HOME MONITORING OF ELDERLY PEOPLE WITH 3D CAMERA TECHNOLOGY

B. Jansen, R. Deklerck
Vrije Universiteit Brussel, Department of Electronics and Informatics, Belgium

Abstract

We advocate the usage of 3D camera technology in the visual monitoring of elderly people in their home environment or in the nursing home. We illustrate the usage of this novel technology by proposing an extension to existing work on visual fall detection and inactivity detection.

1 Introduction

As people are getting older and older, the fraction of the population which needs dedicated medical care is growing very fast, resulting in medical expenses growing without bond [1,2]. As the social security systems in many countries are highly pressurized by this development, the medical care for elderly now has turned into an important political and economical issue.

One approach to reduce hospital loads and control medical expenses consists of stimulating elderly people to live longer in their natural home environment. Moreover, it is often the explicit wish of the elderly to remain at home. Supportive actions are required to stimulate the elderly to keep on living at home. Mainly family, neighbors, social workers and dedicated home care organizations are facing huge responsibilities.

A wide range of home monitoring technologies are currently being developed or commercialized which support the elderly and their caregivers [3-12]. Currently, work is for instance performed on the automatic monitoring of vital signs, on the detection of drug intake, on the monitoring of physiological exercises and many more.

In this paper, we focus on the automatic detection of falls of the elderly, as risks of falls are very high. Falls are the primary cause of loss of independence, both due to the deteriorating medical conditions, but also because of psychological effects involved. Moreover, elderly are sometimes living in social isolation and hence the possibility exists that falls are not noticed by caregivers for several days.

The most widespread approach for the detection of falls is the usage of accelerometer based devices, often encapsulated in small wrist like devices equipped with a wireless connection. Whenever a significant acceleration towards the ground is measured, the alarm call is transmitted to the base station, which in turn forwards the alarm to a service center using conventional telephone lines. The detection rates are very high with only a small amount of false positives. However, the correct functioning of this type of detectors is depending on the patient's cooperation: patients need to wear the devices always and everywhere in a correct manner. In particular elderly, often suffering from dementia or other mental conditions, sometimes simply forget to wear the device after taking a bath and in many other similar conditions.

An approach which is independent of the cooperation of the patient is the usage of camera technology for the visual detection of falls. This paper argues that the usage of novel 3D camera technology will dramatically improve the performance of existing methods for the monitoring of elderly people. 3D cameras are providing high quality depth images. Compared to stereovision cameras, the accuracy of the depth information is much higher, because it is not depending on identifying correspondences between left and right images. 3D cameras are active sensors which are measuring the time of flight of emitted infrared light. Currently, prices are high while resolution is low. However, 3D cameras are using standard CMOS technology and prices are hence expected to drop to a level comparable to nowadays webcams.

2 Approach

A first approach consists of the explicit detection of the transition from standing to laying on the ground. As falls are however instantaneous and frame rates are sometimes rather low, this approach is often not very robust. Therefore, there is a tendency in literature to focus on the detection of abnormal inactivity, which is more general, but often a strong cue for a fall.

State of the art technologies for visual fall detection [13-15] are based on a regular camera which is placed at the ceiling of the room. An alarm is triggered as soon as inactivity is detected over a predefined time interval. This approach is naturally prone to large amounts of false positives, as the alarm is also triggered when people are sleeping in their beds or napping on the sofa etc. Some authors have introduced the concept of *zones of inactivity*, which are regions in which inactivity will not trigger an alarm. The zones of inactivity can be learnt during the first days of operation of the system. We argue that the zones of inactivity as proposed in the literature will not reduce the amount of false positives, while increasing the amount of false negatives; if you kneel on the ground for searching your wedding ring which just fell on the ground, the alarm will be triggered. On the other hand, if an emergency occurs in a known zone of inactivity, no alarm will be triggered.

By using 3D camera technology, we obtain a variety of important features which correlate with suspicious inactivity and / or falls. These features will be fed into a simple classifier which decides whether or not to ring the fall alarm. Currently, we explore three types of features: (1) the orientation of the body, (2) the height of the subject's head above the ground and (3) an extended version of a context model based on zones of inactivity.

2.1 Body orientation

We explore two different approaches for the determination of the main body orientation: (1) the orientation of the main axis of a simple ellipse fitting technique is used to estimate body orientation in 2D. (2) Opposed to this lightweight method, we implemented a state of the art posture recognition technique in 2D and considered only the orientation of the trunk. The first approach is much faster, while the second approach is believed to provide more reliable results.

2.2 Height of the subject's head above the ground

Using a simple least squares calibration, all points expressed in camera coordinates are mapped onto a coordinate system defined by a corner of the room. As the 3D camera does not provide color information, skin detection algorithms could not be used for the automatic detection of the head in the image. Therefore, we opted for a simple and straightforward approach: using background subtraction, we identify the moving person in the image. After thresholding, we fit an ellipse around the biggest segment which is found. The top of the ellipse is considered to be the position of the head.

2.1 Context model

The implementation of the context model is directly based on the model by McKenna and Nait-Charif proposed in [13,15]. On the one hand, our model is more complex as it takes into account more features, on the other hand it is much simpler as it does not assume specific probabilistic distributions of events.

The context model serves to model the zones of inactivity as introduced before. Using a context model one can assess the probability of an emergency given a specific context. Context is defined by the location, the time and the duration of the observed inactivity. Rather than explicitly programming the context model, it is learnt from experience: the ground floor of the patient's room is divided in small grids. A histogram is associated with every cell in the grid. Every frame, it is investigated how long the subject is already located in this cell on the grid. Those counts on the grid are added to the histogram. After a while, the *d*-th bin of the histogram relates to the probability of inactivity over a period of *f(d)* frames.

3 Experiments

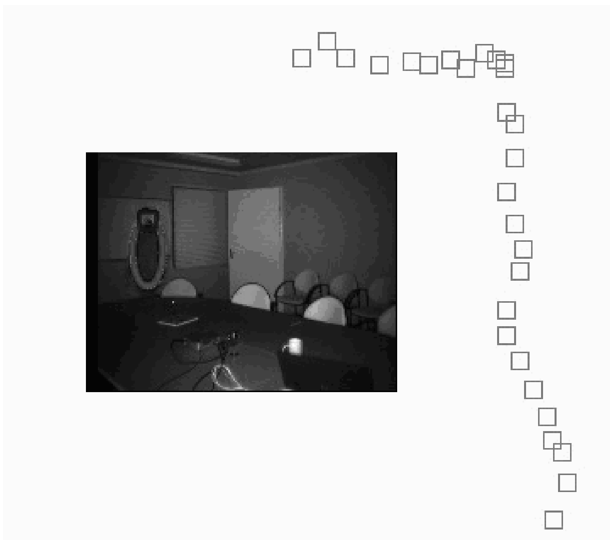

Figure 1: The subject's trajectory in the room.

As the work reported here is work in progress, we currently do not have experiments which compare the overall performance of the fall detector to previous work. Rather, we report on some initial experiments validating the individual components of the system.

3.1 Subject segmentation and position estimation

In figure 1 a sample image is shown which is acquired with the camera. The person is segmented and its trunk and head are identified and marked by the ellipse and the square. The trajectory of the subject is plotted on the ground floor of the room. This type of trajectory will serve as inputs for the context learning model.

3.2 Automatic learning of context models

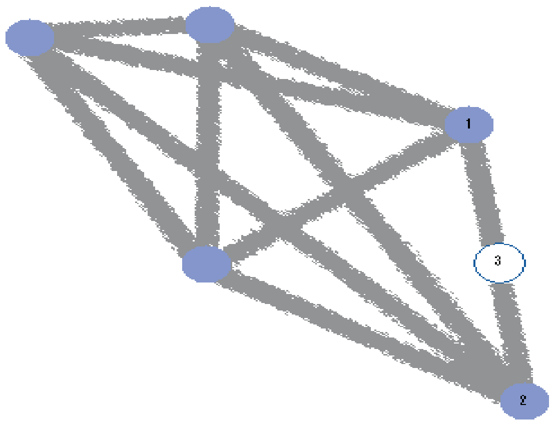

Figure 2: Simulated trajectories between five resting points.

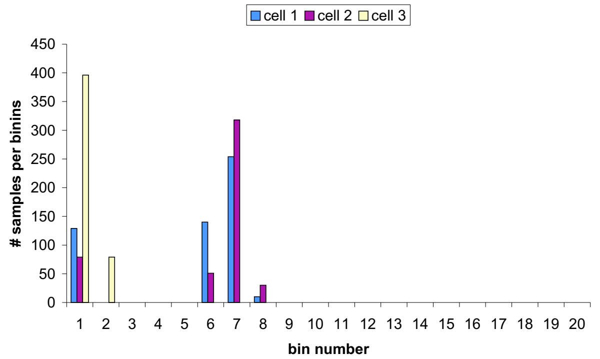

Figure 3: The context model which is learnt for the three cells marked in figure 2.

As the current system is still very preliminary, we were not able yet to perform an experiment in which a real world context model is learnt. In figures 2 and 3, we present results of a computer simulation of a simple room in which a subject is forming simulated trajectories. A square room measuring 500 by 500 units was simulated, with a grid size of 50 by 50. In the room, five random positions were selected as resting points (they represent a sofa, a bed, ...) 1000 trajectories were generated randomly by selecting two different random resting points and connecting them by the shortest path, which was corrupted with normally distributed noise. Every trajectory starts and terminates with 50 time units in the resting point itself. (see figure 2 for a plot of the trajectories). Using those trajectories, a context model was learnt. For three points marked on the trajectories, the resulting histograms are plotted in figure 3. They clearly show a different pattern for the resting points (1 and 3) compared to the point somewhere in the middle of a trajectory (2).

3 Conclusion

We propose a simple fall detection algorithm which is based on three different features: the height of the person's head above the ground floor, the orientation of the body and the detection of irregular inactivity using a context model. Compared to state of the art, our method is more general in the sense that no prior distributions are assumed on the inactivity patterns. Our method is also more robust, since a more advanced context model is used which reduces the amount of false positives without increasing the amount of false negatives.

References

[1] World Health Organization. European health for all database (HFA-DB), 2006.

[2] De Standaard, 26 oktober 2006.

[3] K. C. Sokwoo Rhee, Boo-Ho Yang and H. H. Asada, "The ring sensor: a new ambulatory wearable sensor for twenty-four hour patient monitoring," in Proceedings of the 20th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Hong Kong, 1998.

[4] B.-H. Yang, S. Rhee, and H. H. Asada, "A twenty-four hour tele-nursing system using a ring sensor," in Proc. of 1998 Int. Conf. on Robotics and Automation, Leuven, Belgium, 1998.

[5] M. Ackerman, R. Craft, F. Ferrante, M. Kratz, S. Mandil, and H. Sapci, "Telemedicine technology," Telemedicine Journal and e-Health, vol. 8, no. 1, 2002.

[6] U. Varshney, "Using wireless networks for enhanced monitoring of patients," in Proceedings of the tenth Americas conference on information systems, New York, 2004.

[7] R. Lee, H. Shen, C. Lin, K. Chang, and J. Chen, "Home telecare system using cable television plants, an experimental field trial," IEEE

Transactions on Information Technologies in Biomedicine, vol. 4, no. 1, p. 37 44, 2000.

[8] Pavlopoulos, S. Kyriacou, E. Berler, A. Dembeyiotis, S. Koutsouris, D. , "A novel emergency telemedicine system based on wireless communication technology," IEEE Trans Information Technology in Biomedicine, vol. 2, p. 261267, 1998.

[9] T. Bui and Sankaran, "Group decision and negotiation support in telemedicine: An application of intelligent mobile agents as non-human teleworkers," in Proceedings of the 30th Hawaii International Conference on System Sciences, 1997, pp. 120–129.

[10] K. Z. Haigh and H. A. Yanco, "Automation as caregiver: A survey of issues and technologies," in Proceedings of AAAI 2002 Automation as Caregivers, 2002, pp. 39–53.

[11] F. G. Miskelly, "Assisitive technology in elderly care," Age and Ageing, vol. 30, 2001.

[12] B. U. Toreyin, Y. Dedeoglu, and A. E. Cetin, "Hmm based falling person detection using both audio and video," in IEEE International Workshop on Human-Computer Interaction, Beijing, China, 2005.

[13] S. McKenna and H. Nait-Charif, "Summarising contextual activity and detecting unusual inactivity in a supportive home environment," *Pattern Analysis and Applications*, vol. 7, no. 4, pp. 386–401, 2004.

[14] P. Bromiley, P. Courtney, and N. Thacker, *Design of a Visual System for Detecting Natural Events by the use of an Independent Visual Estimate:a Human Fall Detector*. World Scientific, 2002.

[15] H. Nait-Charif and S. J. McKenna, "Activity summarisation and fall detection in a supportive home environment," in Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.,vol. 4, 2004, pp. 323–326.

[16] Greg Mori and Jitendra Malik, "Estimating human body configurations using shape context matching," in Proceedings of ECCV3, 2002, pp. 666–680.

[17] Greg Mori and Jitendra Malik, Recovering 3d Human Body Configurations Using Shape Contexts, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2006