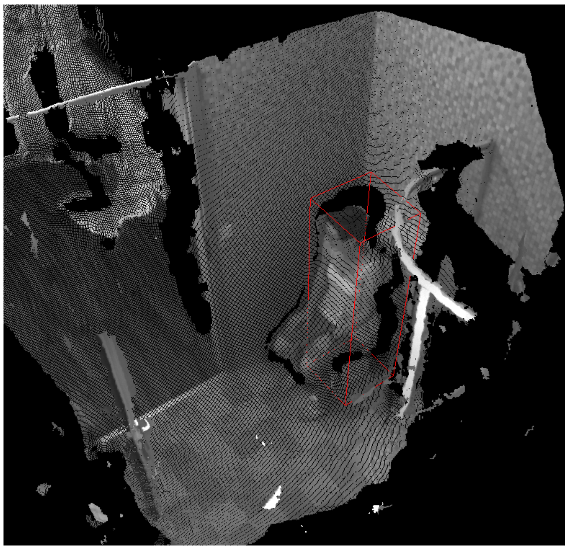

Figure 2: The person is detected very close to the toilet. Moreover, the algorithm determined that the person is sitting. Based on this information, the system reasons that the person is probably using the toilet. This view was generated in debug mode. Due to privacy aspects, it is not visible outside of the smart sensor in the final application.

4 ANALYSIS METHODS

In the following section, the analysis methods of both the person detection algorithm and the reasoning algorithm for the ADLs are presented.

4.1 Person Detection

In order to determine the the accuracy of the person detection algorithm (Richter et al., 2014), the output of this algorithm is compared with the position information provided by a reference system. Each time both a stereo measurement p_s and a reference measurement p_r are obtained, the error e of the stereo measurement is calculated as the Euclidean distance between these two values.

$$e = \|p_r - p_s\| \ . \tag{3}$$

Error histograms are generated for every scenario on the basis of these error measurements. They illustrate the relative occurrence of the errors within different error intervals.

4.1.1 Position Data Acquisition from Reference System

The reference system is completely independent from the stereo sensor system. It is composed of infra-red emitters and cameras that are able to detect special rigid bodies. The reference system provides reliable data at a frame rate of 120 FPS, so that higher walking speeds have no critical influence on the reference measurement. In order to obtain reliable reference data even with occurring reflections and occlusions, eight of such devices are installed in the surveyed volume, so that a good coverage with redundancy is achieved. These emitters and cameras are distributed along a rectangle on the ceiling that surrounds the surveyed area. The rigid body that has been designed for this accuracy analysis is a symmetric construction of four infra-red reflecting markers. This device is installed in the center of a helmet, which a person wears on the head during all recordings. This construction is designed in such a way that it does not influence the stereo measurement. The reference system outputs the 3-D center position of the rigid body. In this study, only the x and the y coordinates of the 3-D position, denoted as $p_r = (x_r, y_r)$, are used.

In order to directly compare the measured position and the reference position, both the reference and the stereo camera system are extrinsicly calibrated, so that they share the same world coordinate system.

4.1.2 Analysis Method

In order to obtain statements about the accuracy, the following different parameter configurations are investigated:

Table 1: Description of the scenario configurations.

Scenario	Description
1	Distance from the stereo sensor, determined for the scenario configurations 2 - 3
2	Moving speed: slow ($\approx 0.3\,\mathrm{m/s}$) vs. fast ($\approx 1.3\,\mathrm{m/s}$)
3	Moving direction: parallel vs. perpendicular to the optical axis of the stereo sensor
4	Exposure time: high (40 ms) vs. low (10 ms)

By setting the mentioned exposure times, a dark scene with low background illumination as well as a bright scene with very high illumination have been reproduced. For scenario 2 to scenario 4, two video sequences with one person walking within the surveyed area have been recorded for each scenario, whereas the exposure time was 20 ms for scenario 2 and 3. All these recordings were used for scenario 1.

Moreover, the mean, minimum and maximum error have been calculated for every scenario. In order to investigate how the distance between sensor and person influences the error, the mean error has been calculated for different radius intervals using all the recordings from scenario 2 to scenario 4. The interval limits are 500 mm, 1000 mm, 1500 mm, 2000 mm, 2500 mm, 3000 mm and 3500 mm (see Figure 3). For each interval, the mean error is calculated and plotted in the middle of the interval.

4.2 Reasoning Algorithm

In order to evaluate the reasoning algorithm, three video sequences with three different persons have been recorded in the bathroom of our testing flat. In these sequences, a typical morning routine is reproduced: Each testing person is entering the bathroom and attending to his or her usual personal hygiene. Thereby, typical activities, such as using the toilet, washing the hands and showering, are performed. Every frame is recorded with a time-stamp. At the same time, the system subsequently outputs the determined ADL, accompanied by the time-stamp, into a file. After the sequence has been recorded, every frame was labelled with the actual ADL. Afterwards, both the labelled ADL as well as the action that has been determined by our system is plotted over time. In this way, the output of the system can be compared against ground-truth data. Moreover, it is possible to identify potential delays the system exhibits.

5 RESULTS AND DISCUSSION

5.1 Person Detection

In this section, the results of the person detection analysis are presented and discussed for each scenario.

Figure 3 illustrates the influence of the distance to the mean of the error, whereas the information between the measured mean values is linearly interpolated.

The graphs show a higher error between the position measured by the stereo camera and the reference position in the range with a small distance to the stereo sensor (≤ 1000 mm). For distances higher than 1000 mm, a tendency of an increasing mean error can be observed. Generally, it can be stated that the mean error will further increase at higher distance, because the accuracy of the stereo data decreases with higher distance.

In Table 2, the mean, the minimum and the maximum error for each parameter configuration of scenario 2 to scenario 4 are listed. The mean errors show values ranging from 74 mm to 87 mm, which is accurate enough for our application, because the

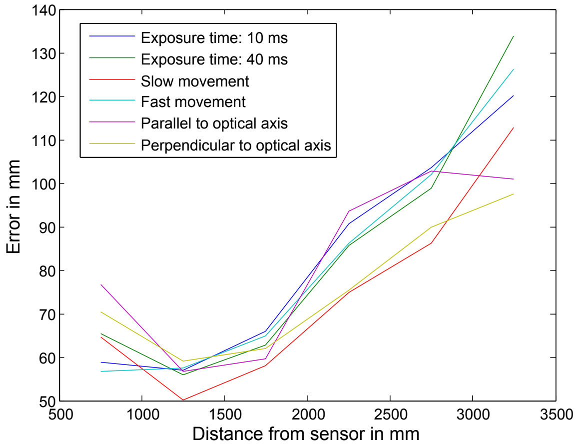

Figure 3: Scenario 1. Influence of the distance on the mean error for the defined intervals. For distances higher than 1000 mm, the error between the position measured by the stereo camera and the reference position increases from about 60 mm to a range from 100 mm to 130 mm.

distances between sink, shower and toilet are several times higher.

Table 2: Determined mean, minimum and maximum errors for scenario 2 to scenario 4. All numbers in mm.

Scenario	Configuration	mean	min.	max.
2	Slow	74	2	339
	Fast	83	2	654
3	Parallel	87	1	416
	Perpendicular	76	1	842
4	Exp. 40 ms	83	1	396
	Exp. 10 ms	86	2	398

In the following, the histograms are presented and discussed for each scenario. In Figure 4, the results for the two different moving speeds are presented. Figure 4 and Table 2 show that the person detection algorithm is more accurate for slower movements.

Figure 5 shows the results for the two different moving directions.

In this experiment, we expected that the parallel movement shows better results than the perpendicular movement. This assumption is based on the fact, that one stereo sensor just provides the 3-D points of the part of a person that is visible to the sensor, i. e. half of the person's surface, either viewed from the side or from the front or back respectively.

When the person is viewed from the side, the points belonging to the arm and shoulder might cause the center of these points to be located near the shoulder instead near the actual body center (head). When the person is viewed from the front or back respectively, the center of the visible surface points is more likely to be located near the actual body center. Therefore, we expected that the center is shifted

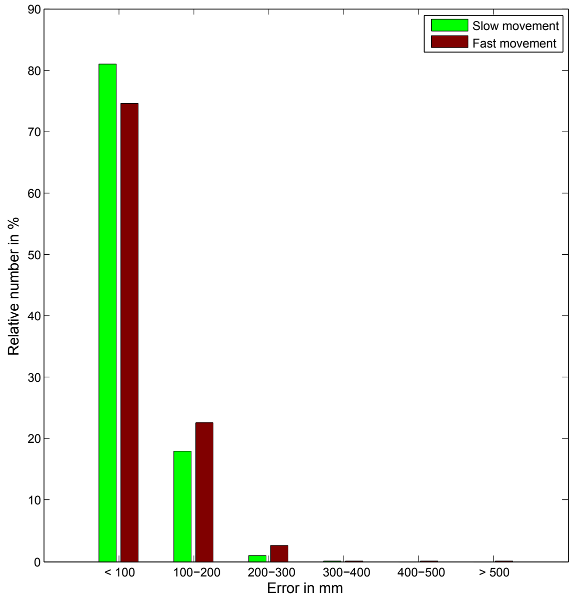

Figure 4: Histogram for scenario 2, moving speed.

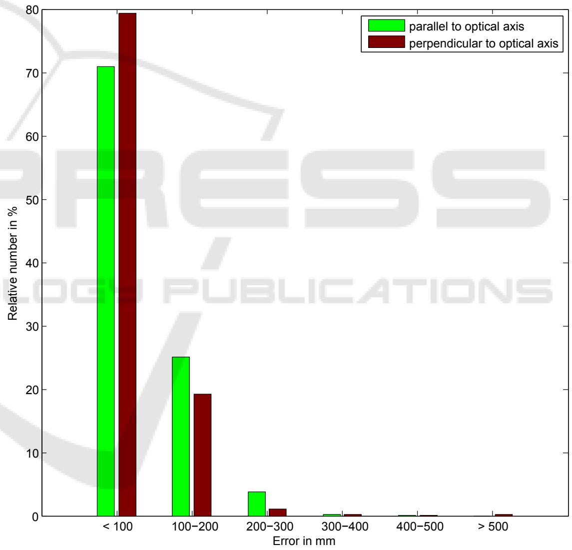

Figure 5: Histogram for scenario 3, moving direction.

to one shoulder when the person moves perpendicular (viewed from the side), while the shift to the person's back or front is expected to be smaller when the person walks parallel to the optical axis (viewed from the front). In contrast to the expectation, the results of the perpendicular movement direction are even better than those of the parallel direction according to Figure 5.

Figure 6 shows the results for the two different exposure times that represent different lighting conditions in an AAL environment (bright scene vs. dark scene).

In this test set-up, the person detection algorithm performs slightly better in the brighter scene with the

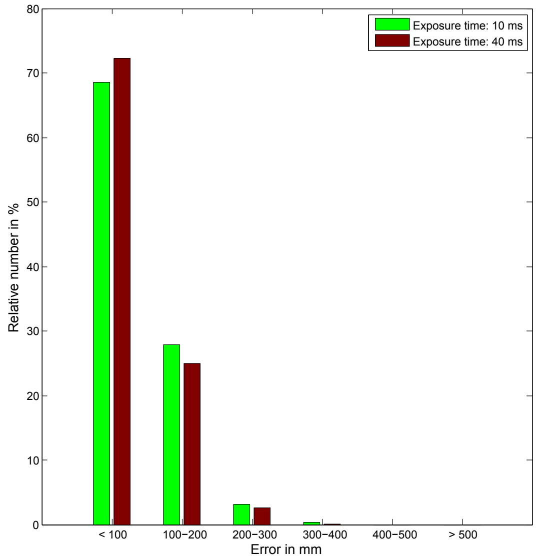

Figure 6: Histogram for scenario 4 (exposure time).

exposure time of 40 ms.

Generally, it can be observed that a configuration change has only a slight effect on the results. The histograms show only a small difference for the parameter changes. Furthermore, Figure 3 shows that the plotted lines are close to each other, which demonstrates that the error is similarly high for the different scenarios. Consequently, the experiments proved that the person detection algorithm is rather robust against these changes. During the experiments, systematic errors are occurring. We assume that the major inaccuracies in the measurements originate from the calibration procedures, i. e., intrinsic and stereo calibration as well as during the extrinsic calibration of the stereo and the reference system. Even a small misalignment of these coordinate systems will result in a measurement inaccuracy that linearly increases for higher distances. Random errors may occur when the measurement system fails to detect the rigid body correctly because of occlusions or reflections.

5.2 Reasoning Algorithm

The following three figures illustrate the results for each recorded sequence. The green-shaded bars represent the real, i. e. the manually labelled, activity. The activity that was determined by the system is marked with an orange line. The ordinate shows the time in seconds, whereas the abscissa shows the activity.

The graphs show that the reasoning algorithm shows results of high quality even for the small bath room in our testing flat. During all sequences, only few minor peaks with a false detection occur. Furthermore, there is a small delay of only a few seconds visible. The delay is caused by a Kalman filter that is used as a tracker in the person detection algorithm. Additionally, the reasoning algorithm is designed as a low-pass filter in order to filter very small peaks. In view of AAL applications with activities

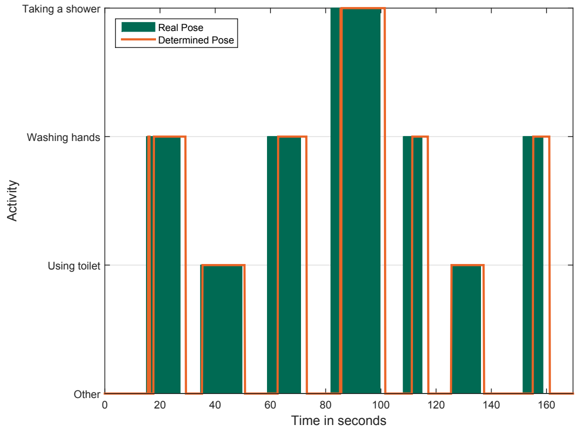

Figure 7: Test person 1: Comparison between labelled (green) and real (orange) activity.

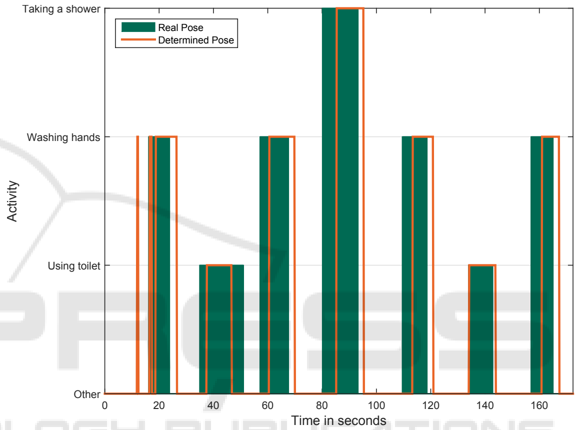

Figure 8: Test person 2: Comparison between labelled (green) and real (orange) activity.

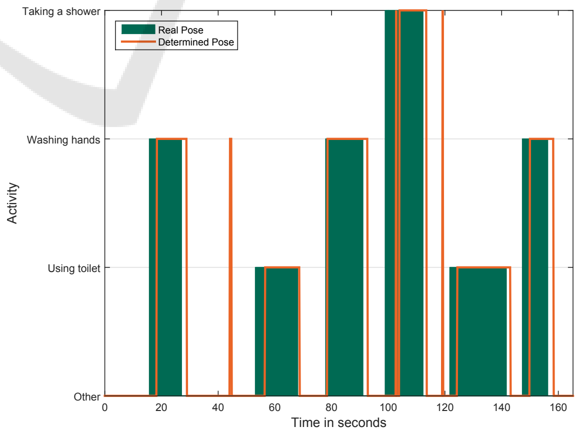

Figure 9: Test person 3: Comparison between labelled (green) and real (orange) activity.

spanning over a time of several minutes, neither the peaks nor the delay have an effect on the functionality. To sum up the results, the reasoning algorithm was able to detect all the activities that were part of the reconstructed morning scene, except only a few small peaks with a false detection. Thus far, no method can be found in literature that addresses activity recognition related to hygiene aspects in AAL environments in a comparable way.

6 CONCLUSIONS AND FUTURE WORK

In this study, an algorithm for reasoning about ADLs has been presented, which evaluates the proximity of relevant objects as well as the person's pose. In order to determine whether the chosen person detection algorithm is theoretically accurate enough – even when objects are close to each other – the accuracy of the algorithm has been analysed with respect to different parameters relevant in AAL scenarios.

As a result, the algorithm has proved sufficient quality. It can consequently be stated that this algorithm is appropriate for our AAL application, where relevant objects are close to each other. Moreover, the accuracy analysis has been designed in an universal fashion, so that other person detection algorithms can be analysed under similar conditions, which facilitates comparisons. The experiments demonstrate that the algorithm is accurate with regard to changing conditions that prevail in AAL environments.

The evaluation of the reasoning algorithm in the testing flat demonstrated that activities normally performed in front of a sink, such as "washing hands, combing, teeth brushing, etc.", "showering" and "using the toilet" could be accurately determined. The tests were conducted in the comparatively small bathroom of our testing flat, so that it can be assumed that our approach would also show good results in larger rooms. We plan to extend the algorithm to more objects in other rooms, in order to recognise further ADLs, such as "preparing a meal", "washing up" or "cooking". In addition, we intend to conduct more tests with probands in our testing flat and to integrate the designed system in real living environments. At this point, we will continue working together with local housing associations and care facilities.

In summary, technical support systems could contribute to a higher quality of care. By giving advice, sending reminding messages to patients and providing care-related information to caring personnel, these modern developments could be beneficial to patients, caring personnel and relatives alike.

ACKNOWLEDGEMENTS

This project is funded by the European Social Fund (ESF).

REFERENCES

Dayangac, E., Wiede, C., Richter, J., and Hirtz, G. (2015). Robust Head-shoulder Detection using Deformable Part-based Models. In *10th International Conference on Computer Vision Theory and Applications (VISAPP-2015)*, pages 236–243.

Harville, M. and Li, D. (2004). Fast, integrated person tracking and activity recognition with plan-view templates from a single stereo camera. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*, volume 2, pages II–398. IEEE.

Pirsiavash, H. and Ramanan, D. (2012). Detecting activities of daily living in first-person camera views. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 2847–2854. IEEE.

Richter, J., Christian, W., and Hirtz, G. (2015). Mobility Assessment of Demented People Using Pose Estimation and Movement Detection. In *ICPRAM Lisbon, Fourth International Conference on Pattern Recognition Applications and Methods*, pages 22–29.

Richter, J., Findeisen, M., and Hirtz, G. (2014). Assessment and Care System Based on People Detection for Elderly Suffering From Dementia. In *Consumer Electronics Berlin (ICCE-Berlin), 2014. ICCE Berlin 2014. IEEE Fourth International Conference on Consumer Electronics*, pages 59–63. IEEE.

Scanaill, C. N., Carew, S., Barralon, P., Noury, N., Lyons, D., and Lyons, G. M. (2006). A Review of Approaches to Mobility Telemonitoring of the Elderly in Their Living Environment. *Annals of Biomedical Engineering*, 34(4):547–563.

Steen, E.-E., Frenken, T., Frenken, M., and Hein, A. (2013). Functional Assessment in Elderlies Homes: Early Results from a Field Trial. *Lebensqualität im Wandel von Demografie und Technik*.

Yous, S., Laga, H., Chihara, K., et al. (2008). People detection and tracking with world-z map from a single stereo camera. In *The Eighth International Workshop on Visual Surveillance-VS2008*.

Zivkovic, Z. (2004). Improved adaptive Gaussian mixture model for background subtraction. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on Pattern Recognition*, volume 2, pages 28–31. IEEE.