# A New Method for Fall Detection of Elderly Based on Human Shape and Motion Variation

Abderrazak Iazzi[1(✉)], Mohammed Rziza[1], Rachid Oulad Haj Thami[1,2], and Driss Aboutajdine[1]

[1] LRIT URAC 29, Faculty of Sciences, University of Mohammed V in Rabat, 4, Avenue Ibn Batouta, BP 1014 Rabat, Morocco
abderrazak.iazzi@gmail.com

[2] RIITM, ENSIAS, University of Mohammed V in Rabat, Avenue Mohammed Ben Abdallah Regragui, Madinat Al Irfane, BP 713 Agdal Rabat, Morocco

**Abstract.** Fall detection for elderly and patient has been an active research topic due to the great demand for products and technology of fall detection in the healthcare industry. Computer vision provides a promising solution to analyze personal behavior and detect certain unusual events such as falls. In this paper, we present a new method for fall detection based on the variation of shape and motion. First, we use the CodeBook method to extract the person silhouette from the video. Then, information of rectangle, ellipse and histogram projection are used to provide features to analyze the person shape. In addition, we represent the person shape by three blocks extracted from rectangle. Then, we use optical flow to analyze the person motion within each blocks. Finally, falls are detected from normal activities using thresholding-based method. All experiments show that our fall detection system achieves very good performances in accuracy and error rate.

**Keywords:** Fall detection · Elderly people · Recognition posture · Monitoring · Healthcare · Background subtraction · CodeBook · Daily activities

## 1 Introduction

Majority of elderly person living alone face high risk situations such as falls. These falls causes high damages such as fractures and dramatic psychological consequences. In the past few years, many works have been carried out in this area. There are a lot of proposed approaches that we can categorize in two kinds: (i) methods based on wearable-sensor device and (ii) methods based on computer-vision device. For the wearable-sensor devices, most fall detection techniques are based on accelerometers, buttons or gyroscope [1,2], but the major problem is that elderly people usually forget to wear them or do not feel comfortable with. Concerning computer-vision-based fall detection system,

Project ANGEL: This work is supported by MERSFC and CNRST.


G. Bebis et al. (Eds.): ISVC 2016, Part II, LNCS 10073, pp. 156–167, 2016.
DOI: 10.1007/978-3-319-50832-0_16

they use scene analysis to identify a laying posture or/and vector analysis to identify abnormal motion. The advantages of using computer-vision are detection or/and identification of several events. Furthermore, the camera installed in the house can detect the fall without the elderly's interaction and with their privacy respected.

The rest of this paper is organized as follows: In Sect. 2, we briefly present some existing video-based fall detection systems. In Sect. 3, we present details of the proposed fall detection system. Section 4 presents results and evaluation of our proposed system. Followed by a general conclusion and discution of future works.

## 2 Related Work

In the literarture, researchs have been done to detect falls using image processing techniques. There are some methods based on analyzing shape variation as [3–6]. Some other approaches are based on analyzing motions variation [7–9]. In order to improve these methods, some authors have recently used both shape and motion variation to detect a fall such as [10–14]. Lee and Chung [10] propose a novel computer vision technique that can first extract objects more accurately, then discriminate between abnormal and normal activities relying on threshoding-based methods. Abnormal event detection based on visual sensor by using shape features variation and 3-D trajectory has been presented to overcome the low fall detection rate. In [11], Chua et al. propose an analytical method to detect a fall. The human shape analysis based on features extracted from the three points (centroid of upper, middle and lower human part) is used to detect possible falls. A longer duration of the inactivity period is used to ensure the person is completely unconscious after a fall. Thus, a fall is confirmed. Other fall detection systems based on learning methods have been proposed by Charfi et al. [12]. The authors define 14 features based on the bounding box such as height and width, aspect ratio, and centroid coordinates of the box. Transformations (Fourier, wavelet) are applied to these features before fall detection through SVM and AdaBoost classification.

Feng et al. [14] proposed a novel vision-based fall detection method for monitoring elderly people in a house care environment. The human body is represented with ellipse fitting, and the silhouette motion is modeled by an integrated normalized motion energy image computed over a short-term video sequence. Then, the shape deformation quantified from the fitted silhouettes is used as features to distinguish different postures of the person. Yu et al. [13] proposed a novel method based on ellipse fitting, shape description and position information. These later features are collected to construct Online one class Support Vector Machine (OCSVM) Model to distinguish a normal posture from an abnormal posture. Then, two rules are used to reduce false alarms.

# 3 The Principal Module of Fall Detection System

Our proposed fall detection system, as shown in Fig. 1, is composed of four modules: (i) video capture, (ii) Detection Moving Object, (iii) Features Extraction and (iv) Recognition Behavior and fall detection. The goal of our system is to detect the fall of people living alone at home. In this purpose, we use a camera sensor which is the first component in our system to collect video and data information of the whole environment, then the detection of moving objects is applied the background subtraction (BS) by using the CodeBook (CB) model method [15]; the first step is to build background model and train it. It is then applied to detect the moving objects. However, the results of the CB method are not satisfying because of noises (due to moving some furniture) and brightness changes. In order to avoid this, we add a post-processing component as shown in Fig. 1 where we update the background model by adding the non-required objects detected. In our system, we use blob-merging method [16] for small area and optical flow for big area. In the component of fall detection, we analyze the features extracted from the silhouette person to detect abnormal activities in order to detect a fall.

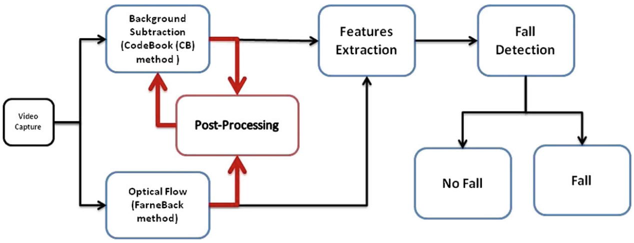



**Fig. 1.** Diagram of the main modules of fall detection system

## 3.1 Background Subtraction

Background subtraction is the first step in our fall detection system. Approaches proposed by [15,17] are so common and widely used for extracting moving objects from the image. For this purpose, we use the CB method described in [15] for its advantages, such as it capability of removing shadows and of giving, good result in complex environments like resistance to acquisition artifacts. In [13], the authors show a qualitative and quantitative comparison between methods presented in [15,17].

In general, the BS results obtained contain different noisy artifacts caused by brightness changes and displacement of objects. For this purpose, we add an additional post-processing component to avoid this problem and to improve the BS results.

### 3.2 Post-Processing

After the detection of moving objects in the image by BS using CB method, the result obtained is not always satisfying in general. The fact that the person the person is not the only object that is detected. In order to remove these non-required objects, we propose to add a post-processing component (cf. Fig. 1) composed of two steps:

**Blob-merging** [16]**:** Detected objects with a number of pixels of their area less than a threshold should be removed. In our system, this threshold is 50 pixels.

**Determining the Human Silhouette:** If the number of detected objects after the first step is more than one because of changes in some furniture background, it is necessary to determine the silhouette of the person through the use of optical flow [18]. In our system, we are not interested in identifying and recognizing the human silhouette (i.e. the elderly living alone in general). Therefore, it is necessary to distinguish humans from other objects. The optical flow is applied between two successive images for the pixels in motion. The blob which has the most moving pixels is determined as the object desired and non-required objects are added to the background model of CB method. Sometimes, two blobs have nearly the same number of moving pixels. We use a position information to determine the object which is the blob of a person. The desired object (blob) is the object which has the smallest distance between its position (center of gravity) and the last recorded position of the person.

### 3.3 Features Extraction

The second component in our system is feature extraction to discriminate various human activities. All features can be divided into global and local features which are extracted from bounding box, ellipse, position information and projection histogram. All these features are used for analyzing the changes of shape. The optical flow is used for analyzing person's motion.

**Shape Change Analysis:** The first feature can be extracted from the bounding box drawn around the person as shown in Fig. 2. The **ratio of Bounding Box** (RB) is the ratio between the height and the width of the bounding box. During persons activities at home, the height and the width of the bounding box will they change their posture; the RB will change as well [19]. The second feature can be extracted from the ellipse fitted to human body silhouette as shown in Fig. 2b. The moment-based method [13] is applied to fit the ellipse. From the ellipse, we extract the person **orientation**($\theta$). From Fig. 2b, we can see that the ellipse fitting can describe the human body posture as using the orientation of the ellipse.

The posture of the person can be determined by using their 2D binary silhouette as shown in Fig. 3. The projection histogram can give more detaills, compared to the bounding box in terms of the person's posture. It projects the silhouette following Y-axis and X-axis giving vertical and horizontal projection

(a) (b) (c) (d)

**Fig. 2.** The result of bounding box and ellipse drawn around the person. (2a) human silhouette, (2b) bounding box result and (2d) ellipse result,

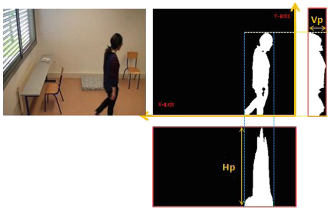



**Fig. 3.** Extracted features from The Projection Histograms of the silhouette person

(Vp and Hp respectively). **The ratio (RH)** between Hp and Vp is considered as feature in our system.
**Motion Analysis:** The optical flow is a visual displacement field that helps to explain variations in a moving image in terms of displacement of image points. In the literature, there are several approaches for motion detection using optical flow such as [18,20], where the authors calculate the flow for each pixel in the first image input and they have used multi-scale for tracking the sparse features.

In our system, we used the algorithm presented in [18]. Tracking over image pyramids allows large motions to be caught by local windows [18]. The optical flow can give two important pieces of information as features to analyze behavior, especially the person's **velocity** and **the direction** of motion to discriminate between two modes: the person's normal and abnormal activities. From the person extracted silhouette, we draw a bounding box around the person, then we divide it into three blocks based on the width and height. The first block contains the person's head, the center block contains the belly and arms and the third block contains the feet.
From the person's blob and these three latter blocks, we compute the velocity of each block based on the result of optical flow algorithm by using this formula:

$$velocity(Block_i) = \frac{\sum velocity(Pixel_{\epsilon blob \cap Block})}{\#Pixels \in blob \cap Block_i} \tag{1}$$

Where $velocity(Pixel_{\epsilon blob \cap Block})$ is the displacement of pixel by using optical flow. In order to compute the person's velocity and based on formula 1, we use the following formula:

$$velocity(person) = \frac{1}{3}\sum_{1}^{3} velocity(Block_i) \tag{2}$$

The second feature extracted from the result of optical flow is the motion direction. We compute motion direction sing the orientation of the person's displacement. Four directions are defined, namely up, left, rigth and down. In our system, we relied on down-direction to detect a fall.

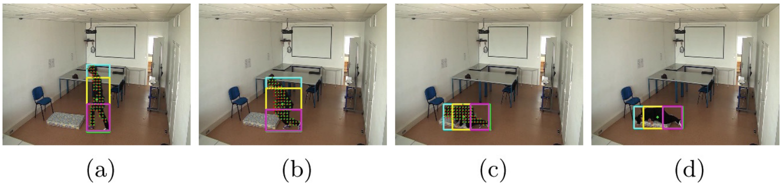


**Fig. 4.** Result of fitting bounding box with three blocks around the person and the result of optical flow.

### 3.4 Fall Detection

After features extraction, the last component in our system is a recognition of the person's behavior. The main goal of this component (cf. Fig. 1) is to analyse different activities of the elderly (i.e. to detect the existence of the fall). This component is based on three basic states. In the first state, we check if the current activity (posture) is an abnormal event, then we determine if it is a right fall by using the features extracted from the human silhouette in the previous component. The abnormal event is characterized by a sudden change in velocity and shape.

Normal and abnormal activities are defined as controlled and uncontrolled movements respectively. The elderly can take different postures (walking/standing, lying, bending and sitting). To move from one posture to another, the normal activities take more time whereas abnormal activities take less time. For the velocity, it is almost stable for normal activities while it is very excessive for abnormal activities where it may change suddenly.

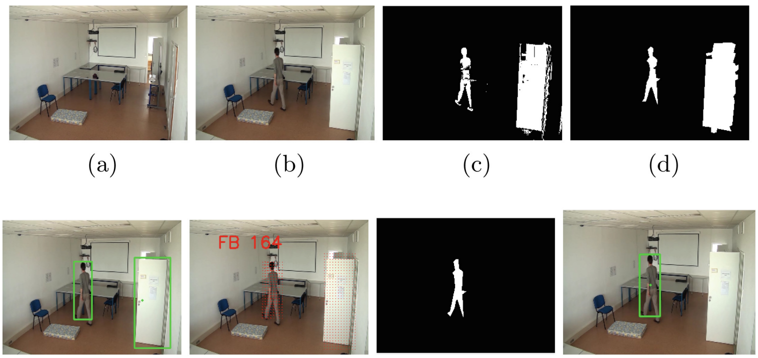
In our system, we study different activities as walking/standing, lying, bending, sitting and falling. Every activity can be determined by using the features extracted from the shape and motion. The fall occurs when an elderly's posture changes from (walking/standing, sitting, bending) to lying and the velocity of the motion is higher than a threshold. Thus, if we detect a big velocity and down-direction, then we check the lying-posture of the person by using the shape-variation in order to confirm if it is a fall or it is normal activity.

## 4 Experimental Results

In this part, we show the performance of our results achieved using our approach to detect a fall. Video processing is done by visual studio C++ 10.0 (with library OpenCv 2.4) in the Intel (R) Core (TM) i5-2430M CP Laptop with 4.00 GB memory. All experiences were applied to two publicly available fall datasets [11,12]. The first Dataset1[1] [12] is composed of 219 videos, including different activities, 95 videos of normal activities and 124 videos of fall activities. All activities were simulated in several location (Home, Lecture-room, Coffee-room, Office) by people of different ages. The second Dataset2[2] [11] is composed of 20 videos including 38 normal daily activities (6 crouch-down, 6 squat-down, 10 walking, 6 running, 4 lie-down, 6 sit-down) and 29 fall as backward falls, forward falls, sideway falls, and falls due to loss of balance.

### 4.1 Background Subtraction Results

In Fig. 5, we show the results of Background subtraction by using CB method and post-processing to improve the performance of detecting moving object. As shown in the figure, the result of CB method needs to update its background model in order to extract only the human silhouette as presented in the last image (h) of the figure.




(a) (b) (c) (d)

(e) (f) (g) (h)

**Fig. 5.** The result of background subtraction and post-processing. (a) background image, (b) current image, (c) result of codebook, (d) result of step 1 in post-processing, (e) result of tracking two object, (f) result of optical flow, (g) final result of background subtraction after step 2 of post-processing, (h) person's track

### 4.2 Feature Analysis and Fall Detection

To analyse each feature (RB, $\theta$, RH, velocity), scatter plots are ploted as shown in Fig. 6. Each feature is ploted as curve corresponding to its variation for video sequence. By analysing the shape of the person by these features, we can describe more accuratly the activities of the person as standing/walking, sitting, bending, lying and falling.

[1] http://le2i.cnrs.fr/Fall-detection-Dataset.
[2] http://foe.mmu.edu.my/digitalhome/FallVideo.zip.

In our system, we define our threshold of each feature describing if an abnormal event is occur or not. The features RB and RH are defined as 1, the threshold of $\theta$ is defined as 50° and the threshold of velocity is defined as 3.

When an abnormal activity occurs, as shown in Fig. 6, the person's velocity is higher than the threshold. The RB, RH and $\theta$ are less than threshold. We use the velocity of the three blocks (head, center, feet) and the difference of velocity Head-feet and Center-feet as shown in Fig. 7. From Fig. 7, we show that the velocity of three blocks (head, center, feet) has to be higher than the threshold which is defined as 3, and the velocity of Head-feet and Center-feet is less than the threshold which is defined as 4. The last feature is the down-direction of motion where the duration of down-direction is higher than the threshold which is defined as 5 frames at least to decide if an abnormal event has occurred. For the final decision, in order to ensure and to confirm possible fall, we use 10 frames as threshold-duration of inactivity period compared with the work [11] which only used 5 frames.

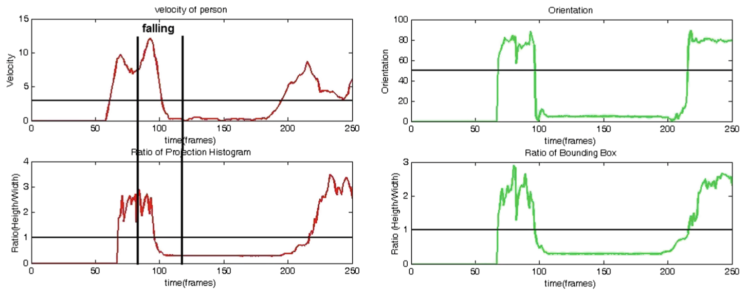



**Fig. 6.** Result of Features extraction from one video sequence in dataset [11] contains four falls. First image is for the velocity of the person using optcal flow, the second curve is for the orientation ($\theta$) of the person, the third curve is for the Ratio of projection histogram (RH) and the last curve is for the Ratio of bounding box (RB)

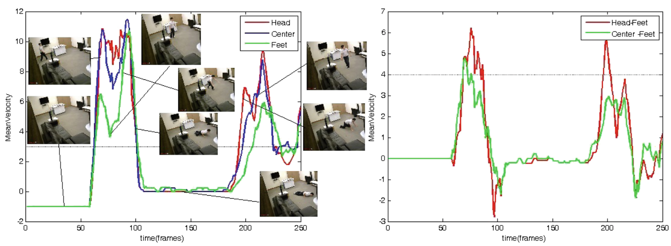



**Fig. 7.** Result of velocity of each block (Head, Center, Feet) of the fall from dataset [11]. First plot is for the velocity of the Head, Center and Feet. The second plot shows the variation of difference between the velocity of the Head and Feet and the velocity of the Center and Feet.

### 4.3 Performance of Our Fall Detection System

To evaluate the performance of our proposed method, we use two famous criteria that are widely used in fall detection systems.

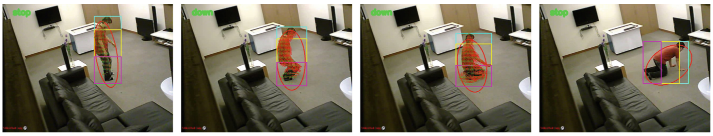
$$Sensitivity = \frac{TP}{TP + FN} \qquad Specifity = \frac{TN}{TN + FP} \tag{3}$$

where TP means that the fall has occured and the system detects as fall, FP means the fall has occurred and the system doesn't detect, TN means the fall has not occurred and the system doesn't detect and FN means the fall has not occurred and the system detects as a fall.

In Table 1, we show the experiment results of our proposed fall detection method. For the datasets, the fall incidents were not detected because the person's(human) body was in straight line during these falls. Thus, no change of features RB, RH and $\theta$. The velocity is higher than threshold, the duration of down-direction of motion is less than a threshold, then the system considered there is no possible fall. Some uncontrolled movement such as brutally crouches-down as shown in Fig. 8d and lie-down on chair activity were detected as falls because of a high velocity (i.e. higher than the threshold) and a high duration of down-direction of motion. Overall, our system can achieve high accuracy in fall detection. For the Dataset1, The detection accuracy is up to 96.34% and rate error is 3.65%.

**Table 1.** Fall detection results of our proposed method

| Recognition system | Dataset | | | |
|---|---|---|---|---|
| | Dataset1 | | Dataset2 | |
| | Fall occur | Not occur | Fall occur | Not occur |
| Positive | 118 | 2 | 28 | 1 |
| Negative | 6 | 93 | 1 | 31 |




(a) (b) (c) (d)

**Fig. 8.** Crouch-down activity

Compared with our Dataset2, the dataset cited in [11] which contains only 21 falls and 31 normal daily activities. As shown in [11] and as we said before, the crouch-down activities are detected as falls because the system detect any abnormal activity similar to a fall. In addition, detecting these activities as a fall is better for any system than to not detect them. With these conditions, we use two scenarios to show the performance of our system. The first scenario is to use the Dataset2 with crouch-down activities and the second scenario without them. Then, this new dataset is composed of 32 normal daily activities. For the first scenario, our system achieve to 88.05% of accuracy and 11.94% of rate error. For the second scenario, our system achieve to 96.72% of accuracy and 3.27% of rate error.

We compare our system, as shown in Table 2, with some proposed systems discussed in the state of the art. For our proposed method, we use the Dataset1 and the second scenario of Dataset2 while the methods proposed in [11,14] only use the second scenarios of Dataset2. The methods proposed in [10,13] use their private own datatset. The results of the systems show that our proposed method has better performance in sensitivity than the threshold-based methods (i.e., [10,11]) and machine-learning-based methods(i.e., [14]). Table 2 also summarizes that the sensitivity of using Yu et al. approach [13] still surpasses in sensitivity given by other approaches.

**Table 2.** Comparison of our proposed method with some methods in the state of the art

| Methods | Performance | | | |
|---|---|---|---|---|
| | Sensitivity | Specifity | Accuracy | Error rate |
| Lee and Chung [10] | 94% | 98% | 97% | 2% |
| Chua et al. [11] | 90.5% | 93.3% | - | 6.3% |
| Yu et al. [13] | 100% | - | - | 3% |
| Feng et al. [14] | 95.2% | 100% | - | - |
| Proposed method (Dataset1) | 95.16% | 97.89% | 96.34% | 3.65% |
| Proposed method (Dataset2) | 96.55% | 96.87% | 96.72% | 3.27% |

### 4.4 Conclusion and Discussion

In this paper, we presented a fall detection system using single camera for monitoring elderly people. We proposed two critical components in our approach. We first proposed a post-processing to improve CB Background Subtraction results by updating a background to cope with changes of background model. Secondly, the recognition of behavior was applied by using simple feature extracted from the human silhouette. The ratio of bounding box, ratio of projection histograms and orientation was used to analyze shape variation and the velocity was used to analyze the motion. The combination of these features gives good results to discriminate between a fall and a normal activity. All experiments are tested on two different datasets and show that our proposed system gives good results to detect a fall and to avoid false alarm without using any supervised or semi-supervised classifier as other methods do for final decision.

However, our system still has some limitations. We first need to improve the background subtraction which suffers from some drawbacks, such as removing shadows and occlusions. In fact, we only rely on the motion and the distance between objects to update the background model. Another limitation is the videos used in all experiments which are only made under good lighting conditions. Finally, as has previously been noted, the performance of our system depends on the best thresholds used to detect a fall. Nevertheless, in real life, our system needs to be adapted for the monitoring of different people.

In the future, we will focus on combining our approach with an SVM classifier in order to improve the performance of fall detection system and be adapted for monitoring different size of persons. Also, using 3D information can describe more precisely the person's posture to recognize more detailes about daily activities of the person.

## References


1. Zhang, T., et al.: Fall detection by embedding an accelerometer in cellphone and using KFD algorithm. Int. J. Comput. Sci. Netw. Secur. **6**(10), 277–284 (2006)
2. Ge, Y., Xu, B.: Detecting falls using accelerometers by adaptive thresholds in mobile devices. J. Comput. **9**(7), 1553–1559 (2014)
3. Miaou, S.G., Sung, P.H., Huang, C.Y.: A customized human fall detection system using omni-camera images and personal information. In: 1st Transdisciplinary Conference on Distributed Diagnosis and Home Healthcare, D2H2, pp. 39–42. IEEE, April 2006
4. Nasution, A.H., Emmanuel, S.: Intelligent video surveillance for monitoring elderly in home environments. In: IEEE 9th Workshop on Multimedia Signal Processing, MMSP 2007, pp. 203–206. IEEE, October 2007
5. Cucchiara, R., Prati, A., Vezani, R.: An intelligent surveillance system for dangerous situation detection in home environments. Intelligenza artificable **1**, 11–15 (2004)
6. Rougier, C., Meunier, J., St-Arnaud, A., Rousseau, J.: Robust video surveillance for fall detection based on human shape deformation. Circuits Syst. Video Technol. **21**, 611–622 (2011)
7. Wu, G.: Distinguishing fall activities from normal activities by velocity characteristics. Biomechanics **33**, 1497–1500 (2000)
8. Babu, R.V., Ramakrishnan, K.R.: Recognition of human actions using motion history information extracted from the compressed video. Image Vis. Comput. **22**, 597–607 (2004)
9. Nait-Charif, H., McKenna, S.J.: Activity summarisation and fall detection in a supportive home environment. In: Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004, vol. 4, pp. 323–326. IEEE, August 2004
10. Lee, Y.S., Chung, W.Y.: Visual sensor based abnormal event detection with moving shadow removal in home healthcare applications. Sensors **12**, 573–584 (2012)
11. Chua, J.L., Chang, Y.C., Lim, W.K.: A simple vision-based fall detection technique for indoor video surveillance. Sig. Image Video Process. **9**(3), 623–633 (2015)
12. Charfi, I., Miteran, J., Dubois, J., Atri, M., Tourki, R.: Definition and performance evaluation of a robust svm based fall detection solution. In: 2012 Eighth International Conference on Signal Image Technology and Internet Based Systems (SITIS), pp. 218–224. IEEE November 2012

13. Yu, M., Yu, Y., Rhuma, A., Naqvi, S.M.R., Wang, L., Chambers, J.A.: An online one class support vector machine-based person-specific fall detection system for monitoring an elderly individual in a room environment. IEEE J. Biomed. Health Inform. **17**, 1002–1014 (2013)
14. Feng, W., Liu, R., Zhu, M.: Fall detection for elderly person care in a vision-based home surveillance environment using a monocular camera. Sig. Image Video Process. **8**, 1129–1138 (2014)
15. Kim, K., Chalidabhongse, T.H., Harwood, D., Davis, L.: Background modeling and subtraction by codebook construction. In: 2004 International Conference on Image Processing, ICIP 2004, vol. 5, pp. 3061–3064. IEEE, October 2004
16. Gonzalez, R.: Digital Image Processing. Pearson Education, Upper Saddle River (2008)
17. Zivkovic, Z.: Improved adaptive Gaussian mixture model for background subtraction. In: Proceedings of the 17th International Conference on Pattern Recognition, ICPR 2004, vol. 2, pp. 28–31. IEEE, August 2004
18. Farnebäck, G.: Two-frame motion estimation based on polynomial expansion. In: Bigun, J., Gustavsson, T. (eds.) SCIA 2003. LNCS, vol. 2749, pp. 363–370. Springer, Heidelberg (2003). doi:10.1007/3-540-45103-X_50
19. Töreyin, B.U., Dedeoğlu, Y., Çetin, A.E.: HMM based falling person detection using both audio and video. In: Sebe, N., Lew, M., Huang, T.S. (eds.) HCI 2005. LNCS, vol. 3766, pp. 211–220. Springer, Heidelberg (2005). doi:10.1007/11573425_21
20. Bouguet, J.Y.: Pyramidal implementation of the Lucas Kanade feature tracker: description of the algorithm. Technical report (2000)