Video Analytic for Fall Detection from Shape Features and Motion Gradients

Muhammad Jamil Khan and Hafiz Adnan Habib

***Abstract*—**Unintentional falls are common causes of serious injuries and health threats to patients as well as senior citizens living alone. Advanced computer vision algorithms and low cost cameras can be used for assessment of health hazards such as falls. In this paper, we propose a new approach to deal with this particular problem i.e. to detect unintentional falls, which are one of the greatest risks for seniors living alone. Our proposed approach is based on a combination of motion gradients and human shape features variation. Our proposed system provides promising results on video sequences of simulated falls.

***Index Terms*—fall recognition, elders, Motion gradient Image, length-width feature**

1. Introduction

Development of helping technology for patients and senior citizen is an active research issue due to that healthcare industry has a massive demand for such goods and technology [1][2]. At the same time advances of cameras and computer vision algorithms make such development viable. More importantly this development is greatly funded and pushed by governments in many developed countries [1]. The immediate treatment of injured people by fall is very critical. Hence, we should detect falls as soon as possible so that injured can get an immediate treatment. Although the perception of a fall is in the common sense, it is difficult to illustrate it precisely due subtle and multipart nature of body movements that requires exact and reliable measuring techniques, and thus to categorize its means of detection so the evaluation of unintentional fall is difficult. It can be illustrated as the hasty change from standing position to the reclining or almost lengthened pose but is not organized movement like lying. Machine vision involves posture analysis and classification for fall detection and deal with following stages environment modeling, person's detection and tracking, understanding and describing behaviors. Video cameras are relatively low cost, rich source of machine vision algorithms that can be used for evaluation of health hazards such as falls.

2. Related Work

Some research work in area of unintentional fall detection has been done using surveillance cameras. For example, a simple method consists to analyze the bounding box coordinates representing the person in a single image [6]. Anderson use MHI and ellipse enclosed human body to detect falls [4]. J. Q. Lin uses panoramic cameras to observe the elderly fall detection [7]. Although J. Q. Lin accomplished the detection of fall action but his method is very difficult to discriminate between intentional fall e.g. lying down and unintentional falling activity. The false detection is high. Zhong presented an unsupervised method in [8] for detecting abnormal activity using a fusion of many simple features. Using a method similar to document-keyword analysis, he divided the video into equal length segments and classified the extracted features into prototypes, from which a prototype-segment co-occurrence matrix was calculated and used to determine abnormal activity. Zhong tested this procedure for a few different scenarios, one of which included typical activity in a nursing home. To overcome this problem, some researchers [9]**Error! Reference source not found.** have mounted the camera on the ceiling. Lee and Mihailidis [9] achieve fall detection by estimating the shape and the 2D velocity of the person, and define inactivity zones like the bed. Nait-Charif and Mckenna [10] tried to track a person using a ellipse model and a particle filter in a supportive home environment and detect abnormal actions such as falls. However this method using surveillance camera sensor was not sufficient to discern between fall activities and normal daily activities such as sitting down rapidly.

ISBN:978-988-18210-2-7

3. System Overview

We proposed a new method based on the Motion gradient Image and variation in extracted shape features.

3.1 Motion Gradient Image

Our technique is based on the fact that the motion is large when a fall occurs [5]. So, the first step of our system is to detect large motion of the person on the video sequence using the Motion History Image. A large motion can be occurred when a person running or sitting quickly so using motion gradient to estimate the speed of motion and direction in which motion occurred or normal optical flow to find regional orientation. This method is described in section 4.

3.2 Change in the Human Shape

When a large motion is detected in a specific direction, we analyze the shape of the person in the video sequence. During a fall, the human shape changes and, at the end of the fall, the person is generally on the ground with few and small body movements. A change in the human shape can discriminate if the large motion detected is normal (e.g. the person walks or sits) or abnormal (e.g. the person falls). The extraction of the human shape is described in section 5.The different steps of proposed fall recognition system combining motion gradient image and human shape are described in section 6.

4. Motion Gradient Image

Motion gives crucial information about fall, because no serious fall occurs without a large movement. Based on this observation, we decided to extract some motion information from the video sequence. Capturing a foreground understandable silhouette of the moving object or person is obtained through application of background subtraction technique [11] . As the person moves, copying the most recent extracted silhouette as the highest values in the motion history image creates a layered history of the resulting movement; in general this highest value is just a floating point timestamp of time elapsing since the application was launched in milliseconds.

4.1 *timed* Motion History Image

In this paper, proposed method use a floating point Motion History Image [12] where new silhouette values are copied in with a floating point timestamp. This MHI representation is updated as follows:
If current silhouette is at (x,y)

$$tMHI_{\delta}(x,y) = \begin{cases} \tau & if current silhouette at(x,y) \\ 0 & elseif \quad tMHI_{\delta}(x,y) < (\tau - \delta) \end{cases} \quad (1)$$

where τ is the current time-stamp, and δ is the maximum time duration constant associated with the template. This method makes our representation independent of system speed. We title this demonstration the *timed* Motion History Image (*t*MHI).

Fig. 1 Time Motion History Image

4.2 Making Motion Gradients Image

Observe the right image in Fig.1 (*t*MHI) that if we took the gradient of the *0*MHI, we would get direction vectors pointing in the direction of the movement of the body. Note that these gradient vectors will point orthogonal to the moving object boundaries at each “step” in the *t*MHI giving us a normal optical flow representation. Gradients of the *t*MHI can be calculated efficiently by convolution with separable Sobel filters in the X and Y directions yielding the spatial derivatives: $S_x(x,y)$ and $S_y(x,y)$. Gradient orientation at each pixel is then:

$$\phi(x,y) = \arctan \frac{S_y(x,y)}{S_x(x,y)} \qquad (2)$$

We must be careful, though, when calculating the gradient information because it is only suitable at positions within the *t*MHI. The surrounding boundary of the *t*MHI should not be used because inclusion of non-silhouette pixels would be took part in corrupting the result of gradient calculation. Only *t*MHI inside pixels of silhouette pixels should be examined. Additionally, we must not use gradients of MHI pixels that have a contrast which is too low or too high in their local neighborhood. Applying the above criteria to the raw gradients yields a masked region of valid gradients. Next work is needed to estimate the magnitude of motion that as follow

$$M(x,y) = \sqrt{S_x^2(x,y) + S_y^2(x,y)} \qquad (3)$$

After calculating the motion gradients, we can then extract motion features to varying scales.

4.3. Global Motion Orientation

Calculation of the global orientation should be weighted by normalized *t*MHI values to give more influence to the most current motion [12] within the template. A simple calculation for the global weighted orientation is as follows:

$$\bar{\phi} = \phi_{ref} + \frac{\sum_{(x,y)} angDiff(\phi(x,y), \phi_{ref}) * norm(\tau, \delta, tMHI_{\delta}(x,y))}{\sum_{(x,y)} norm(\tau, \delta, tMHI_{\delta}(x,y))} \quad (4)$$

where $\bar{\phi}$ is the global motion orientation, ϕ_{ref} is the base reference angle (peaked value in the histogram of orientations), $\bar{\phi}$ **(***x*,*y***)** is the motion orientation map found from gradient convolutions, **norm**($\tau, \delta, tMHI_{\delta}(x,y)$)is a normalized *t*MHI value (linearly normalizing the *t*MHI from 0-1 using the current time-stamp τ and duration **δ**), and $angDiff(\phi(x,y), \phi_{ref})$ is the minimum, signed angular difference of an orientation from the reference angle.

5. Human Shape Representation

5.1. Human Segmentation

We employ a simple silhouette length width ratio based approach to segment human body. The length–width ratio is derived from lateral histograms of segmented people. Horizontal histogram is first computed. Find the row with the maximum number of pixels. Then search from the maximum row to the top of the image. If the row data is less than a threshold, we mark it as the top of the body (H_{top}). Then search from the maximum row to the image bottom. If the row data is less than a threshold, it is marked as the moving object bottom (H_{bottom}). Second, a vertical image histogram is performed. Finding the left (H_{left}) and right (H_{right}) moving object boundaries are found in the same way. An illustrative result is shown in Fig. 2. Third, the moving object height and width are found as follows:

$$H_{length} = H_{bottom} - H_{top} \quad (4)$$

$$H_{width} = H_{right} - H_{left} \quad (5)$$

Finally, the length–width ratio ($L-W$ ratio) can be derived as follows:

$$L - Wratio = \frac{H_{length}}{H_{width}} \quad (6)$$

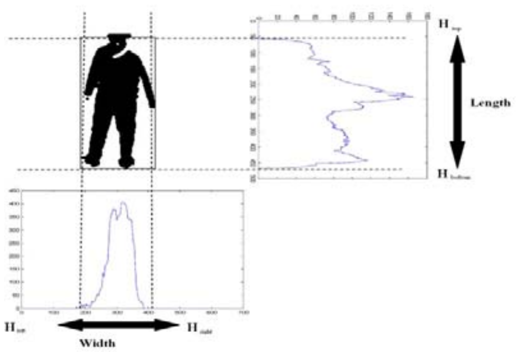


Fig.2 Segmentation of Human body

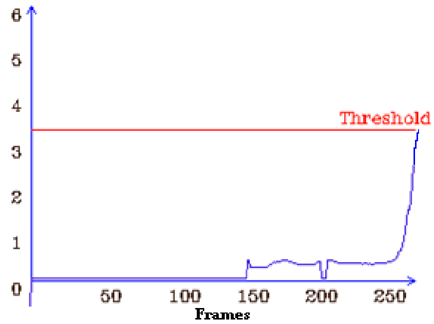
5.2 Shape feature Extraction

After segmentation of the foreground object, the next work is related to feature extraction to carry on fall behavior recognition. We should consider the features to extract from two aspects: First, they should be so simple as to be easy to recognize and have a better real-time. Second, they should be complete enough to classify the behavior of fall [6]. Suppose the width of the foreground object's bounding box in the nth frame is *w* and the height is *h* and the width is w' and the height is h' in the (n+1)th frame. The features we extract are shown below:

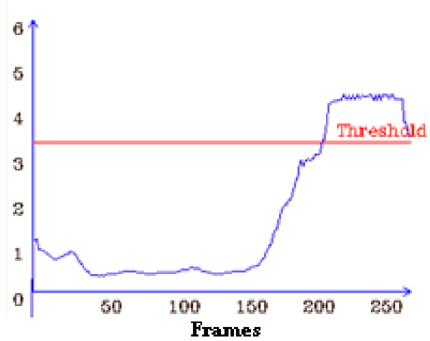
$$\alpha = \frac{w}{h} \quad (7)$$

$$\beta = \sqrt{(w - w')^2 + (h - h')^2} \quad (8)$$

α used to distinguish standing and lengthen postures and β depicts how fast changes occur. Therefore, combining α and β indicate falling activity occurrence. The curves of α and β when the falling takes place are shown in Fig.3 (a) and Fig.4 (a). In the process of fall event, α has been increasing; simultaneously the change span of β is large. The curves, of α and β, indicate falling activity occurrence. The curves of α and β when the person is lying down are shown in Fig.3 (b) and Fig.4 (b).

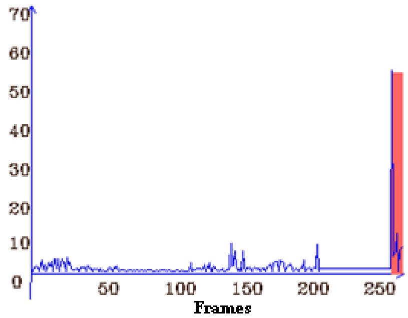


(a) The change of α in fall

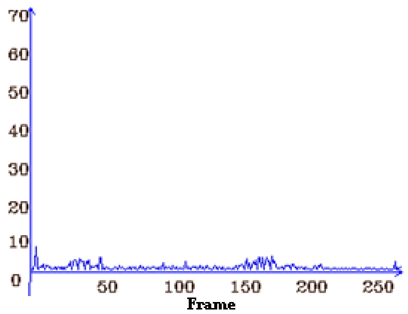


(b) The change of α in lying down

Fig.3 The change of α in falling and lying down



(a) The change of β in fall



(b) The change of β in lying down

Fig.4 The change of β in falling and lying down

6. Fall Recognition System

6.1 Fall Detection System Overview

Our proposed fall detection system is based on the Motion Gradient Image and variation in the shape features of the monitored person. An overview of our fall recognition algorithm is shown in Fig. 5. The three main steps of the algorithm are:

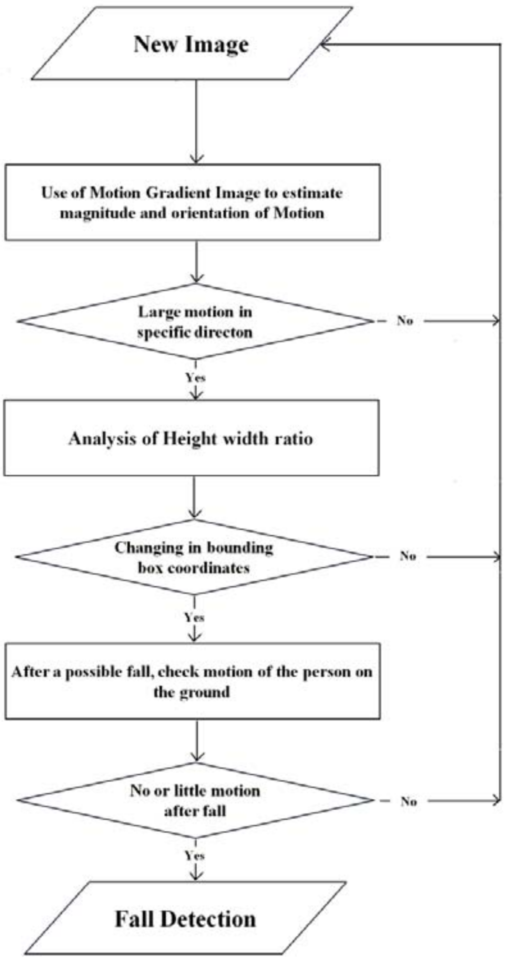


Fig.5 System diagram

Motion Gradient Estimation The estimation of the motion of the person allows detecting large motion like falls. But a large motion can also be a characteristic of a walking person, so we need to analyze further to discriminate a fall from a normal movement. To discriminate fall motion from other we use Global Motion Orientation to detect the direction of motion.

Analysis of Human Shape An analysis on the moving object is performed to detect a change in the human shape, width to height ration α up to a certain threshold considered to distinguish fall from other activities.

Lack of motion after a fall The second analysis of the moving object is to check if there is a lack of motion just a few seconds after the fall. In the next subsections, we describe in more details each one of these steps.

6.2 Motion Gradient Estimation

As we want to estimate the motion of the person, we compute a directional optical flow vector and estimate the intensity of motion. After that system calculates the global motion orientation of body.

6.3 Analysis of Human Shape

If a large motion is detected and global orientation angle is exceeding a threshold, we analyze more precisely the change in the human shape features to discriminate a fall from another normal activity such as lying down. For this purpose we use shape features α and β. So following two rules

1. α exceeds threshold
2. β changes fast

6.4 Lack of motion after a fall

A last verification is accomplished by checking if the person is motionless on the ground after a possible fall. As soon as a fall is detected, we look for α and β during the 5 seconds following the fall. If an inactive α and β curves are detected with little motion then we confirm the fall. If the α , β and motion gradients still continue to change during these 5 seconds, we consider that this cannot be a fall. All the criterions below threshold must be respected to detect an unmoving α and β:

Few motions in the blob of the person No big motion occurs during specific time duration.

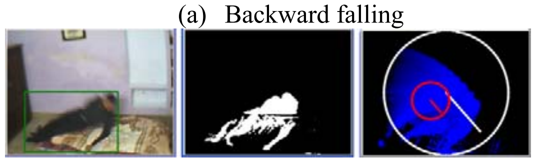
An unchanged α and β we defined unchanged α and β if variation in α and β values are too small

7. Experimental Results

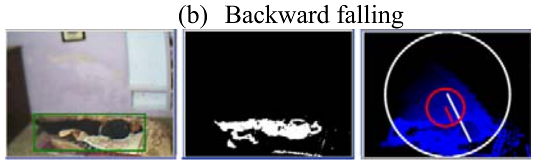
Our system is designed to work with a single uncalibrated camera. As we want a low-cost system, our video sequences were acquired using a USB webcam. As you will see further down, our method gives good results in spite of the low-quality images of our acquisition system. In experiments distance of monitored person to the camera is roughly 4-5 meters. Our fall detection system is implemented in C++ using the OpenCV library [3].

In order to validate the overall system performance we applied the proposed approach to a set of recorded videos.

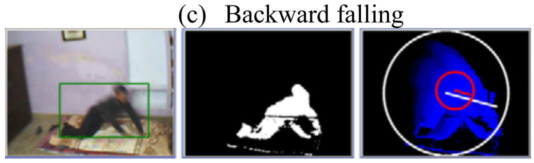
(a) Backward falling



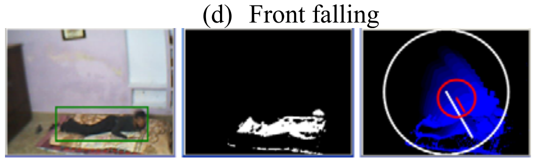
(b) Backward falling



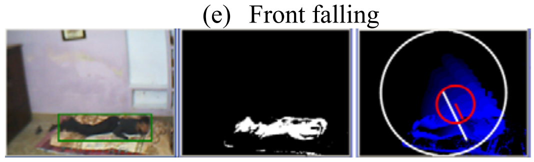
(c) Backward falling



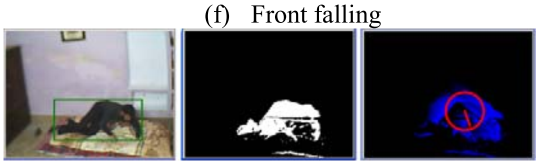
(d) Front falling



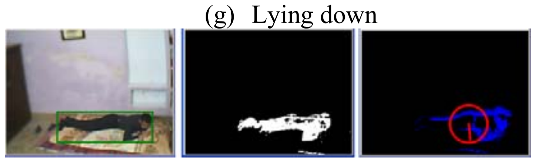
(e) Front falling



(f) Front falling



(g) Lying down



(h) Lying down

Fig 6. Experimental image sequence of fall

7.1 Backward fall:

In the fig. 6(a), (b) & (c), the person falls backward without refraining. This type of fall is extremely dangerous; the person can have a serious head injury. In this simulation, a mattress is used to protect the person during the simulated fall. This fall is detected because of significant global motion in downward direction, and no movement occurs after the fall.

ISBN:978-988-18210-2-7

7.2 Lying down

In the Fig. 6 (d), (e) & (f), the person falls in front and tries to eliminate the fall effect by stretching arms. So in this case intensity of motion is not high as in the case of backward fall. However α and β variation and motion direction $\emptyset$ predict fall easily.

7.3 Front fall

In the Fig. 6 (g) & (h), the person is lying down, In this small motion occurs and β variation is smooth. So result will not be a fall.

Type	Result (+1)	Result(-1)
Front Fall : 20	19	1
Backward Fall : 20	20	0
Lying Down : 20	18	2

Table 1. Recognition results

8. Conclusion and Discussion

Fall recognition for the elders is an important task that can be approached by designing a video sensor network that is capable of segmenting a human from its background and tracking it over time. In proposed scheme to detect unexpected falls of elderly persons and patients introduced. The combination of Motion Gradients and change in shape gives crucial information on human activities. Our proposed fall detection system has proven its robustness on realistic image sequences of simulated falls.

We claim that the developed system has the following distinctive features compared to existent fall detection system. One of the main advantages of the proposed system in comparison with other human fall detection system is that we have considered vast spectrum of angles in which a person fall from static camera view. Our proposed system is best fit for elderly persons living alone. Another part of this system is the communications of data transmission. With the communications of VoIP or GSM, the alarm system will inform specific related people when detecting the falls occurs. The future work contains alarm system with VoIP or GSM communications. The superior outcome will be demonstrated soon in the future discussion.

Acknowledgment

This work was supported by Video and Image Processing Laboratory, UET, Taxila, Pakistan.

References

[1] N. Noury, , "fall detection - Principles and Methods" Conf. Proc. IEEE Eng Med Biol Soc. 2007.

[2] Arie Hans Nasution and S. Emmanuel; Intelligent Video Surveillance for Monitoring Elderly in Home Environments, International Workshop on Multimedia Signal Processing , October 2007.

[3] Intel open source computer vision library http://www.intel.com/research/mrl/research/opencv.

[4] D. Anderson, "Recognizing Falls from Silhouettes". Proceedings of the 28th IEEE EMBS Annual International Conference, Aug 30-Sept 3, 2006

[5] C. Rougier, "Fall Detection from Human Shape and Motion History using Video Surveillance". (AINAW'07)

[6] B. Huang, "A Method for Fast Fall Detection Based on Intelligent Space". Proceedings of the IEEE International Conference on Automation and Logistics,September 2008

[7] J. Q. Lin. "The Behavior Analysis and Detection of Falling". Taiwan, 2004

[8] H. Zhong, M. Visontai, J. Shi, "Detecting Unusual Activity in Video," *IEEE Computer Society Conf. on Computer Vision and Pattern Recognition*, June 2004.

[9] T. Lee and A. Mihailidis, "An intelligent emergency response system: preliminary development and testing of automated fall detection". *Journal of telemedicine and telecare*, 2005.

[10] H. Nait-Charif and S. Mckenna, "Activity summarization and fall detection in a supportive home environment," In International Conference on Pattern Recognition, 2004.

[11] K. Kim and L. Davis. "Real-time foreground-background segmentation using codebook model". *Real-Time Imaging*, June 2005.

[12] Davis, J. "Recognizing movement using motion histograms" MIT Media lab Technical Report. 1999.