Vision-Based Fall Detection Through Shape Features

Chih-Yang Lin
Dept. of Bioinformatics & Medical Engineering,
Asia University,
Taichung, Taiwan
andrewlin@asia.edu.tw

Shang-Ming Wang
Dept. of Computer Science & Information Engineering,
Asia University,
Taichung, Taiwan
interman6113@gmail.com

Jia-Wei Hong
Dept. of Computer Science & Information Engineering,
Asia University,
Taichung, Taiwan
stu2488@gmail.com

Li-Wei Kang
Graduate School of Engineering Science and Technology, and Dept. of Computer Science & Information Engineering,
National Yunlin University of Science and Technology, Yunlin, Taiwan
lwkang@yuntech.edu.tw

Chung-Lin Huang
Department of M-Commerce and Multimedia Applications
Asia University,
Taichung, Taiwan
clhuang@asia.edu.tw

***Abstract*—A major cause of deaths among the elderly relates to accidental falls. Such falls are of particular medical concern to this population because they often result in severe injuries, since senior citizens usually live alone and cannot ask for help when accidents happen. In this paper, we propose a fall detection system with the help of a Gaussian mixture background model to build the background before motion history image (MHI) is applied to analyze the fall behavior. Finally, two extra features, acceleration and angular acceleration, are proposed to more accurately determine whether a fall event has happened.**

Keywords: fall detection; motion history image; motion features

I. INTRODUCTION *(HEADING 1)*

Falls are the main cause of accidental deaths among the elderly. Injuries caused by falls not only harm the elderly, but also increase the burden on caregivers. The causes of falls among this population can be broadly divided into internal and external factors. Internal factors include vision, hearing, and bone degeneration, while external factors include furnishings and environments.

Many fall detection methods in existing research focus on wearable device sensors. However, the elderly usually forget or do not want to wear such sensors. Therefore, researchers have increasingly proposed vision-based detection methods through surveillance cameras to detect fall events. For example, Rougier et al. [1] used wall-mounted cameras to extend the monitor area and detect falls using motion history image (MHI) and human shape variations. The MHI is a fusion of moving images, including crucial information on human activities. The more recent the motion in an image, the more intense the pixel. The variation in the human shape was evaluated by an ellipse shape. Nait-Charif and McKenna [2] tracked head movement to detect fall events based on particle filter algorithms and likewise used the ellipse to analyze fall trajectories. Rougier et al. [3] further proposed silhouette detection (foreground segmentation method) and edge points extraction (Canny edge detector) to obtain an edge image. The moving edge points were then extracted by comparing the current edge image with the previous edge image. Finally, Gaussian mixture model was used to model fall events. The fall result was decided by votes from four cameras. Unfortunately, since Rougier et al.'s method was based on shape matching, a complicated background would have made the matching ineffective.

In this paper, we propose a vision-based detection method to detect falls. The proposed method uses a mixture of Gaussians to segment background and foreground, and applies MHI to analyze fall behaviors. Several fall features, including acceleration and angular acceleration, are proposed to determine fall events. This paper is organized as follows: Section 2 explains the details of the proposed algorithm and framework, Section 3 presents our experimental results, and Section 4 summarizes our work.

II. PROPOSED METHOD

A. *Fall Detection System Overview*

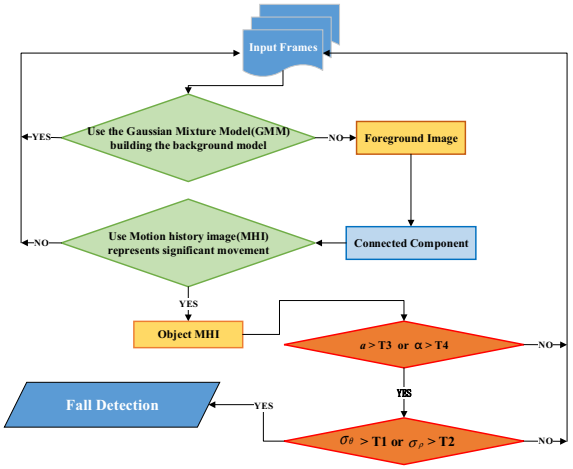

Figure 1. The architecture of the fall detection system.

978-1-5090-2179-6/16 $31.00 © 2016 IEEE
DOI 10.1109/BigMM.2016.22

In this paper, our approach is based on MHI and variations in the shape of the human object. An overview of our fall detection system is shown in Fig. 1, and is detailed in the following sections.

B. Background Modeling

Gaussian mixture model (GMM) is often used to establish a background model and segment foreground objects. Grimson and Stauffer [4] used multiple Gaussian functions to describe a complex background, and to adjust parameters to adapt to changes in lighting. In this paper, GMM is used to segment background and foreground and readers can refer to [4] for more details.

C. Motion History Image

Optical flow [5] analyzes light variations in consecutive images to predict movement areas. However, optical flow is not suitable for real-time detection because it is computationally expensive. Therefore, in this paper, we use motion history image (MHI) to detect falls in real-time environments.

Bobick and Davis [6] proposed using MHI to represent the motion of a body movement. MHI is an image involving information about multiple consecutive images in a period of time. The most recent movement gains the largest intensity, and then the intensity is decreased gradually as time passes. The formula for MHI is shown in Eq. (1), and Fig. 2 shows the MHI images. In Eq. (1), $D(x, y, t)$ is the difference in binary image derived from the current frame and the previous frame. τ is a time range between 1 and N, and N is the number of frames in a period of time. The intensity of each pixel in MHI, H_τ, shows the movement history based on N frames.

$$H_\tau(x,y,t) = \begin{cases} \tau & if\ D(x,y,t) = 1 \\ \max(0, H_\tau(x,y,t-1) - 1) & otherwise. \end{cases} \quad (1)$$

(a) (b)

Figure 2. (a) Walk and (b) Fall event.

This work was supported by Ministry of Science and Technology, Taiwan, under Grants MOST 103-2221-E-468-007-MY2.

D. Human Shape Analysis

The center of the ellipse

The shape of the human body can be approximated by an ellipse using [7]. The ellipse contains center (x, y), with orientation θ, and length a and b of its major and minor semi-axes. The center of an ellipse can be calculated by moments. For a continuous image $f(x, y)$, the moments are given by:

$$m_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} x^p y^q f(x,y) dx dy \quad (2)$$

$$with\ p, q = 0,1,2, \ldots$$

The center of the ellipse can be obtained by computing the coordinates of the center of the mass with the first and zero order spatial moments: $x=m_{10}/m_{00}$, $y=m_{01}/m_{00}$. The centroid (x, y) is used to compute the central moment as follows:

$$\mu_{pq} = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x-\bar{x})^p (y-\bar{y})^q f(x,y) dx dy \quad (3)$$

The angle of the ellipse

The angle of the ellipse serves as key information in fall detection. The angle between the human shape's long axis and horizontal x-axis can be calculated by the following formula:

$$\theta = \frac{1}{2} \arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right) \quad (4)$$

The semi-axes of the ellipse

To recover the major semi-axis, a, and the minor semi-axis, b, of the ellipse, we need to compute I_{min} and I_{max}, which represent the shortest and the longest moments of inertia. They can be computed by evaluating the eigenvalues of the covariance matrix [7]:

$$J = \begin{bmatrix} \mu_{20} & \mu_{11} \\ \mu_{11} & \mu_{02} \end{bmatrix} \quad (5)$$

The eigenvalues I_{min} and I_{max} of the co-variance matrix are given as:

$$I_{min} = \frac{\mu_{20} + \mu_{02} - \sqrt{(\mu_{20} - \mu_{02})^2 + 4{\mu_{11}}^2}}{2} \quad (6)$$

$$I_{max} = \frac{\mu_{20} + \mu_{02} - \sqrt{(\mu_{20} - \mu_{02})^2 + 4{\mu_{11}}^2}}{2} \quad (7)$$

The major semi-axis a and the minor semi-axis b of the ellipse are defined as:

$$\mathrm{a} = \left(\frac{4}{\pi}\right)^{\frac{1}{4}} \left[\frac{(I_{max})^3}{I_{min}}\right]^{\frac{1}{8}} \quad (8)$$

$$b = \left(\frac{4}{\pi}\right)^{\frac{1}{4}} \left[\frac{(I_{min})^3}{I_{max}}\right]^{\frac{1}{8}} \quad (9)$$

The ellipse is used to approximate the whole human shape, and to calculate the angle of a fall in the images. The created ellipse is shown in Fig. 3.

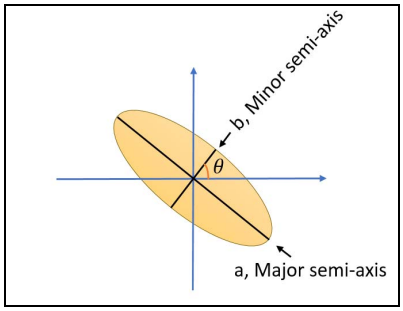

Figure 3. Approximat ellipse.

The ellipse, defined by center (x, y), orientation θ, major and minor semi-axis (a and b), can be used to analyze the human shape.

Fig. 4. shows the approximate ellipse and the corresponding bounding box in MHI. Traditional bounding box is defined by the length and width, scaled by the human shape and size. This definition of the bounding box makes it difficult to detect the change in the shape angle or the change in motion speed when the fall event happens. Therefore, in this paper, we use an ellipse to approximate the movement of the human shape.

(a) (b)

Figure 4. (a) Corresponding bounding box (b) Approximate ellipse for MHI.

Rougier et. al. [1] analyzed changes in the human shapes to distinguish fall events from other activities. If a large motion was detected (C_{motion}>65%), and the orientation standard deviation σ_θ of the ellipse was larger than 15 or the ratio standard deviation σ_ρ was larger than 0.9, a motion was considered a possible fall event.

If the above conditions were satisfied, further steps should be taken to confirm an actual fall event. Therefore, five seconds after a possible fall event happened, Rougier et. al.'s method would verify the fall by checking the ellipse of the human shape against two requirements: C_{motion}>15%, and σ_θ < 15 degrees.

In this paper, two features, acceleration and angular acceleration, are added to the fall conditions. As long as the possible fall event (σ_θ > 15 or σ_ρ > 0.9) satisfies one of the extra conditions, the event will be considered a fall event.

E. Acceleration and Angular Acceleration

In the fall event, there will be large variances in the values of speed and angle. Acceleration is defined as an object moving without a consistent velocity. In real cases, fall events must have different speeds from normal activities. Therefore, acceleration can be used to describe the fall event as shown in Fig. 5. Similarly, angular acceleration as shown in Fig. 6, which shows the difference in angular velocity for the object across a period, can also describe how quickly the object rotates. Acceleration a is defined as follows ($v0$ and $v1$ are the initial velocity and the final velocity, respectively):

$$a = \frac{v1 - v0}{\Delta t}, \quad (10)$$

In addition, the angular acceleration α is defined as follows:

$$\alpha = \lim_{\Delta t \to 0} \frac{\Delta \omega}{\Delta t} = \frac{d\omega}{dt} \quad (11)$$

In this paper, when the acceleration is higher than 10, it will be regarded as a possible fall event; similarly, if the angular acceleration is greater than 0.4, it is a possible fall event.

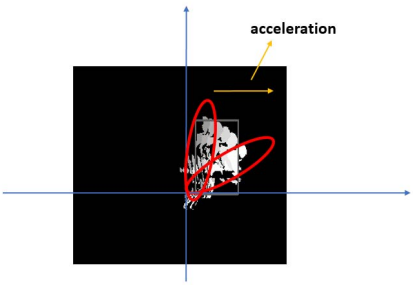

Fig. 5. Example of acceleration.

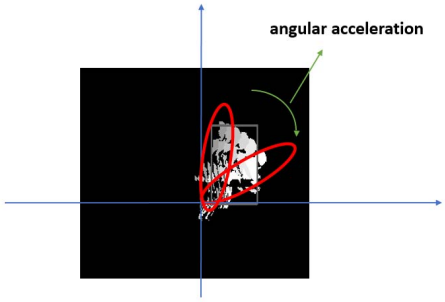

Figure 6. Example of angular acceleration.

III. EXPERIMENTAL RESULTS

In the experiments, several scenarios are tested, including falling, sitting down and walking. The fall criteria in this paper is set to ($\sigma_\theta > 15$ or $\sigma_\rho > 0.9$) and ($a > 10$ or $\alpha > 0.4$). Fig. 7 shows the human shape for a fall with a significant motion. The measured a 、α 、σ_θ and σ_ρ are shown in Fig. 7. This is determined to be a fall event.

Figure 7. Fall: $a = 26.5$, $\alpha = 0.84$, $\sigma_\theta = 15.4$, $\sigma_\rho = 1.33$.

Fig. 8 shows a person who sits down. Although σ_ρ is greater than the threshold, a, α, and σ_θ are below the thresholds, so the event is not considered a fall.

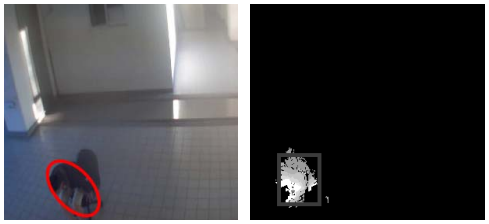

Figure 8. Sit down: $a = 4.12$, $\alpha = 0.31$, $\sigma_\theta = 9.42$, $\sigma_\rho = 0.93$.

Fig. 9 shows the scenario of a person walking slowly. Since all parameters were below the thresholds, there is no fall event detected.

Figure 9. Walk: $a = 0.5$, $\alpha = 0.13$, $\sigma_\theta = 3.18$, $\sigma_\rho = 0.21$.

Fig. 10. is a fall event that moves toward the camera. In Rougier et. al.'s [1] method, the fall event would be triggered when $C_{motion} < 65\%$, σ_θ was between -10 and 10, and σ_ρ between 0.9 and 1. However, the case would not have satisfied the fall requirements defined in Rougier et. al.'s method. In contrast, by incorporating acceleration and angular acceleration, the proposed method can successfully detect this fall scenario.

Another disadvantage of Rougier et. al.'s method involves using the frame difference method to extract moving pixels (foreground object), which is not an appropriate way to accurately model an ellipse. Moreover, C_{motion} is very unstable because it is highly dependent on the results of moving pixels. Instead, the proposed method applies GMM to extract foreground objects, which is more reliable and accurate for constructing an ellipse.

Figure 10. Fall toward the camera: $a = 28.4$, $\alpha = 0.32$, $C_{motion} = 42\%$, $\sigma_\theta = 8.47$, $\sigma_\rho = 0.91$.

IV. CONCLUSIONS

In this paper, we propose a new fall event detection method that applies GMM to construct the background model, and then employs MHI to analyze human activities and integrate the orientation standard deviation and the ratio standard deviation from the ellipse. Finally, acceleration and angular acceleration are proposed to improve fall detection accuracy. All of the employed features are cost-effective, thereby allowing the proposed method to be used in a real-time environment. In future work, we will extend the results to multiple cameras to overcome the problem of different angles or occlusion, further increasing the accuracy of the fall detection system.

REFERENCES

[1] J. L. Barron, D.J. Fleet and S.S. Beauchemin, "Performance of optical flow Techniques," *International Journal of Computer Vision*, vol. 12, no. 1, pp. 43-77, 1994.

[2] A. F. Bobick and J.W. Davis, "The recognition of human movement using temporal templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 3, pp. 257-267, Mar. 2001.

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

[4] W. Pratt. *Digital Image Processing*. 3rd edition, John Wiley & Sons, New York, 2001.

[5] C. Rougier, J. Meunier, A. St-Arnaud, J. Russeau, "Fall detection from human shape and motion history using video surveillance," *Advanced Information Networking and Applications Workshops*, vol. 2, pp. 875-880, 2007.

[6] C. Rougier , J. Meunier , A. St-Arnaud and J. Rousseau, "Robust video surveillance for fall detection based on human shape deformation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 21 , no. 5 , pp.611 -622 , 2011.

[7] C. Stauffer and W. E. L. Grimson, "Adaptive background mixture models for real-time tracking," *In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. 22–29, 1999.