Fall Detection using Ceiling-mounted 3D Depth Camera

Michal Kepski[2] and Bogdan Kwolek[1]
[1]*AGH University of Science and Technology, 30 Mickiewicza Av., 30-059 Krakow, Poland*
[2]*University of Rzeszow, 16c Rejtana Av., 35-959 Rzeszów, Poland*

Keywords: Video Surveillance and Event Detection, Event and Human Activity Recognition.

Abstract: This paper proposes an algorithm for fall detection using a ceiling-mounted 3D depth camera. The lying pose is separated from common daily activities by a k-NN classifier, which was trained on features expressing head-floor distance, person area and shape's major length to width. In order to distinguish between intentional lying postures and accidental falls the algorithm also employs motion between static postures. The experimental validation of the algorithm was conducted on realistic depth image sequences of daily activities and simulated falls. It was evaluated on more than 45000 depth images and gave 0% error. To reduce the processing overload an accelerometer was used to indicate the potential impact of the person and to start an analysis of depth images.

1 INTRODUCTION

The aim of behavior recognition is an automated analysis (or interpretation) of ongoing events and their context from the observations. Behavior understanding aims at analyzing and recognizing motion patterns in order to produce high-level description of actions. Currently, human behavior understanding is becoming one of the most active and extensive research topics of artificial intelligence and cognitive sciences. The strong interest is driven by broad spectrum of applications in several areas such as visual surveillance, human-machine-interaction and augmented reality. However, this problem is very difficult because of wide range of activities, which can occur in any given context and the considerable variability within particular activities, taking place at different time scales. Despite its difficulty, there has been made a considerable progress in developing methods and models to improve understanding of human behavior. The current research concentrates more on natural settings and moves from generic techniques to specific scenarios and applications (Pantic et al., 2006; Weinland et al., 2011).

One aspect of human behavior understanding is recognition and monitoring of activities of daily living (ADLs). Several methods were proposed to distinguish between activities of daily living and falls (Noury et al., 2008). Falls are a major health risk and a significant obstacle to independent living of the seniors (Marshall et al., 2005) and therefore significant effort has been devoted to ensuring user-friendly assistive devices (Mubashir et al., 2013). However, despite many efforts made to obtain reliable and unobtrusive fall detection, current technology does not meet the seniors' needs. One of the main reasons for non-acceptance of the currently available technology by elderly is that the existing devices generate too much false alarms. This means that some daily activities are erroneously reported as falls, which in turn leads to considerable frustration of the seniors. Additionally, the existing devices do not preserve the privacy and unobtrusiveness adequately.

Most of the currently available techniques for fall detection are based on wearable sensors. Accelerometers or both accelerometers and gyroscopes are the most frequently used sensors in wearable devices for fall monitoring (Noury et al., 2007). However, on the basis of inertial sensors it is hard to separate real falls from fall-like activities (Bourke et al., 2007). The reason is that the characteristic motion patterns of fall also exist in many ADLs. For instance, the squat or crouch also demonstrate a rapid downward motion and in consequence the devices that are only built on inertial sensors frequently trigger false alarms for such simple action. Thus, a lot of research was devoted to detecting of falls using various sensors. Mubashir et al. (Mubashir et al., 2013) done a survey of methods used in the existing fall detection systems. Single CCD camera (Rougier et al., 2006),

Kepski M. and Kwolek B..
Fall Detection using Ceiling-mounted 3D Depth Camera.
DOI: 10.5220/0004742406400647
In *Proceedings of the 9th International Conference on Computer Vision Theory and Applications* (VISAPP-2014), pages 640-647
ISBN: 978-989-758-004-8

multiple cameras (Aghajan et al., 2008), specialized omni-directional ones (Miaou et al., 2006) and stereo-pair cameras (Jansen and Deklerck, 2006) were investigated in vision systems for fall detection. However, the proposed solutions require time for installation, camera calibration and are not cheap.

Recently, Kinect sensor was used in prototype systems for fall detection (Kepski and Kwolek, 2012; Mastorakis and Makris, 2012). It is the world's first low-cost device that combines an RGB camera and a depth sensor. Unlike 2D cameras, it allows 3D tracking of the body movements. Thus, if only depth images are used it preserves the person's privacy. Because depth images are extracted with the support of an active light source, they are largely independent of external light conditions. Thanks to the use of the infrared light the Kinect is capable of extracting the depth images in dark rooms.

In this work we demonstrate how to achieve reliable fall detection with low computational cost and very low level of false positive alarms. A body-worn tri-axial accelerometer is utilized to indicate a potential fall (impact shock). A fall hypothesis is then verified using a ceiling-mounted RGBD camera. We show that the use of such a camera leads to many advantages. For ceiling-mounted RGBD cameras the fully automatic procedure for extraction of the ground plane (Kepski and Kwolek, 2013) can be further simplified as the ground can be determined easily on the basis of farthest subset of points cloud. Moreover, we show that the distance to the ground of the topmost points of the subject undergoing monitoring (head-floor distance) allows us to ignore a lot of false positives alarms, which would be generated if only the accelerometer is used. We show how to extract informative attributes of the human movement, which allows us to achieve very high fall detection ratio with very small ratio of false positives. To preserve the privacy of the user the method utilizes only depth images acquired by the Kinect device.

2 FALL DETECTION USING SINGLE INERTIAL SENSOR

A lot of different approaches have been investigated to achieve reliable fall detection using inertial sensors (Bourke et al., 2007). Usually, a single body-worn sensor (tri-axial accelerometer or tri-axial gyroscope, or both embedded in an inertial measurement unit) is used to indicate person fall. The most common method consists in using a tri-axial accelerometer. At any point in time, the output of the accelerometer is a linear combination of two components, that is the acceleration component due to gravity and the acceleration component due to bodily motion. The first approach to fall detection using accelerometry was done by Williams et al. (Williams et al., 1998). The accelerometer-based algorithms simply raise the alarm when a certain threshold value of the acceleration is reached. In practice, there exist many problems inherently connected with such kind of algorithms, including lacking of adaptability.

After accidental fall, the individual's body is usually in a different orientation than before the fall, and in consequence the resting acceleration in three axes is different from the acceleration in the pose before the fall. In (Chen et al., 2005) an accelerometer-based algorithm, relying on change in body orientation has been proposed. If the root sum vector of the three squared accelerometer outputs exceeds a certain threshold, it signals a potential fall. After detecting the impact, the orientation change is calculated over one second before the first impact and two seconds after the last impact using the dot product of the acceleration vectors. The angle change can be set arbitrarily based on empirical data, as suggested by the authors. However, despite many efforts to improve this algorithm, it does not provide sufficient discrimination between real-world falls and ADLs.

In this work we assume that a fall took place if the signal upper peak value (UPV) from the accelerometer is greater than $2.5g$. We investigated also the usefulness of the orientation change for different thresholds. However, the potential of this attribute seems to be small, particularly in actions consisting in lowering the body closer to the ground by bending the knees, for instance while taking or putting the object from the floor. A review of the relevant literature demonstrates that for a single inertial device the most valuable information can be obtained if the device is attached near the centre of mass of the subject. Therefore, the accelerometer was attached near the spine on the lower back using an elastic belt around the waist.

Compared to vision-based motion analysis systems, wearable sensors offer several advantages, particularly in terms of cost, ease of use and, most importantly, portability. Currently available smartphones serve not only as communication and computing devices, but they also come with a rich set of embedded sensors, such as an accelerometer, gyroscope and digital compass. Therefore, they were used in prototype systems for fall detection. However, despite many advantages, the inertial sensors-based technology does not meet the seniors' needs, because some activities of daily living are erroneously reported as falls.

3 FALL DETECTION USING CEILING-MOUNTED RGBD CAMERA

Various types of cameras were used in vision systems for fall detection (Rougier et al., 2006; Aghajan et al., 2008; Miaou et al., 2006; Jansen and Deklerck, 2006). However, the video technology poses a major problem of acceptance by seniors as it requires the placement of the cameras in private living quarters, and especially in the bedroom and the bathroom, with consequent concerns about privacy. The existing systems require time for installation, camera calibration and are not cheap. Recently, Kinect sensor was demonstrated to be very useful in fall detection (Kepski and Kwolek, 2012; Mastorakis and Makris, 2012). In (Kepski and Kwolek, 2013) we demonstrated an automatic method for fall detection using Kinect sensor. The method utilizes only depth images acquired by a single device, which is placed at height about 1 m from the floor. It determines the parameters of the ground plane equation and then it calculates the distance between the person's centroid to the ground. In contrast, in this work we employ a ceiling-mounted Kinect sensor. In subsequent subsections we demonstrate that such a placement of the sensor has advantages and can lead to simplification of the algorithms devoted to distinguishing the accidental falls from ADLs.

3.1 Person Detection in Depth Images

Depth is very useful cue to attain reliable person detection since humans may not have consistent color and texture but have to occupy an integrated region in space. Kinect combines structured light with two classic computer vision techniques, namely depth from focus and depth from stereo. It is equipped with infrared laser-based IR emitter, an infrared camera and a RGB camera. The IR camera and the IR projector compose a stereo pair with a baseline of approximately 75 mm. A known pattern of dots is projected from the IR laser emitter. These specs are captured by the IR camera and compared to the known pattern. Since there is a distance between laser and sensor, the images correspond to different camera positions, and that in turn allows us to use stereo triangulation to calculate each spec depth. The field of view is 57° horizontally and 43° vertically, the minimum measurement range is about 0.6 m, whereas the maximum range is somewhere between 4-5 m. It captures the depth and color images simultaneously at a frame rate of about 30 fps. The RGB stream has size 640×480 and 8-bit for each channel, whereas the depth stream is 640×480 resolution and with 11-bit depth.

The software called NITE from PrimeSense offers skeleton tracking on the basis of RGBD images. However, this software is targeted for supporting the human-computer interaction, and not for detecting the person fall. In particular, it was developed to extract and to track persons in front of the Kinect. Therefore, we employ a person detection method (Kepski and Kwolek, 2013), which with low computational cost extracts the individual on images acquired by a ceiling-mounted Kinect. Another rationale for using such a method is that it can delineate the person in real-time on PandaBoard (Kepski and Kwolek, 2013), which is a low-power, low-cost single-board computer development platform for mobile applications.

The person was delineated on the basis of a scene reference image, which was extracted in advance and then updated on-line. In the depth reference image each pixel assumes the median value of several pixels values from the earlier images. In the setup phase we collect a number of the depth images, and for each pixel we assemble a list of the pixel values from the former images, which is then sorted in order to extract the median. Given the sorted lists of pixels the depth reference image can be updated quickly by removing the oldest pixels and updating the sorted lists with the pixels from the current depth image and then extracting the median value. We found that for typical human motions, good results can be obtained using 13 depth images (Kepski and Kwolek, 2013). For Kinect acquiring the images at 25 Hz we take every fifteenth image.

In the detection mode the foreground objects are extracted through differencing the current depth image from such a depth reference map. Afterwards, the person is delineated through extracting the largest connected component in the thresholded difference between the current depth map and the reference map.

Figure 1 illustrates extraction of the person using the updated depth reference image of the scene. In the first and the second row are depicted example color images and their corresponding depth images, in the third row are shown the difference images between the depth images and the depth reference images, which are shown in the last row. As we can observe, if the layout of the scene changes, for instance due to the shift a chair to another location, see Fig. 1 b,f, the depth difference images temporally contain not only the person, but also the shifted object, see Fig. 1 j,k. As we can observe in the subsequent frames, on the basis of the refreshed depth-reference image, see also Fig. 1 n), which now contains the moved chair, the algorithm extracts only the

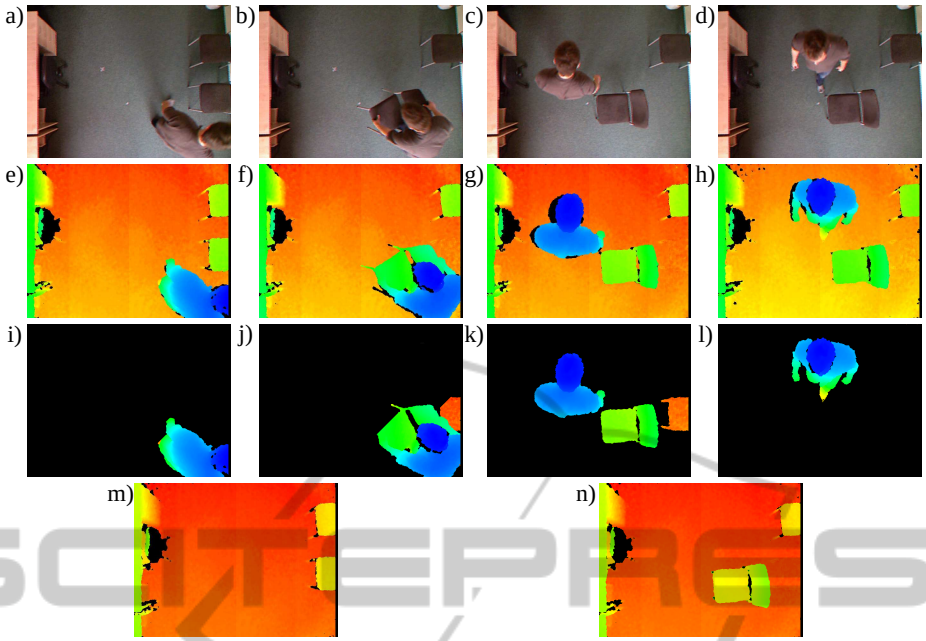

Figure 1: Extraction of the person on the basis of the updated depth reference image. Input images a-d), corresponding depth images e-h), difference between depth images and depth-reference images i-l), updated on-line depth-reference images m-n).

person undergoing monitoring, see Fig. 1 l.

3.2 Lying Pose Recognition

The recognition of lying pose has been achieved using a classier trained on features representing the extracted person in the depth images. A data-set consisting of images with normal activities like walking, taking or putting an object from floor, bending right or left to lift an object, sitting, tying laces, crouching down and lying has been composed in order to train classifiers responsible for checking whether a person is lying on the floor. Thirty volunteers with age under 28 years attended in preparation of the data-set. In total 100 images representing typical activities of daily living were selected and then utilized to extract the features.

In most vision-based algorithms for lying pose recognition the ratio of height to width of the rectangle surrounding the subject is utilized. In contrast, in our algorithm we employ the ratio of major length to major width, which is calculated on the basis of the binary image I representing the person undergoing monitoring. The major length and width (eigenvalues) were calculated in the following manner (Horn, 1986):

$$l = 0.707\sqrt{(a+c)+\sqrt{b^2+(a-c)^2}}$$
$$w = 0.707\sqrt{(a+c)-\sqrt{b^2+(a-c)^2}} \quad (1)$$

where

$$a = \frac{M_{20}}{M_{00}} - x_c^2,\ \ b = 2(\frac{M_{11}}{M_{00}} - x_c y_c),\ \ c = \frac{M_{02}}{M_{00}} - y_c^2,$$
$$M_{00} = \sum_x \sum_y I(x,y),\ \ M_{11} = \sum_x \sum_y xyI(x,y),$$
$$M_{20} = \sum_x \sum_y x^2 I(x,y),\ \ M_{02} = \sum_x \sum_y y^2 I(x,y),$$

whereas x_c, y_c were computed as follows: $x_c = \frac{\sum_x \sum_y xI(x,y)}{\sum_x \sum_y I(x,y)},\ \ y_c = \frac{\sum_x \sum_y yI(x,y)}{\sum_x \sum_y I(x,y)}$.

In total, we utilized three features:

- H/H_{max} - a ratio of head-floor distance to the height of the person
- *area* - a ratio expressing the person's area in the image to the area at assumed distance to the camera
- l/w - a ratio of major length to major width, calculated on the basis of (1).

Figure 2 depicts a scatter plot matrix for the employed attributes, in which a collection of the scatter plots is organized in a two-dimensional matrix simultaneously to provide correlation information among the attributes. In a single scatter plot two attributes are projected along the x-y axes of the Cartesian coordinates. As we can observe on the discussed plots, the overlaps in the attribute space are not too significant. We considered also another attributes, for instance, a filling ratio of the rectangles making up the person's bounding box. The worth of the considered features

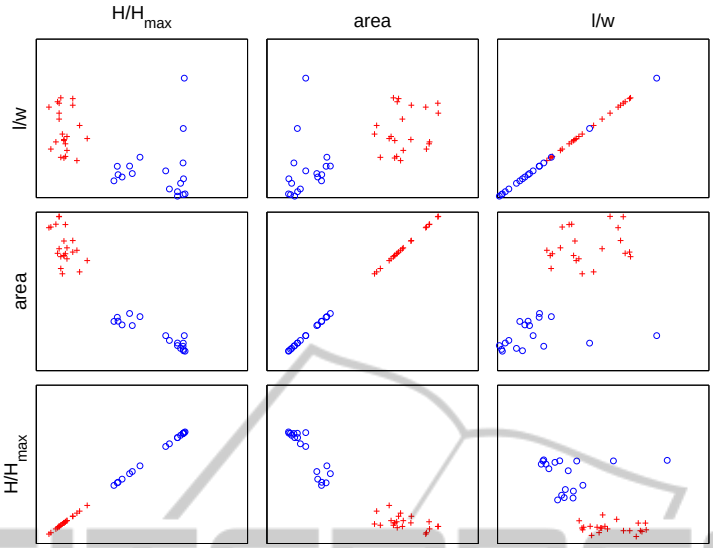

Figure 2: Multivariate classification scatter plot for features used in lying pose recognition.

was evaluated on the basis of the information gain (Cover and Thomas, 1992), which measures the dependence between the feature and the class label. In assessment of the discrimination power of the considered features and selecting most discriminative ones we used `InfoGainAttributeEval` function from the Weka (Cover and Thomas, 2005). The features selected in such a way were then utilized to train classifiers responsible for distinguishing between daily activities and accidental falls.

3.3 Dynamic Transitions for Fall Detection

In the previous subsection we demonstrated how to detect a fall on the basis of lying pose recognition, through analysis the content of single depth image. However, the human fall is a dynamic process, which arises in relatively short time. The relevant literature suggests that the fall incident takes approximately 0.4 s to 0.8 s. During a person fall there is an abrupt change of head-floor distance with accompanying change from a vertical orientation to a horizontal one. The distance of the person's centroid to the floor also changes significantly and rapidly during the accidental fall. Thus, depth image sequences can be used to extract features expressing motion patterns of falls. Using depth image sequences we can characterize the motion between static postures, and in particular between the pose before the fall and the lying pose. In particular, motion information allows us to determine whether a transition of the body posture or orientation is intentional or not.

In the depth image sequences with ADLs as well as person falls we analyzed the feature ratios H/H_{max}, area and l/w, and particularly their sudden changes that arise during falling, e.g. from standing to lying. To reduce the ratio of false positive alarm of the system relying only on features extracted in a single depth image, we introduced a feature reflecting change of the H/H_{max} over time. That means that aside from the static postures our algorithm also employs information from dynamic transitions, i.e. motions between static postures. In particular, this allows us to distinguish between intentional lying postures and accidental falls. Our experimental results show that the ratio of $H(t)/H(t-1)$ for $H(t)$ determined in the moment of the impact and $H(t-1)$ determined one second before the fall is very useful to distinguish the fall from many common ADLs. Figure 3 depicts change over time of the discussed features for typical activities, consisting in walking, crouching down, knelling down, sitting, bending, standing and falling. A peak in walking phase that is seen in the plot of H/H_{max} is due to a raise of the hand. As we can observe, for ceiling mounted camera the area ratio changes considerably in the case of the fall. In the case of a person fall, the peak value of $H(t)/H(t-1)$ is far below one.

The ratio $H(t)/H(t-1)$ can be determined using only vision-based techniques, i.e. through analysis of pairs of depth images. The Kinect microphones can be used to support the estimation of the moment of the person impact. Inertial sensors are particularly attractive for such a task because currently they are embedded in many smart devices. With the help of such fall indicators the value $H(t)$ can be computed at low computational cost.

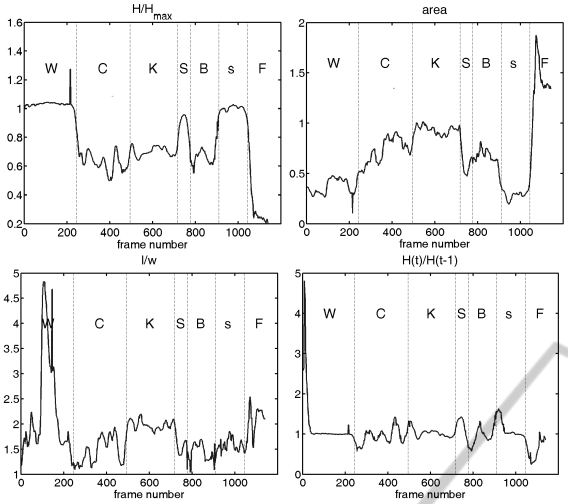

Figure 3: Curves of feature change during the ADLs and falls. W-walking, C-crouching down, K-knelling down, S-sitting, B-bending, s-standing, F-falling.

4 EXPERIMENTAL RESULTS

At the beginning of the experimental validation of the proposed algorithm we assessed the usefulness of an accelerometer as an indicator of potential fall. The actors performed common daily activities consisting in walking, taking or putting an object from floor, bending right or left to lift an object, sitting, tying laces, crouching down and lying. The accelerometer was worn near the spine on the lower back using an elastic belt around the waist. The motion data were acquired by a wearable smart device (x-IMU or Sony PlayStation Move) containing accelerometer and gyroscope sensors. Data from the device were transmitted wirelessly via Bluetooth and received by a laptop computer. Figure 4 depicts the histogram of the UPV values for the carried out activities during half an hour experiment. As we can see, the values $2.5 - 3g$ were exceeded several times. This means that within half an hour of typical person's activity a considerable number of false alarms would be generated if the fall detection was carried out only on the basis of the accelerometer. In particular, we noticed that all fall-like activities were indicated properly. Thus, the accelerometer can be used as reliable indicator of the person impact. In consequence, the computational overhead can be reduced significantly, as the depth image analysis can only be performed in case of signaling with low cost the potential fall. On the other hand, the accelerometer acknowledged its usefulness in the update of the depth reference image. For that reason, for a person at rest no update of the depth reference image is needed. Overall, the accelerometer acknowledged its usefulness in activity summarization. However, typical accelerometer can be inconvenient to be worn during the sleep and this in turn results in the lack of ability to monitor the critical phase of getting up from the bed.

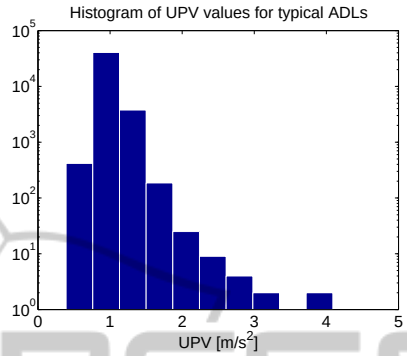

Figure 4: Histogram of UPV for typical ADLs registered during half fan hour experiment.

The algorithm for lying pose recognition has been evaluated in 10-fold cross-validation using features discussed in Section 3.2. We trained k-NN, SVM, KStar and multilayer perceptron classifiers responsible for checking whether the person is lying on the floor. All falls were distinguished correctly. We also trained and evaluated our fall detection algorithm on images acquired by Kinect with Nyko zoom range reduction lens. Figure 5 illustrates the images of the same scene, which were acquired by Kinect with and

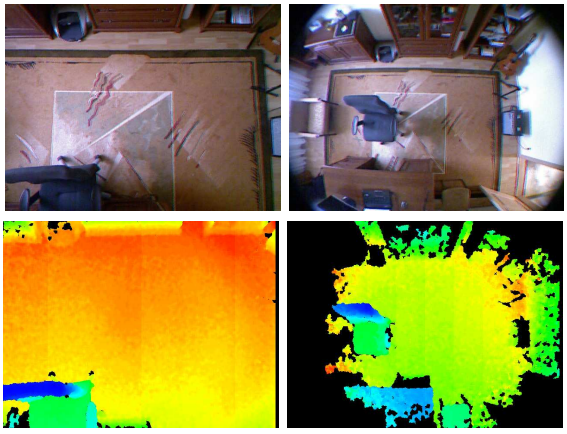

Figure 5: The same scene seen by Kinect (left) and by Kinect with Nyko zoom range reduction lens (right). In upper row are shown color images, whereas in bottom one are corresponding depth images.

without the Nyko lens. As we can notice, owing to the Nyko lens the monitored area is far larger. In particular, using the Kinect sensor mounted at height of about 2.5 m it is possible to encompass and to shot

the ground of the whole room. Five young volunteers attended in the evaluation of the algorithm for fall detection through analysis of lying pose. For the testing of the classifiers we selected one hundred of images with typical human actions of which half depicted persons lying on the floor. It is worth noting that in the discussed experiment all lying poses were distinguished properly from daily activities.

Figure 6 illustrates change of H/H_{max} over time for an intentional fall and an intentional lying on the floor. As we can observe, the time needed for approaching the value $2.5g$ is different, i.e. approaching the value for which the considerable part of the body has a contact with the floor. The discussed plot justifies the usefulness of $H(t)/H(t-1)$ feature in distinguishing the intentional lying on the floor from the accidental falling.

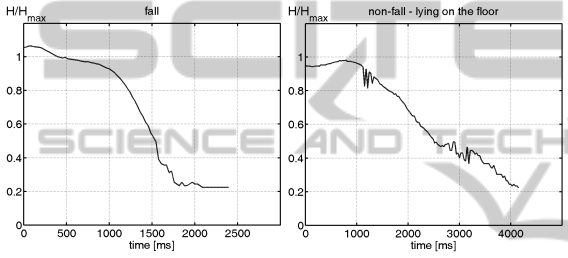

Figure 6: Curves of H/H_{max} change over time for an intentional fall and lying on the floor.

The lying pose detector was evaluated by five students. The experiments demonstrated that the algorithm has very high detection ratio, slightly smaller than 100%. However, the students found a couple of poses that were similar to lying poses, and which were recognized as lying poses. The use of the inertial sensor as indicator of the fall reduces considerably the possibility of false alarms of the system-based on lying pose detection at the cost of obtrusiveness.

In the next stage of the experiments we combined the detector of lying pose with a detector using the proposed temporal feature. Through combining the discussed detectors we achieved reliable fall detection with very small false alarm ratio. As expected, the students found a small amount of daily actions, which in some circumstances can lead to false alarm, mainly due to imperfect detection of the moment of the body impact on the basis of only vision techniques.

Having on regard that fall detection system should have negligible false alarm ratio as well as low consumption power, in the extended evaluation of the system we employed an inertial sensor to sense the impact, and particularly to determine more precisely the moment at which it takes place. In response to potential fall signaled by the inertial sensor the lying pose classifier is executed to verify if the depth image contains a person in a lying pose. If yes, the classifier using the temporal feature is executed to check whether this was a dynamical person action. In this case, the value $H(t)$ is calculated at the moment of the impact, i.e. at time in which the acceleration exceeds the value of $2.5\ g$.

The depth camera-based system for fall detection has been tested using image sequences with daily activities and simulated-falls performed by young volunteers. Intentional falls were performed in a room towards a carpet with thickness of about 2 cm. A comprehensive evaluation showed that the system has high accuracy of fall detection and very low level (as low as 0%) of false alarms. In half an hour experiment, in which more than 45000 depth images with fall-like activities and simulated falls, all daily activities were distinguished from falls like daily activities. The image sequences with the corresponding acceleration data are available at: http://fenix.univ.rzeszow.pl/∼mkepski/ds/uf.html. The datasets contain 66 falls of which half of them concerned persons falling of a chair. No false alarm was reported and all intentional falls were indicated appropriately. The classification was done by k-NNs. The algorithm has also been tested in an office, where the simulated falls were performed onto crash mats. In a few minutes' long video with walking, sitting/standing, executed 20 times, crouching, executed 10 times, taking or putting an object from floor, repeated 10 times, and 20 intentional falls, all falls were recognized correctly using only depth images.

The experimental results are very promising. In order to be accepted by seniors, a system for fall detection should be unobtrusive and cheap, and particularly it should have almost null false alarm ratio as well as preserve privacy. The proposed algorithm for fall detection was designed with regard to factors mentioned above through careful selection of its ingredients. In comparison to existing systems (Mubashir et al., 2013), it has superior false alarm ratio with almost perfect fall detection ratio, and meets the requirements that should have a system to be accepted by the seniors. Moreover, due to low computational demands, the power consumption is acceptable for seniors. Having on regard our previous implementation of a fall detection system on PandaBoard as well as the computational demands of the current algorithm it will be possible to implement the algorithm on new PandaBoard and to execute the algorithm in real-time. The advantage will be low power consumption and easy setup of the system.

The depth images were acquired by the Kinect sensor using OpenNI. The system was implemented in C/C++ and runs at 30 fps on 2.4 GHz I7 (4 cores,

Hyper-Threading) notebook, powered by Windows. For ceiling-mounted Kinect at the height of 2.6 m from the floor the covered area is about 5.5 m^2. With Nyko lens the area covered by the camera is about 15.2 m^2. The most computationally demanding operation is extraction of the depth reference image of the scene. For images of size 640×480 the computation time needed for extraction of the depth reference image is about 9 milliseconds.

5 CONCLUSIONS

In this work we demonstrated an approach for fall detection using ceiling-mounted Kinect. The lying pose is separated from common daily activities by a classifier, trained on features expressing head-floor distance, person area and shape's major length to width. To distinguish between intentional lying postures and accidental falls the algorithm employs also motion between static postures. The experimental validation of the algorithm that was conducted on realistic depth image sequences of daily activities and simulated falls shows that the algorithm allows reliable fall detection with low false positives ratio. On more than 45000 depth images the algorithm gave 0% error. To reduce the processing overload an accelerometer was used to indicate the potential impact of the person and to start analysis of depth images. The use of accelerometer as indicator of potential fall simplifies computation of the motion feature and increases its reliability. Owing the use only depth images the system preserves privacy of the user and works in poor lighting conditions.

ACKNOWLEDGEMENTS

This work has been supported by the National Science Centre (NCN) within the project N N516 483240.

REFERENCES

Aghajan, H., Wu, C., and Kleihorst, R. (2008). Distributed vision networks for human pose analysis. In Mandic, D., Golz, M., Kuh, A., Obradovic, D., and Tanaka, T., editors, *Signal Processing Techniques for Knowledge Extraction and Information Fusion*, pages 181–200. Springer US.

Bourke, A., O'Brien, J., and Lyons, G. (2007). Evaluation of a threshold-based tri-axial accelerometer fall detection algorithm. *Gait & Posture*, 26(2):194–199.

Chen, J., Kwong, K., Chang, D., Luk, J., and Bajcsy, R. (2005). Wearable sensors for reliable fall detection. In *Proc. of IEEE Int. Conf. on Engineering in Medicine and Biology Society (EMBS)*, pages 3551–3554.

Cover, T. M. and Thomas, J. A. (1992). *Elements of Information Theory*. Wiley.

Cover, T. M. and Thomas, J. A. (2005). *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, San Francisco, 2nd edition.

Horn, B. (1986). *Robot Vision*. The MIT Press, Cambridge, MA.

Jansen, B. and Deklerck, R. (2006). Context aware inactivity recognition for visual fall detection. In *Proc. IEEE Pervasive Health Conf. and Workshops*, pages 1–4.

Kepski, M. and Kwolek, B. (2012). Fall detection on embedded platform using Kinect and wireless accelerometer. In *13th Int. Conf. on Computers Helping People with Special Needs*, LNCS, vol. 7383, pages II:407–414. Springer-Verlag.

Kepski, M. and Kwolek, B. (2013). Unobtrusive fall detection at home using kinect sensor. In *Computer Analysis of Images and Patterns*, volume 8047 of *LNCS*, pages I:457–464. Springer Berlin Heidelberg.

Marshall, S. W., Runyan, C. W., Yang, J., Coyne-Beasley, T., Waller, A. E., Johnson, R. M., and Perkis, D. (2005). Prevalence of selected risk and protective factors for falls in the home. *American J. of Preventive Medicine*, 8(1):95–101.

Mastorakis, G. and Makris, D. (2012). Fall detection system using Kinect's infrared sensor. *J. of Real-Time Image Processing*, pages 1–12.

Miaou, S.-G., Sung, P.-H., and Huang, C.-Y. (2006). A customized human fall detection system using omni-camera images and personal information. *Distributed Diagnosis and Home Healthcare*, pages 39–42.

Mubashir, M., Shao, L., and Seed, L. (2013). A survey on fall detection: Principles and approaches. *Neurocomputing*, 100:144 – 152. Special issue: Behaviours in video.

Noury, N., Fleury, A., Rumeau, P., Bourke, A., ÓLaighin, G., Rialle, V., and Lundy, J. (2007). Fall detection - principles and methods. In *Int. Conf. of the IEEE Eng. in Medicine and Biology Society*, pages 1663–1666.

Noury, N., Rumeau, P., Bourke, A., ÓLaighin, G., and Lundy, J. (2008). A proposal for the classification and evaluation of fall detectors. *IRBM*, 29(6):340 – 349.

Pantic, M., Pentland, A., Nijholt, A., and Huang, T. (2006). Human computing and machine understanding of human behavior: a survey. In *Proc. of the 8th Int. Conf. on Multimodal Interfaces*, pages 239–248.

Rougier, C., Meunier, J., St-Arnaud, A., and Rousseau, J. (2006). Monocular 3D head tracking to detect falls of elderly people. In *Int. Conf. of the IEEE Engineering in Medicine and Biology Society*, pages 6384–6387.

Weinland, D., Ronfard, R., and Boyer, E. (2011). A survey of vision-based methods for action representation, segmentation and recognition. *Comput. Vis. Image Underst.*, 115:224–241.

Williams, G., Doughty, K., Cameron, K., and Bradley, D. (1998). A smart fall and activity monitor for telecare applications. In *IEEE Int. Conf. on Engineering in Medicine and Biology Society*, pages 1151–1154.