

A HYBRID HUMAN FALL DETECTION SCHEME

Yie-Tarnng Chen, Yu-Ching Lin, Wen-Hsien Fang

Department of Electronic Engineering
National Taiwan University of
Science and Technology
Taipei, Taiwan, R.O.C.

Email: ytchen@mail.ntust.edu.tw, m9602116@mail.edu.tw, whf@mail.ntust.edu.tw

ABSTRACT

This paper presents a novel video-based human fall detection system that can detect a human fall in real-time with a high detection rate. This fall detection system is based on an ingenious combination of skeleton feature and human shape variation, which can efficiently distinguish “fall-down” activities from “fall-like” ones. The experimental results indicate that the proposed human fall detection system can achieve a high detection rate and low false alarm rate.

Index Terms— fall detection, skeleton, human behavior analysis

1. INTRODUCTION

We address the problem of a video-based fall incident detection system for senior citizens. For these people, a fall-down incident normally occurs suddenly, within approximate 0.45 to 0.85 seconds. When that happens, both the posture and shape of the victim change rapidly. The victim then lies on the floor and becomes inactive. Hence, a drastic change in the human posture and shape are important features for human fall detection. However, modeling human postures with low computational complexity is a challenge issue.

Many efforts have been made on fall detection[1]-[4]. Simple features derived from shape analysis such as the aspect ratio of the bounding box[1], a fall angle and a vertical projection histogram have been used for fall detection. Rougier [2] proposed a fall detection approach based on the Motion History Image (MHI) and change in human shape. Hidden Markov Models (HMMs) have also been proposed in fall detection [3]. Hsieh [4] investigated the triangulation-based skeleton extraction approach to analyze human movements. However, this system is not specifically designed for falling-down incident detection. No single approach based on simple features can detect all kinds of human falls. Most existing video-based fall detection systems based on simple features instead of the human skeleton suffer from high

false alarms because they fail to differentiate “fall-like” activities from “fall-down” ones. On the other hand, the high computational cost of human skeleton extraction deters this approach from real-time human fall detection. To provide a reliable fall detection system, an ingenious combination of several approaches, which can increase the detection accuracy while still satisfying the real-time constraint, becomes a major research issue.

Inspired by efficient human skeleton extraction [4] and the shape analysis in fall detection approach [2], we propose a novel video-based human fall detection system that combines posture estimation and shape analysis. We attempt to use the skeletal information to differentiate “fall-down” from “fall-like” activities. To alleviate the high computational cost of human skeleton extraction, we extract human posture based on the 2-D skeleton models instead of the complex 3-D ones. We then apply the Douglas-Peucker algorithm to reduce the number of pixels in the human contour, which can effectively speed up the human skeleton extraction. Simultaneously, we acquire the human skeleton every 0.4 seconds instead of in each frame. The major contribution of this paper is to propose a novel real-time fall detection approach, which is an intelligent combination of different fall detection approaches to increase the detection accuracy while still satisfying the real-time constraint.

The rest of this paper is organized as follows. In Section 2, the fall detection model is proposed. The results of experiments are presented in Section 3, and Section 4 concludes this paper.

2. FALL-DOWN DETECTION SCHEME

Foreground object segmentation and object tracking with occlusion handling are important steps before human fall detection. The human skeleton is first extracted through the depth-first search of Delaunay meshes. The distance between the two sampling skeletons beyond a threshold determines a posture change. We use an ellipse to approximate the human shape and the orientation of the ellipse and the ratio of its ma-

Thanks to this work was supported by National Science Council of R.O.C. under contract NSC 98-2221-E-011-071.

jor and minor semi-axes to detect the human shape change. We then confirm a human fall incident by monitoring the inactivity of a person for a period of time.

2.1. Posture Analysis with Human Skeleton

We use the Douglas-Peucker algorithm to approximate the contour of the foreground object with fewer vertices. We next perform the constrained Delaunay triangulation to partition the foreground object into triangular meshes, and extract the human skeleton through the depth first search on centers of triangular meshes. Finally, we calculate a distance map of the human skeleton, and detect the posture change by calculating the distance map of two human skeletons every 0.4 second.

2.1.1. Douglas-Peucker Algorithm

The Douglas-Peucker algorithm approximates a curve by recursively dividing the poly-line. In the beginning, we connect the two endpoints of a poly-line as the initial approximation and calculate the perpendicular distance of each intermediate point to the approximation line. If each distance is less than a predefined threshold, ξ , the approximation straight line is an acceptable solution. All endpoints are kept, and other intermediate points are discarded. Otherwise, if any perpendicular distance is larger than a threshold, ξ , the approximation is then unacceptable and we continue to choose a point that is farthest away as a new point and subdivide the original poly-line into two shorter poly-lines. This process continues until the approximation is acceptable.

2.1.2. Delaunay Triangulation

Definition: let S be a set of points in the plane of triangulation. T is a Delaunay triangulation of S if for each edge of T , there exists a circle C with the following properties: (1) endpoints of edge e are on the boundary of C , and (2) no other vertex of S is in the interior of C .

After simplifying the human contour, we use the Delaunay triangulation, well-studied in computational geometry, to partition the foreground object into triangular meshes. The details of the Delaunay triangulation algorithm can be found in [4].

2.1.3. Human Skeleton Extraction

We first connect all triangle meshes and obtain the centroid of each triangular mesh. We can then find a spanning tree as the human skeleton via the depth first search from the root node of the triangular meshes [4].

Human Skeleton Extraction Algorithm:

1. Calculate the centroid C of a human posture.

2. Construct a graph G by connecting all centers of any two connected triangular meshes, which share a common edge.
3. Find the node with the largest y coordinate and degree = 1 among all nodes in G as the root node R .
4. Perform the depth first search on G and obtain a spanning tree F .
5. Find all leaf nodes L and branch nodes B of the spanning tree F .
6. Extract the skeleton S by connecting any two nodes in $\{R, L, B, C\}$ if there exists a path between these two nodes.

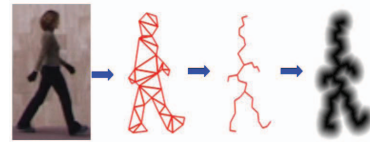


Fig. 1. an example of the Delaunay triangulation, the human skeleton extraction, and the distance map of a skeleton

2.1.4. Distance Map of Human Skeleton

The distance map, also known as the distance transform, is employed to compute the distance of two sampling skeletons. In the distance transform of a binary map, the value of a pixel is the shortest distance to all pixels in the foreground object. The distance map of the binary image S_1 , denoted as DM_{S_1} , can be represented as follows :

$$DM_{S_1}(p) = \min_{q \in S_1} Dist(p, q) \quad (1)$$

where $Dist(p, q)$ is the Euclidian distance between pixel p and pixel q . Before calculating the distance between two skeletons, we must normalize them to the same size. Consequently, the distance of the two skeletons, S_1 and S_2 , denoted as $Dist(S_1, S_2)$, can be calculated as follows :

$$Dist(S_1, S_2) = \frac{1}{|DM|} \sum_p |DM_{S_1}(p) - DM_{S_2}(p)| \quad (2)$$

where $|DM|$ represents the image size of the distance map.

2.2. Human Shape Analysis

We use an ellipse to approximate the human shape [5] instead of a bounding box. An ellipse can provide more precise shape information than a bounding box, especially, when a human carries an object.

2.2.1. Ellipse Fitting

We can obtain the contour of a human and approximate the human shape with an ellipse using moments [2]. The ellipse is defined by its center (\bar{x}, \bar{y}) , the orientation θ and the length I_a and I_b of its major and minor semi-axes.

For a gray level image value $f(x, y)$, the moments are given by :

$$m_{p,q} = \int \int x^p y^q f(x, y) dx dy \quad p, q = 0, 1, 2.. \quad (3)$$

The center of gravity is obtained by computing the coordinates of the center of mass with the first and spatial moments of zero-order: The center (\bar{x}, \bar{y}) is used to compute the central moment as follows :

$$\mu_{p,q} = \int \int (x - \bar{x})^p (y - \bar{y})^q f(x, y) dx dy \quad (4)$$

The orientation of the ellipse is given by the tilt angle between the major axis and the x-axis of the person, and it can be computed with the second order central moments:

$$\theta = \frac{1}{2} \arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right) \quad (5)$$

We then compute the major and the minor semi-axes of the ellipse, which corresponding, respectively, the maximum and minimum eigenvalues of the covariance matrix:

$$J = \begin{pmatrix} \mu_{2,0} & \mu_{1,1} \\ \mu_{1,1} & \mu_{0,2} \end{pmatrix} \quad (6)$$

The maximum eigenvalue I_{min} and the minimum eigenvalue I_{max} are given ,respectively, by :

$$I_{min} = \frac{\mu_{2,0} + \mu_{0,2} - \sqrt{4\mu_{1,1}^2 + (\mu_{2,0} - \mu_{0,2})^2}}{2} \quad (7)$$

$$I_{max} = \frac{\mu_{2,0} + \mu_{0,2} + \sqrt{4\mu_{1,1}^2 + (\mu_{2,0} - \mu_{0,2})^2}}{2} \quad (8)$$

The major and the minor semi-axes of the ellipse are then given, respectively, by :

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

$$I_b = \left(\frac{4}{\pi}\right)^{\frac{1}{4}} \left(\frac{I_{min}^3}{I_{max}}\right)^{\frac{1}{8}} \quad (10)$$

Finally, the ratio of the major and the minor semi-axes of the ellipse, γ , can be defined as follows

$$\gamma = \frac{I_a}{I_b} \quad (11)$$

2.2.2. Ellipse Features for Fall-down Detection

Two features, derived from the orientation of the ellipse, θ , and the ratio of major and minor of the ellipse semi-axes, γ , are used to discriminate a falling incident from daily normal activities. To avoid the influence of the size of foreground objects on feature thresholds, we use the change rate of the ellipse features within a slide window instead of a fixed threshold.

The change rate in the orientation of the ellipse, R_θ , and the ratio of major and minor of the ellipse semi-axes, R_γ , are then given in (12) and (13) respectively.

$$R_\theta = \frac{|\theta(SW) - \theta(n)|}{\theta(SW)} \quad (12)$$

where $\theta(n)$ represents the orientation of the object's ellipse in the n -th frame, and $\theta(SW)$ represents the average orientation in a slide window.

$$R_\gamma = \frac{|\gamma(SW) - \gamma(n)|}{\gamma(SW)} \quad (13)$$

where $\gamma(n)$ represents the proportion of the object's ellipse in the n -th frame, and $\gamma(SW)$ represents the average proportion of the object's ellipse in a slide window.

2.3. Fall-Down Confirmation

The final verification of a fall-down incident is to check if the person is inactive for a period of time after a possible fall. We check the following two conditions to confirm to a fall-down.

1. The motion of the human object is smaller than the threshold for a period of time.
2. Both R_θ in (12) and R_γ in (13) are smaller than the threshold for a period of time.

3. EXPERIMENTAL RESULTS

We implement the proposed skeleton-based fall detection system on Intel's OpenCV library. All test videos are acquired from a single stationary and un-calibrated camera in the MPEG-1 format with 320*240 pixels resolution, and 30 frame rates per second. Human activities in testing videos include fall incidents such as forward falls, backward falls, and side-way falls, and daily activities such as walking, running, and squatting. Our experiments were run on a computer with windows XP, Intel Pentium D 3.2GHz CPU and 2 GB RAM.

Performance metrics in fall detection experiments are the detection rate and the false alarm rate.

$$\text{detection rate} = \frac{TP}{(TP + FN)}$$

$$\text{false alarm rate} = \frac{FP}{(TN + FP)}$$

Table 1. The experimental results of the proposed fall detection with the skeleton distance = 0.056

events	video no.	detected falling	detected non-falling
falling	22	20	2
sit-down	8	2	6
squat	8	0	8
walking	8	0	8
running	8	0	8

where true positive (TP) and true negative (TN) are the counts of correct detection while false positive (FP) and false negative (FN) are the counts of incorrect prediction.

In our experiments, we compared the proposed fall detection system with three different fall detection approaches: the skeleton match [4], the change rate of sampling skeletons and the shape analysis [2]. The shape analysis approach only uses the change rate in the ellipse angle and the ratio between the major and minor semi-axes of the ellipse to detect human falling.

The experimental results for the proposed falling detection are shown in Table 1. Two fall-down incidents cannot be detected because they are slow-speed fall incidents, which cannot flag as a fast posture change in the proposed approach. On the other hand, two brutally sit-down activities are flagged fall incidents. Table 2 summarizes the performance of the four human fall detection approaches in terms of the detection rate and the false alarm rate. Table 3 compares the proposed hybrid human detection approach with the shape analysis approach [2] in terms of the execution time, the detection rate and the false alarm rate. We can observe from Tables 2 and 3 that the proposed approach, a combination of different fall detection schemes, can achieve high detection accuracy and lower false alarm rate within a reasonable execution time.

4. CONCLUSIONS

We presented a novel video-based human fall detection system, which can achieve real-time detection while still maintaining high detection accuracy. This system combines posture analysis and shape analysis to detect fall-down incidents. A human skeleton is first extracted from a human

Table 2. The comparisons of four fall detection schemes

detection schemes	detection rate	false alarm rate
the skeleton match	0.727	0.031
the change rate of sampling skeletons	1	0.34
the shape analysis	0.909	0.16
the proposed scheme	0.909	0.06

Table 3. Comparison of the proposed scheme and the shape analysis in terms of the execution time, the detection rate and the false alarm rate

metrics	the proposed scheme	the shape analysis
detection rate	90.9%	90.9 %
false alarm rate	6.25 %	15.6 %
execution Time	4.21 sec	1.36 sec

posture. The distance between two sampling skeletons beyond a threshold flags a posture change. We then used an ellipse to approximate the human shape. The orientation and the ratio of the major and the minor semi-axes of the ellipse are used to detect the human shape change. We confirmed a human fall incident by monitoring the inactivity of a person for a period of time. The contribution of this paper is to combine the skeleton change and shape change into the proposed fall-down detection framework to provide a reliable fall detection system. Experimental results indicate that the proposed human fall detection system can achieve a high detection rate and low false alarm rate with reasonable computational costs.

5. REFERENCES

- [1] A. William, D. Ganesan, and A. Hanson, "Aging in place: fall detection and localization in distributed smart camera network," in *Proc. the 15th Int. Conf. on Multimedia*, 2007, pp. 892 – 901.
- [2] Caroline Rougier, Jean Meunier, Alain St-Arnaud, and Jacqueline Rousseau., "Fall detection from human shape and motion history using video surveillance," in *Proc. 21st Int. Conf. on Advanced Information Networking and Applications Workshops*, 1987, pp. 215–222.
- [3] B. U. Toreyin, Y. Dedeoglu, and A. E. Cetin, "HMM-based falling person detection using both audio and video," in *Proc. Int. workshop on Human-Computer Interaction 2005*, 2005, pp. 211 – 220.
- [4] Y. T. Hsu, H. Y. M. Liao, C. C. Chen, and J. W. Hsieh, "Video-based human movement analysis and its application to surveillance systems," *IEEE Transactions on Multimedia*, vol. 10, pp. 372–392, March 2008.
- [5] and E. Ramasso C. Panagiotakis and G. Tziritas, "Shape-motion based athlete tracking for multilevel action recognition," in *Proc. Int. Conf. on Articulated Motion and Deformable object*, 2006, pp. 385 – 394.