Person Identification from Gait Analysis with a Depth Camera at Home

Amandine Dubois[1] and Jean-Pierre Bresciani[1]

Abstract— **The aim of our project is to develop a markerless system to detect falls and evaluate the frailty of elderly people at home. In previous work, we developed an algorithm detecting falls and daily life activities based on depth images provided by Microsoft's Kinect sensor. We also developed another algorithm based on the same features for gait analysis. However, an ambient system installed at home for frailty evaluation should be able to identify the individuals that one wishes to monitor. This paper proposes a method to identify individuals based on the depth images of gait sequences. The gait sequences are detected using previously presented results on activity recognition based on Hidden Markov Models (HMMs). The visibility of the person in the sequence is assessed from the likelihood of the sequence. We propose to perform the identification of the person from her height and gait in sequences in which she walks being fully visible. The gait pattern of the person is modeled using a HMM built from features of the trajectory of the centre of mass. A specific HMM is built for each person to be identified. This approach also allows us to identify unknown individuals who do not correspond to any of the built HMMs. We test the algorithm with 10 known and 2 unknown individuals. The results show that the presented method differentiates accurately enough the unknown and known individuals, and in the last case identifies correctly the individuals. In other words, our algorithm is able to identify the person of interest among other known (family, caregivers) or unknown persons (occasional individuals).**

I. Introduction

One of the major issues of the years to come is the increase of elderly people in loss of autonomy. This work is a contribution to the development of an autonomous ambient system for frailty assessment and fall prevention at home. The system that we propose to install at home is based on the analysis of depth images provided by Microsoft's Kinect sensor. Kinect is an active RGB-D sensor that performs a depth reconstruction from infrared images with the advantage of working at night. This system is designed to detect whether the monitored person has fallen or performs a risky activity such as going up on a chair. The system is also designed to assess possible increase of frailty by analyzing gait parameters, such as the speed of the gait, and evaluating their evolution over time. Previous results ([1], [2]) have shown that it was possible to identify the activity of a person and measure gait parameters from the analysis of the depth images. One of the issues in the development of an autonomous ambient system is the identification of the monitored person. Specifically, the monitored person could share her house with other people or receive visitors such as family members or nurses. Therefore, in order to efficiently monitor the person of interest, one has to be able to identify this person of interest among other people.

[1]Department of Medicine, University of Friburg, Fribourg, Switzerland.

The person identification subject is usually applied to the detection of pedestrians and the surveillance of banks, military installations or airports ([3]). In such cases, the problem is to determine if a person has already been observed over several cameras among many unknown person. Our problem is different. We want to identify the person in the field of view of the sensor to know if she is the person to be monitored. The number of persons to differentiate in a home is usually low, it is possible to make a database of signals recorded on the persons of interest. In the literature, either the face or the gait are often used to differentiate individuals ([4], [5]). Face recognition requires the use of RGB images and is not always possible since the person do not always face the camera. For this reason, we decided to focus on gait parameters for person identification.

Murray [6] and Winter [7] explain that gait is a characteristic and possibly individual trait of a person. The contours of the person [8] or the length of the leg, the torso or the arm [9] can be used to differentiate individuals. Preis *et al.* [10] used a Kinect in a room and conclude that by using basic classifiers naive Bayes, the height, length of legs, torso and the left upper arm are sufficient to identify a person with a probability of 91 %. They used the skeleton extraction provided by Microsoft's SDK.

We propose to perform person identification using the same features that we use to perform activity recognition and gait parameter extraction. In particular, we decided to use the trajectory and speed of the centre of mass along vertical axis and the highest point of the body. These parameters are invariant to the angle of the sensor and can always be extracted, even in the relatively closed and obstructed environment such as an apartment room.

The idea is to identify each known person when she walks, or decide that the person may be unknown. As in previous work on activity recognition [1], walking sequences are given by a Hidden Markov Model (HMM). This HMM can be used to know if the person is fully visible or hidden by an obstacle, which can decrease the probability to recognize the person. In this article, we present likelihood analysis allowing to identify and keep only the sequences in which the person walks being fully visible from head to foot. Then, we propose a model, based on the height and the gait pattern of each individual of interest. The height and gait pattern features allow to built a HMM specific to each person. When a person is detected in field of view of the sensor, the algorithm analyzes her gait with the built individual models. If the detected person fits the model of the person of interest, the person will be considered identified and we can record her gait parameters. If the detected person does not match any model the person will be considered unknown.

This paper is organized as follows. Section II presents previous work on activity recognition and gait parameters extraction as well as the chosen features for representing the person. Section III presents the algorithm allowing to identify the individuals. Section IV presents the evaluation of our algorithm.

II. Background

This work is the continuation of previous work on activity recognition [1] and extraction of the gait parameters [2]. The depth sensor reconstructs the scene as a set of 3D points in the sensor coordinate system. We proposed a simple floor detection algorithm in order to perform a reference change of the 3D points and be independent of the height of the sensor in subsequent processes. We use the least squares method to retrieve the plane passing through an arbitrary number of the lowest points of the scene. The floor coordinate system is calculated online from the plane equation without requiring the use of a specific calibration process. The "running average" method is used for learning the background before extracting the person by subtraction. Then, some features representing the person are calculated. Mainly, the centre of mass is calculated as the average of all the mobile points detected as belonging to the person. A HMM is used to analyze the activity of the person. The model allows to identify eight activities of daily life: walking, lying (on a bed or couch for example), sitting, falling, lying down, squatting, going up on an obstacle (a chair or a footboard for example) and bending. These activities are distinguished using 3 parameters: the vertical position of the centre of mass of a person, the vertical speed and standard deviation of the mobile pixels belonging to the silhouette. The trajectory of the centre of mass along the vertical axis is used to extract the step length, the pace and the speed of the gait. These parameters are obtained by calculating the Euclidean distance and the time between two consecutive vertical maxima of the centre of mass. The maxima corresponds to the position of the supporting foot. The aim is to detect the activities and to analyze the gait of one or several persons of interest.

III. Method

We choose to identify persons based on their gait. The major drawback of this approach in a house environment is the presence of obstacles hiding body part and distorting results of the gait analysis. We need to recognize these occlusion moments to analyze the gait only when the person is fully visible. In this part, we present the method for recognizing visible gait sequences, i.e., the moments in which gait can be analyzed. Then, we present the method used to identify the person of interest.

A. Detection of visible gait sequences

The HMM described in article [1] can recognize the activity performed by a person in the field of view of the sensor. The parameters of the HMM are learnt from a database containing the different activities of the daily life performed by 26 fully visible individuals. The Forward-Backward algorithm is used to know in which state the person is. Phase *Forward* calculates the joint probability $\alpha_t(i)$ for each state i, at each time step t, of observing the sequence $O_{1..t}$ and being in state i. With this calculation we obtain the likelihood which allows to know if the model matches the observations. The likelihood is a measure of the probability of the observation O given the model λ and denoted as $P(O_{1..t}|\lambda)$. The likelihood corresponds to following calculation:

$$Likelihood(t) = \Sigma_{i=1}^{N} \alpha_t(i).$$

We compare the likelihood values obtained when the person walks being hidden by obstacles with those obtained when the person is fully visible. The likelihood values at each time step, performed in conditions "Hidden" and fully "Visible", are shown in Figure 1. The log-likelihood is generally used to represent the results. In the figure, the person walks being visible until 44.1 s, then continues walking being hidden behind chairs. We can see that likelihood values are smaller when the person is hidden than when she is entirely visible. In other words, the model corresponds better to the observations when the person is fully visible. This result is consistent since the model has learnt from a database containing only situations where the person is fully visible. In the situation "Walking - hidden", the transition from visible to hidden appears clearly in the likelihood values. From these likelihood values, it is possible to recognize a situation where the person is hidden using a threshold below which the person will be hidden by an obstacle.

The HMM presented in [1] can be used to identify the walk sequences and the visibility of the person in these sequences can be assessed from the likelihood. This allows us to identify sequences from which we can perform the identification of the person.

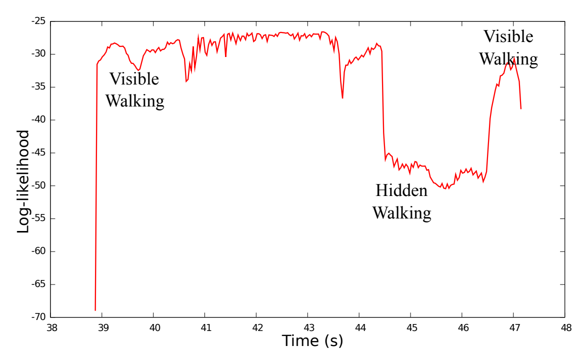


Fig. 1. Likelihood curve for a walk performed by a subject being entirely, then partially visible.

B. Identifying the individuals

Each individual is analyzed from his height and gait pattern. The depth sensor allows to acquire the 3D positions of a set of points corresponding to the silhouette of the person.

The centre of mass is estimated as the geometric centre of the silhouette. The gait of the person is characterized by the vertical trajectory of the centre of mass.

HMMs provide an effective framework for modeling phenomena or processes governed by a dynamic hidden state. Only partial observation on the process is available and the state can be inferred with a certain probability from a sequence of observations. In our case, each individual is represented by a HMM of three states. We use multivariate Gaussian distributions as observation probability density functions. Each individual is characterized by the following features:

- the position of the highest point of the body (corresponding to the top of the head when the person walks);
- the vertical speed of the centre of mass;
- the vertical speed of the centre of mass after applying a low-pass filter (this corresponds to the third coefficient of a Haar transform).

These observations have the advantage to be independent from the angle of the camera. The parameters of each model (the transition matrix, the initial probabilities and the parameters of the observation function) are learnt automatically using the Baum-Welch algorithm from reference gait sequences recorded with several individuals. Inference is made running the algorithm *Forward* on the new gait sequences. This algorithm calculates the likelihood of the sequence for a given model. The most likely individual is the one that corresponds to the model giving the highest likelihood. However, if the likelihood is too small, we consider that the sequence is performed by an unknown individual. A small likelihood means that the sequence does not match any model, i.e., none of the individuals of the database. Because the likelihood depends on the length of the sequence, we use the log-likelihood divided by the length of the sequence for the comparisons. We denote this value as the *normalized log-likelihood* in the next sections.

IV. Evaluation of the algorithm

A. Description of the experiment

Twelve subjects took part in the experiment among which five women and seven men aged from 21 to 54. The study was performed in accordance with the ethical standards laid down in the 1964 Declaration of Helsinki. The experiment was held in a room where the subjects walked, one at a time, perpendicularly to and at a distance of 2 m from the Kinect sensor.

The instruction given to the subjects was to walk "normally" in a straight line. For each subject, twenty walk sequences were recorded. We built the gait model of ten of the twelve subjects. Each model was learnt using 90 % of the subject's sequences. The remaining 10 % of the subject's sequences were used to evaluate the performance of the method. No model was learnt for the remaining two subjects. With the gait sequences of these two subjects, we wanted to evaluate the capacity of the algorithm to detect unknown individuals.

B. Results

We have tested which state ("known" or "unknown") and which name was given by the inference algorithm for 60 sequences not belonging to the training base. We evaluated the algorithm with 20 sequences of known individuals and 40 sequences of unknown individuals. Figure 2 shows the normalized log-likelihood values for the sequences belonging to known individuals (left) as well as unknown individuals (right). We can see that there is only a small overlap between the two boxplots. The normalized log-likelihood value that separates known and unknown individuals is -14.5. In order to test whether log-likelihood values obtained for known and unknown individuals were statistically different, and because the number of available values differed (namely 20 values for known and 40 for unknown), we used a bootstrap procedure. Specifically, we took 2000 random samples of 20 log-likelihood values among the 40 values available, and for each sample, we performed a Wilcoxon Rank Sum test (Mann-Whitney U test) to compare the sample values with the 20 log-likelihood values of known individuals. We used this non-parametric test because the distribution of log-likelihood values deviated from normality. All 2000 p-values computed were below 0.001, indicating that all random samples of unknown normalized log-likelihood values were significantly different from the 20 known values.

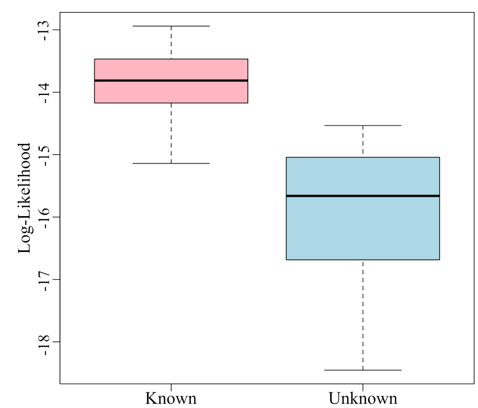


Fig. 2. Normalized log-likelihood values for known and unknown persons.

The result for all the sequences is shown in Table I. This table represents the number of sequences correctly classified. The two rows correspond to the type sequences provided to the algorithm ("known" or "unknown"). The columns indicate the classification provided by the algorithm. We can see that three sequences belonging to known subjects are not identified. The known individuals are classified as unknown.

V. Discussion

The goal of the method is to identify the person of interest so as to further analyze her health condition every day. For that, we need a model identifying correctly this person among other people (such as relatives or nurses). This model should make as few false detections as possible. We consider the error of confusing a person with another more critical than mistaking a known individual for an

RESULTS / SEQUENCES	"Unknown"	"Known"	
		Correct name	Wrong name
20 "Known"	3	17	0
40 "Unknown"	40	0	0

TABLE I
RESULTS OF THE CLASSIFICATION BY THE ALGORITHM OF THE DIFFERENT SEQUENCES.

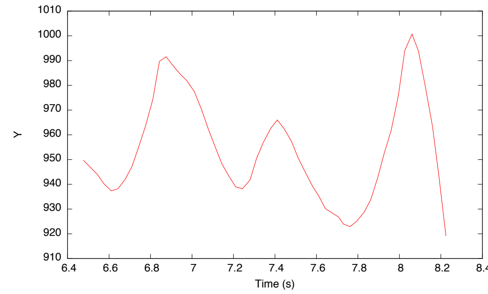
unknown individual. To be sure to identify only the person of interest, we can increase the discrimination threshold between "known" and "unknown". So we will consider the person only when she is detected as close to her model.

Another way to enhance the performance of identification would be to increase the size of the training base. Some subjects were very different from the others and easy to recognize with a high likelihood. But some subjects of the training base were similar which resulted in a smaller likelihood, close to that of unknown subjects. The discrimination threshold between "known" and "unknown" is difficult to determine as shown in Figure 2. By increasing the number of gait sequences, we would have a more representative model of each person. In our case, the algorithm has learnt from only 18 gait sequences by subject and 1 sequence was composed of only 2 to 3 steps depending on the subject.

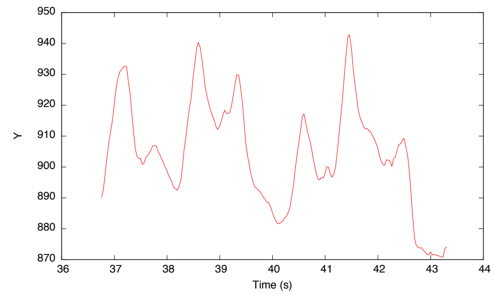
In the future, we would like to test our model integrating elderly people. As compared to young subjects (Figure 3(a)), elderly people usually present trajectories of the centre of mass that are less regular and more specific (Figure 3(b)). We believe that our algorithm is robust enough for identifying elderly people. In our experiment all the subjects were young and thus more similar than what we expect in real conditions.

VI. Conclusion

This work is a part of the development of a system for helping elderly people to stay longer at home. In previous work, we developed an activity recognition algorithm to detect falls and another to extract gait parameters to evaluate the frailty risk. These algorithms should be applied only to the person to monitor. Here, the goal was to identify one or more person of interest among all the people that may enter the living place. From the depth sensor localized at home, we track the persons and detect when they walk with a previously presented algorithm. Likelihood values are used to detect sequences in which the person is not hidden by an obstacle, which increases the probability of identification. Then, we use gait analysis to differentiate individuals. The proposed algorithm uses a HMM for each individuals built from the highest point of the silhouette and features of the trajectory of the centre of mass. To evaluate the accuracy of the recognition provided by the algorithm, we set up an experiment where twelve subjects walked several times in front of the depth sensor. The results show that three individuals are misclassified as unknown of the model. We believe that this method will be accurate to identify the



(a) Walk of a young subject.



(b) Walk of an elderly subject.

Fig. 3. Trajectory of the centre of mass along the vertical axis.

person of interest among others. The next step is to validate the algorithm with elderly people in real conditions.

References

[1] A. Dubois and F. Charpillet, “Human activities recognition with RGB-Depth camera using HMM,” in *35th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2013.

[2] A. Dubois and F. Charpillet, “A gait analysis method based on a depth camera for fall prevention,” in *36th Annual International Conference of the Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2014.

[3] M. Farenzena, L. Bazzani, A. Perina, V. Murino, and M. Cristani, “Person re-identification by symmetry-driven accumulation of local features,” in *Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2360–2367, IEEE, 2010.

[4] W. Zhao, R. Chellappa, P. J. Phillips, and A. Rosenfeld, “Face recognition: A literature survey,” *Computing Surveys*, 2003.

[5] J. Wang, M. She, S. Nahavandi, and A. Kouzani, “A review of vision-based gait recognition methods for human identification,” in *International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, pp. 320–327, IEEE, 2010.

[6] M. Murray, “Gait as a total pattern of movement,” *American Journal of Physical Medicine*, 1967.

[7] D. Winter, *Biomechanics and motor control of human gait: normal, elderly and pathological*. 2nd ed. Waterloo: University of Waterloo Press, 1991.

[8] S. A. Niyogi and E. H. Adelson, “Analyzing and recognizing walking figures in XYT,” in *Computer Vision and Pattern Recognition (CVPR)*, pp. 469–474, IEEE, 1994.

[9] A. F. Bobick and A. Y. Johnson, “Gait recognition using static, activity-specific parameters,” in *Computer Vision and Pattern Recognition (CVPR)*, IEEE, 2001.

[10] J. Preis, M. Kessel, M. Werner, and C. Linnhoff-Popien, “Gait recognition with kinect,” in *1st International Workshop on Kinect in Pervasive Computing*, 2012.