Fall Detection in indoor environment with Kinect sensor

Vitoantonio Bevilacqua, Nicola Nuzzolese, Donato Barone

Polytechnic of Bari
Department of Electrical and Information Engineering
(DEI)
Bari, Italy
vitoantonio.bevilacqua[at]poliba.it,
eng.donato.barone[at]gmail.com,
n.nuzzolese[at]gmail.com;

Michele Pantaleo, Marco Suma, Dario D'Ambruoso, Alessio Volpe A.M.T. Services s.r.l.

A.M.T. Services s.r.l.

Bari, Italy
michele.pantaleo[at]amtservices.it,
marco.suma[at]amtservice.it,
dario.dambruoso[at]amtservices.it,
alessiovolpe[at]msn.com;

Claudio Loconsole, Fabio Stroppa
PERCRO Laboratory
Scuola Superiore Sant'Anna
Pisa, Italy
fbs88italy[at]gmail.com,
c.loconsole[at]sssup.it.

Abstract - Falls are one of the major risks of injury for elderly living alone at home. Computer vision-based systems offer a new, low-cost and promising solution for fall detection. This paper presents a new fall-detection tool, based on a commercial RGB-D camera. The proposed system is capable of accurately detecting several types of falls, performing a real time algorithm in order to determine whether a fall has occurred. The proposed approach is based on evaluating the contraction and the expansion speed of the width, height and depth of the 3D human bounding box, as well as its position in the space. Our solution requires no pre-knowledge of the scene (i.e. the recognition of the floor in the virtual environment) with the only constraint about the knowledge of the RGB-D camera position in the room. Moreover, the proposed approach is able to avoid false positive as: sitting, lying down, retrieve something from the floor. Experimental results qualitatively and quantitatively show the quality of the proposed approach in terms of both robustness and background and speed independence.

Keywords—fall detection, RGB-D cameras, depth sensor, older people , kinect

I. INTRODUCTION

The fall of frail people (elderly or people suffering from motor disabilities) is both a clinical and a social problem. Falls, often due to unstable equilibrium, cause high morbidity and mortality among elderly. As far as the resulting trauma, it often is a source of a potential disability with repercussions on the psychological sphere of the subject himself, especially in case of aged people.

In the U.S., nearly three-quarters of the deaths due to accidental falls occur among over-65s (which represent the 13% of the entire population). Only half of the hospitalized

elderly who have suffered a fall, have survived more than one year.

In the corresponding literature, several studies, such as [1], have described the epidemiology of elderly falls in different contexts. Although the majority of these falls do not cause major injuries, about 5% causes a fracture or requires hospitalization of the person. Moreover, the rates of falls and associated complications significantly increase with age and double in people older than 75 years. TABLE I illustrates the main causes (in percentage) of falls in the elderly [5] over 12 fall study works for a total of 3,628 falls.

TABLE I. MAIN CAUSES OF FALLS IN ELDERLY OVER A TOTAL OF 3,628 REGISTERED FALLS

Causes	Average Percentage
accidental - environment related	31 %
gait disturbance, decreased muscle strength	17 %
dizziness, vertigo	13 %
breakdown	9 %
commotion	5 %
postural hypotension	3 %
visual disturbance	2 %
syncopation	0,3 %
other specified causes	15 %
unknown causes	5 %

The category "other specified causes" includes arthritis, acute illness, drugs, alcohol, pain, epilepsy and fell out of bed.

Other studies show that the context of use of this application is characterized by those places where the fall event is more frequent, such as homes, hospitals, assisted living environments and nursing homes. Following the fall detection

of a subject, would eventually be necessary to communicate it outside from the accident site, in order to assist the elder.

In these indoor environments, it is essential to provide monitoring services that use a non-invasive sensors, and to pay particular attention to the issue of privacy protection, innovative services of surveillance and monitoring.

A common requirement for automatic fall detection system is the speed of execution of fall detection algorithms. In fact, they should allow to detect a potentially dangerous situation for fall in constrained environments. The existing approaches addressing this issue can be mainly subdivided into two classes:

- invasive methods, such as motion sensors worn by the subject [2, 8] or floor vibration detectors [3] which usually require a complex setup;
- computer vision-based systems providing a low cost and ease to use solution.

In particular, referring to the latter systems, RGB-D cameras can be fully exploited in fall detection scenarios, rather than common 2D cameras because they provide a 3D point cloud featuring a higher degree of information required for analyzing the human body dynamics for fall detection.

Main research work based on computer vision-based system are that of Rougier *et al.* [10], in which falls are detected calculating the subject's speed in a 3D environment using the human centroid differential height (with respect to the ground). Planinc *et al.* [7] calculate the main orientation of the subject's body in 3D space by using the skeleton information in addition of the ground reference. The mentioned methods have been improved by the Mastorakis *et al.* [9], that use the subject 3D bounding box without referring it to the ground. Our work, conducted within the MET-AAL project with AMT Services s.r.l. and Agilex s.r.l. companies, is inspired and starts from the latter one, proposing an innovative solution for reducing misclassification.

The Italian project "METhodology and instruments for pervasive model in Ambient Assisted Living" (MET-AAL), funded by Apulia Region (POR FESR 2007-2013 Obiettivo Convergenza - ASSE I - Linea 1.2 - Azione 1.2.4 "Investiamo nel vostro futuro" – Leader Matrix s.p.a. company, Conversano, Bari, Italy) aims to define and create a platform for ambient intelligence featuring a pervasive technology useful to provide support, assistance and services to people with reduced autonomy. Indeed, the project tries to provide a technological solutions to exploit the residual abilities of the elderly persons and patients affected by chronic disease. In particular, in this paper we introduce a fall detection algorithm designed and experimentally tested as a module of home automation platform of the MET-AAL project. The specific aim of the algorithm is to provide an accurate and low-cost solution to detect falls in order to be sustainable even in public hospitals or in domestic houses.

II. THE PROPOSED APPROACH

The proposed system is composed of two main subsystems: the RGB-D camera and the fall detector module based on

available libraries for camera management and computer vision procedures. Next sections are dedicated to the description of the two subsystems (Section A), of the proposed indoor scenario (Section B).

A. The fall detector HW/SW

The proposed approach is based on a RGB-D camera and makes use of a Microsoft Kinect device [4]. The device, in Fig. 1, is composed of a RGB camera and an infrared (IR) depth sensor, both characterized by a resolution of 640×480 at 30 fps. It was developed by PrimeSense [11], which provides also the software library for full-body 3D motion capture



Fig. 1. Kinect Sensor.

As specificated by Microsoft, the maximum spatial workspace of Kinect IR sensor is 3.5m, even if the actual effective range depends on the environment and on its illumination condition. The fall detector subsystem, instead, is a software module mainly based on two web-available open libraries: OpenNI [12] and OpenCV [13]. The first is used for Kinect device management, whereas the second has been used to extend and improve the OpenNI functionality.

B. The proposed indoor scenario

The prototype indoor scenario is a 5x5m room (see Fig. 2) in which the proposed system was set up used to monitor the fall of people with physical. The Kinect device was positioned in the room such that it was possible to ensure the maximum field of view of the camera minimizing the shadow (and occlusion) areas.

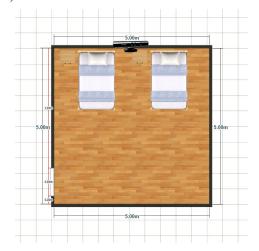


Fig. 2. The prototype room.

Fig. 3 shows three images of falls captured during the simulation test carried out in the experimental laboratory of the MET-AAL project.



Fig. 3. Indoor scenario.

More in detail, the device was positioned on a bracket of the wall, at a height of 2 meters (with respect to the floor), such that the lens makes an angle of -27° with respect to the horizontal axis (see Fig. 4).

The video samples to be processed include three types of fall (backward, forward and sideways) and, "seating", "lying on the floor" (backward, forward and sideways) and "picking up an item from the floor" activities conducted by ten different subjects.

C. The algorithm for fall detection

The movements of the human body are very articulated and, therefore, they are very complex to estimate. The proposed algorithm operates a simplification of the human body motion through a robust and reliable parameterization.

Moreover, our algorithm do not need the estimation of the coordinates of the floor (requirement necessary in other studies, e.g. [6]) as reference point with respect to which it is possible to calculate the position and posture of the person in the scene. This allows to have a more flexible, as well as a lighter algorithm under a computational point of view.

This point is in line with the real-time requirement of an automatic fall detection system, which has to feature a fast dynamics (high frame rate and low processing time) in order to

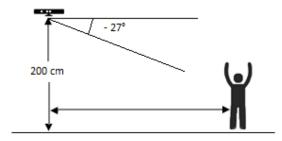


Fig. 4. Sensor Placement.

avoid missed detections.

D. 3D Bounding box extraction

The initial human detection and tracking are operated by a standard OpenNI function. OpenNI provides Depth MetaData, which are used to create a 3D bounding box which fully encloses the user in the scene.

The width, height and depth, generally named as box parameters, (TABLE II) of the 3D bounding box are estimated as the differences of the maximum and minimum points along the X, Y and Z dimensions, respectively. This information is per se sufficient for detecting all three types of fall (backward, forward and sideways). This algorithm runs inside the main loop of the depth map analyzing process.

TABLE II. ESTIMATION OF WIDTH, HEIGHT AND DEPTH

Width	W = Xmin - Xmax
Height	H = Ymin - Ymax
Depth	D = Zmin - Zmax

E. 3D Bounding box data analysis

For each frame, then, the algorithm estimates the 3D bounding box parameters, leading to a permanent condition in which the user is continuously "wrapped" into a bounding box.

The box parameters, reported in the first columns of Table II, are inputs of our fall detection algorithm. Several studies show that the height of the 3D bounding box is reducing during the fall and the width (W) and/or the depth (D) is increasing in case of lateral fall, or vice versa in case of forward (or backward) falls. Combining the two increasing dimensions of the 3D bounding box W and D, it is possible to define a composition of width-depth in order to have a new parameter WD equal to:

$$WD = \sqrt{W^2 + D^2} \tag{1}$$

In our first prototype, the fall detection algorithm used also the following other parameters: the first derivative (speed) of height (Vh), and of the width-depth (VWD) composition of the box, defined as:

$$V_h = \frac{h_k - h_{k-1}}{t_k - t_{k-1}}$$
 (2)

$$v_{WD} = \frac{WD_k - WD_{k-1}}{t_k - t_{k-1}}$$
 (3)

where k is the frame number.

However, these two previous parameters are affected by noise mainly due to the low accuracy of the sensor. For this reason, we propose to implement a Kalman Filter for an improved estimation of the height and width-depth speed.

To detect falls, we follow this procedure (see Fig. 5).

The first step of the algorithm is to recognize and track a human being in the scene.

Then, for each processed frame is checked whether the calculated speed values are greater than a certain threshold for a time interval greater than n frames, where n is the selected number of frames in the Sampling Window (S.W.).

In case it is very likely that the subject falls down, the software starts to track the V_h value to identify the inactivity of the subject. If V_h is less than a specified threshold in more than 10 frames, it means that the subject lies on the floor.

The signal released by OpenNI is quite noisy. Kalman filter has been used for data cleaning.

Since the algorithm described in [9] generates a lot of false positives, in the next section is explained how we reduce and almost eliminate them.

F. False Positives Management

To avoid false positives in a lot of normal actions like *to stand up from a chair, enter in the scene* and so on, we added a further strategy to detect the fall.

This procedure consists of using another parameter to detect the actual falls: the real world y-coordinate of the top left vertex of the user bounding box. This is a simple idea, since the y-coordinate of this point is actually similar to the y-coordinate of a possible head centroid, we monitor the variation during time of this value (V_y) to obtain information about the direction and the intensity of the movement.

If this parameter is lesser than a specified threshold, we can argue that the subject has fallen. In particular, since a fall in any direction means a really fast decrease of this value (and a simultaneous fulfillment of the other two conditions), this threshold has a negative value.

$$v_{y} = \frac{Y_{i} - Y_{i-1}}{t_{i} - t_{i-1}}$$
 (4)

Since this point is also affected by noise, the speed described above has been again filtered by Kalman.

By using this type of further check, we are able to avoid a lot of false positives that could be generated by simple actions.

III. EXPERIMENTAL RESULTS

We tested the fall-detection algorithm on different situations, using sixty videos of ten people, in which we simulated the most common fall cases of all the kind and various daily actions. The test was carried out with videos of indoor environments in order to make the scene more realistic. The typical scenario was:

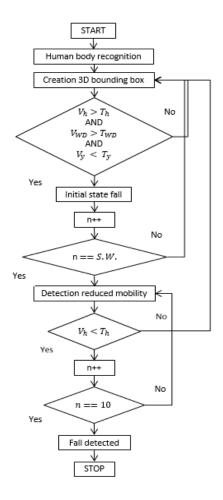


Fig. 5. Flow Chart of fall algorithm.

- a human enters in the scene;
- Kinect recognizes body silhouette;
- the algorithm builds a rectangle that contains the human silhouette meanwhile the Kalman filter reduces spikes in the variation of the rectangle;
- the human being could fall in any direction and the system has to recognize the event.

It was important to try the fall in different directions in relation to the position of the sensor (forward, back, sideways), to ensure the robustness of the algorithm.

First problem occurred when the entrance in the scene of the subject (see Fig. 6) leads to a false positive, because the built rectangle varies very quickly. For this reason we resolve this issue including the monitoring of the *Ist-order derivative* of the y-coordinate of the highest point of the silhouette: in particular when the positive-value of the 1st-order derivative occurred, it surely that was not a fall.

We measure principally four values: the top most y-axis point of the rectangle, the speed vector of this point, the height speed variation and the diagonal speed variation of the bounding box. These four parameters are monitored in each frame and if the conditions explained before are satisfied the fall is detected. Technically, we use 20-25 fps, so we assure that each type of fall can be detected in a few seconds.

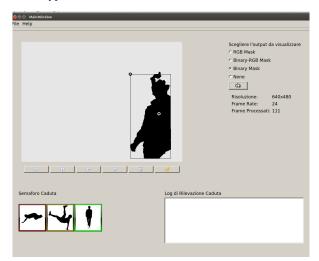


Fig. 6. The GUI of the software during the elaboration.

On the top of Fig. 7 graphics show relevant spikes when a fall occurs. The blue shape over the red shape shows the Kalman-filter effect: if it was not applied, all that spikes would be false positives. On the bottom of Fig. 7 the GUI shows a detected fall: the red-light semaphore switches on to signal the event. In this case, the importance of the Kalman-filter effect (blue shape) helps us to identify correctly the fall although there are red spikes. In the last case the recorded subject is falling forward, and the system detects the event (yellow-light semaphore): the light will become red if the position "lying on the ground" persists. The Fig. 8 shows the sequence of fall detection even in low-light conditions, that is the typical scenario in which the elderly fall down.

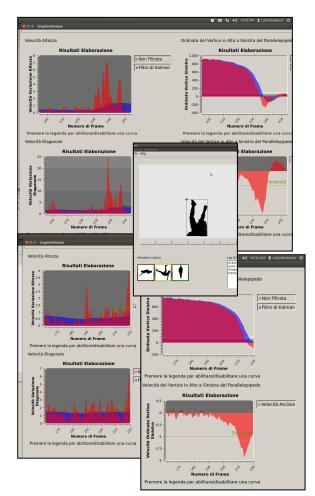


Fig. 8. Graphics when fall occurs or is detected.



Fig. 7. Fall detection in low-light conditions.

Finally, we want to point out the importance of the false positives removal in some particular cases, for example when someone throws himself on the bed or sits down on a chair.

By using simultaneously the threshold $(T_h,\,T_{WD},\,T_y\,)$ and a sampling window, we obtain the following table varying the above parameters.

TABLE III. EXPERIMENTAL RESULT

Th	TwD	Ty	S.W.	Accuracy	Sensitivity	Specificity
1.1	1.4	0.5	5	85%	83%	89%
1.1	1.4	0.5	6	87%	83%	95%
1.1	1.4	0.5	7	85%	80%	95%
1.1	1.4	0.5	8	85%	80%	95%
1.1	1.5	0.5	5	87%	83%	95%
1.1	1.5	0.5	6	87%	83%	95%
1.1	1.5	0.5	7	87%	83%	95%
1.1	1.5	0.5	8	85%	80%	95%
1.1	1.6	0.5	5	87%	83%	95%
1.1	1.6	0.5	6	87%	83%	95%
1.1	1.6	0.5	7	80%	73%	95%
1.1	1.6	0.5	8	76%	68%	95%
1.1	1.7	0.5	5	85%	80%	95%
1.1	1.7	0.5	6	85%	80%	95%
1.1	1.7	0.5	7	83%	78%	95%
1.1	1.7	0.5	8	80%	73%	95%

To measure the performance of our detection process we use accuracy, specificity, sensitivity and the false positive rate according to the following formulas:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$
 (5)

$$Specificity = \frac{TN}{TN + FP} \tag{6}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{7}$$

$$FPR = \frac{FP}{FP + TN} = 1 - Specificity$$
 (8)

The performances of the proposed fall detection system are shown in the TABLE III. The percentage of accuracy is between 76% and 87%, and the false positive rate (FPR) is about 5%.

Based on the positive results obtained, whereas the dataset consisting of only about 60 videos and the used sensor is very cheap, this software can be compared with similar ones available in the literature. For example, a similar study reported in [14] was carried out and presents a multi-sensor system for the detection of people falls in the home environment. Two kinds of expensive sensors are used: a wearable wireless accelerometer with onboard fall detection algorithms and a time-of-flight chamber.

This work correctly allows to detect falls with recovery as non-falls with a reliability of 97.3% and an efficiency of

80.0%. In addition, from the studies conducted, it is no possible to detect the human posture during or in the final state of the fall

IV. CONCLUSIONS AND FUTURE WORKS

In this paper we proposed an improved fall detection tool based on the previous one found in literature. The choice to take advantage from Kalman filter and from simple parameters such as speed and landmarks, allow us to implement a more robust and efficient software. To work in low-light conditions and to avoid rapid-downward-movement-misclassification are good features to use in particular scenarios as hospitals, recreation centers for the elderly, and domestic houses. Finally we remark that the proposed algorithm is a part of a single module developed in our Ambient Assisted Living (AAL) platform composed of a voice commands recognizer that can also be used to confirm the event of the fall, in case of cognitively capable person, and then to reduce, from a functional point of view, the possibility of false alarms.

REFERENCES

- [1] Rubenstein, Laurence Z. "Falls in older people: epidemiology, risk factors and strategies for prevention." Age and ageing 35.suppl 2 (2006): ii37-ii41.
- [2] Noury, N., Fleury, A., Rumeau, P., Bourke, A., Laighin, G., Rialle, V., Lundy, J.:Fall detection - principles and methods. In: 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBS), pp. 1663-1666 (2007).
- [3] Alwan, M., Rajendran, P., Kell, S., Mack, D., Dalal, S., Wolfe, M., Felder, R.: A smart and passive floor-vibration based fall detector for elderly. 2nd Information and Communication Technologies 1, 1003-1007 (2006).
- [4] BBC News-Microsoft Kinect 'fastest selling device on record' http://www.bbc.co.uk/news/business-12697975.
- [5] Rubenstein, Laurence Z., and Karen R. Josephson. "The epidemiology of falls and syncope." Clinics in geriatric medicine 18.2 (2002): 141-158.
- [6] Rougier, Caroline, et al. "Fall detection from depth map video sequences." Toward Useful Services for Elderly and People with Disabilities. Springer Berlin Heidelberg, 2011. 121-128.
- [7] Planinc, Rainer, and Martin Kampel. "Introducing the use of depth data for fall detection." Personal and Ubiquitous Computing (2012): 1-10.
- [8] Tong, Lina, et al. "A research on automatic human fall detection method based on wearable inertial force information acquisition system." Robotics and Biomimetics (ROBIO), 2009 IEEE International Conference on IEEE, 2009.
- [9] Mastorakis, Georgios, and Dimitrios Makris. "Fall detection system using Kinect's infrared sensor." Journal of Real-Time Image Processing (2012): 1-12.
- [10] Rougier, Caroline, et al. "Fall detection from depth map video sequences." Toward Useful Services for Elderly and People with Disabilities. Springer Berlin Heidelberg, 2011. 121-128.
- [11] Primesense. http://www.primesense.com/.
- [12] OpenNI. http://www.openni.org/.
- [13] OpenCV. http://www.opencv.org/.
- [14] G. Diraco, A. Leone, P. Siciliano, M. Grassi, P. Malcovati, "A multi-sensor system for fall detection in ambient assisted living contexts", SENSORNETS, page 213-219. SciTePress, (2012).