A VIDEO-BASED HUMAN FALL DETECTION SYSTEM FOR SMART HOMES

Yie-Tarng Chen*, Yu-Ching Lin, and Wen-Hsien Fang

ABSTRACT

In recent years, the global population has begun to age rapidly. Automatic fall detection for senior citizens has become an important issue for smart homes. 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 features and human shape variations, which can efficiently distinguish "fall-down incidents" 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.

Key Words: smart home, fall detection, shape analysis.

I. INTRODUCTION

The aging global population has rapidly increased in recent years. In Taiwan, in 2007 the senior citizen population was more than 10 percent of the total population. It is expected that the senior citizen population will reach 3.68 million in 2020. Hence, an automatic human fall detection system for senior citizens has become an important issue for smart homes.

We address this problem using 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 as shown in Fig. 1. 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 challenging issue. In previous research, simple features derived from shape analysis were used for fall detection instead of the human skeleton, a popular approach in human behavior analysis. The high computational cost of human skeleton extraction deters this approach from use in real-time human fall 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 suffer from high false alarm rates because they fail to differentiate "fall-like" activities from "fall-down" ones. To provide a reliable fall detection system, an ingenious combination of several approaches is necessary. Hence, the intelligent combination of different fall detection approaches, which can increase detection accuracy while still satisfying the real-time constraint, has become a major research issue in human fall detection.

Many efforts have been made concerning fall detection. Wearable sensor-based devices (Zhang *et al.*, 2006) such as accelerometers (Lindemann *et al.*, 2005) have been used to detect abnormal acceleration. However, this approach becomes useless if senior citizens do not wear the sensor device or the batteries within the sensors are dead. Several video-based fall detection technologies have been developed. For instance, the aspect ratio of the bounding box (William *et al.*, 2007; Töreyin *et al.*, 2005; Vishwakarma *et al.*, 2007), the ratio between the x-axis and y-axis, is used to detect a fall. When a fall-down incident occurs, the aspect ratio of the bounding box will change. A fall angle, the angle between ground and the person, is another popular approach (Vishwakarma *et al.*, 2007). When this

*Corresponding author. (Tel: 886-2-27376420; Fax: 886-2-27376424; Email: ytchen@mail.ntust.edu.tw)

The authors are with the Department of Electronic Engineering National Taiwan University of Science and Technology, No. 43, Sec. 4, Keelung Road, Taipei 10607, Taiwan, R.O.C.

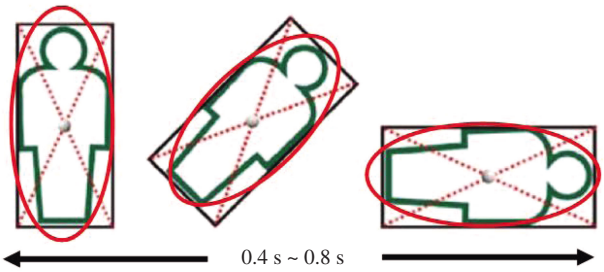

Fig. 1 The features of the fall-down incident

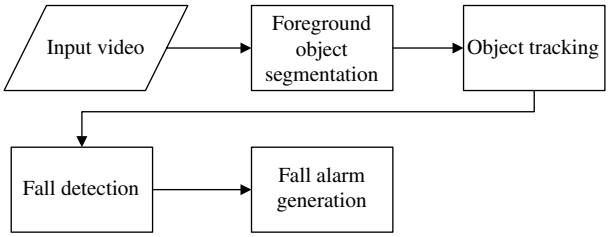

Fig. 2 The flowchart of the proposed fall detection system

angle is less than 45 degrees, a fall alarm is triggered. However, this approach may fail if the person falls toward the camera. The centroid of the person is also used to detect a fall-down incident. During a fall, the centroid changes significantly and rapidly. In light of this, a vertical projection histogram, $V(x)$, is defined as follows:

$$H(x, y) = \begin{cases} 1, & \text{if } (x, y) \text{ is a pixel within the human object;} \\ 0, & \text{otherwise.} \end{cases} \quad (1)$$

$$V(x) = \sum_{y} H(x, y) \quad (2)$$

has been suggested in (Lin *et al.*, 2007) to detect a fall. When a fall occurs, the vertical projection histogram will change significantly. Horizontal and vertical gradients (Vishwakarma *et al.*, 2007) are yet another technique to detect the fall. When a fall-down incident occurs, the vertical gradient is less than the horizontal gradient. Several Hidden Markov Models (HMMs) have been proposed for use in fall detection (Töreyin *et al.*, 2005). In (Töreyin *et al.*, 2005), three-state HMMs are used to classify fall events. The feature parameters of HMMs are extracted from temporal wavelet signals describing the bounding box of the moving object. Anderson also used multiple features extracted from the silhouette such as the magnitude of motion vector, the determinant of the covariance matrix and the ratio of width to height of the bounding box to train HMMs to verify walking, kneeling, getting-up and falling. Hiseh *et al.* (2008) proposed the triangulation-based skeleton extraction approach to analyze human movements. However, this system is not specifically designed for falling-down incident detection. Rougier *et al.* (2007) proposed a fall detection approach based on the Motion History Image (MHI) (Bobick *et al.*, 2001) and changes in human shape.

Inspired by efficient human skeleton extraction (Hsu *et al.*, 2008) and shape analysis in fall detection (Rougier *et al.*, 2007), we propose a novel video-based human fall detection system that combines posture estimation and shape analysis. We attempt to use 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 2-D skeleton models instead of complex 3-D ones despite the fact that 3-D human models provide more information than 2-D ones. We then apply the Douglas-Peucher algorithm (Douglas *et al.*, 1973) to reduce the number of pixels in the human contour, which can effectively speed up human skeleton extraction since the computational cost of the 2-D extraction is proportional to the number of pixels in the human contour. Simultaneously, we acquire the human skeleton every 0.4 seconds instead of in each frame. Consequently, we can detect the change in skeletons in human fall detection instead of skeleton matching. The objective of the human fall detection system is to detect a fall-down incident rather than a fall-down posture. In summary, the proposed fall detection scheme consists of posture change detection, shape change detection and inactivity detection. We use a human skeleton to detect a large posture change. We then use an ellipse to fit a human body and use the orientation and the ratio of height and width of the ellipse to analyze the change in human shape. Finally, we confirm the human fall incident by sensing immobility in a person for a period of time. The major contribution of this paper is to propose a novel real-time fall detection approach for elderly people, which is an intelligent combination of three fall detection approaches to increase detection accuracy while still satisfying the real-time constraint.

The rest of this paper is organized as follows. Section II proposes a hybrid fall detection scheme. The results of experiments are presented in Section III, and Section IV concludes this paper.

II. FALL-DOWN DETECTION SCHEME

Foreground object segmentation and object tracking with occlusion handling are important steps before human fall detection as shown in Fig. 2. Light conditions or the shadows of moving objects can

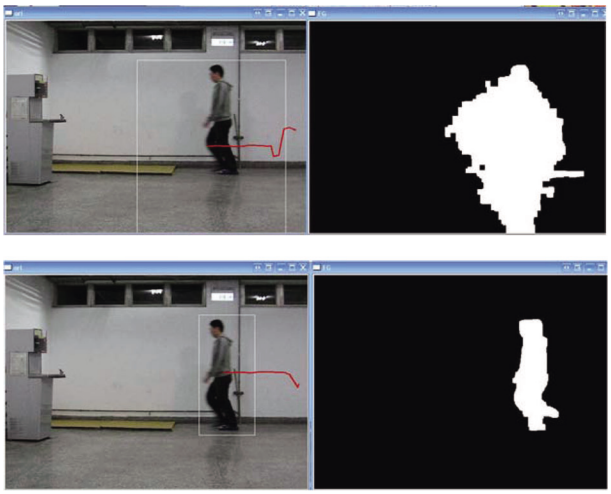

Fig. 3 The comparison of the background subtraction and the Bayesian approach under the shadow of a foreground object

seriously affect the accuracy of foreground object segmentation. Moreover, looking at an object from a different visual angle can produce a different shape. The same scene may include several objects in an occlusion situation where a far object is covered by a near object. In foreground object segmentation, we use the Bayesian approach (Huang *et al.*, 2003) instead of the popular background subtraction approach because shadows of moving objects can seriously affect object segmentation in the background subtraction approach. A comparison of the background subtraction approach and the Bayesian approach under the shadow of a foreground object is shown in Fig. 3. However, the details of foreground object segmentation and object tracking are beyond the scope of this paper.

The proposed human fall detection system, shown in Fig. 4, integrates posture analysis, shape analysis and inactivity detection. We assume that a person is immobile on the floor after a fall and all video sequences are captured from a stationary camera. 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 major 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.

1. Posture Analysis with Human Skeleton

We use the Douglas-Peucker algorithm (Douglas *et al.*, 1973) to approximate the contour of foreground objects with fewer vertices. We next perform

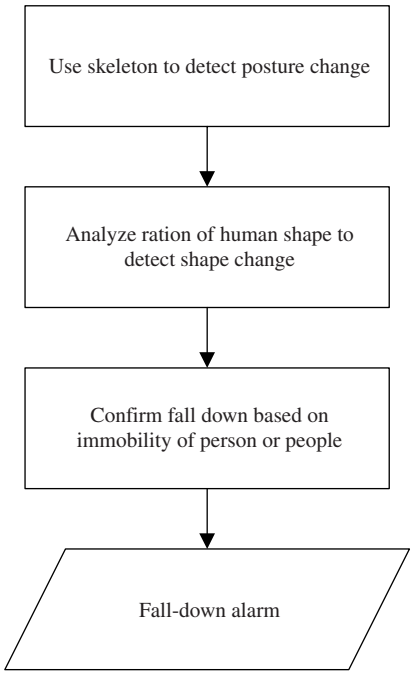

Fig. 4 The flowchart of the fall-down detection

constrained Delaunay triangulation to partition a foreground object blob 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. The human skeleton extraction procedure is illustrated in Fig. 5.

(i) Douglas-Peuker 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 with the largest perpendicular distance as a new point and subdivide the original poly-line into two shorter poly-lines. This process continues until the approximation is acceptable. An example of the Douglas-Peucker algorithm is presented in Fig. 6.

(ii) Delaunay Triangulation

Definition: let S be a set of points in the plane

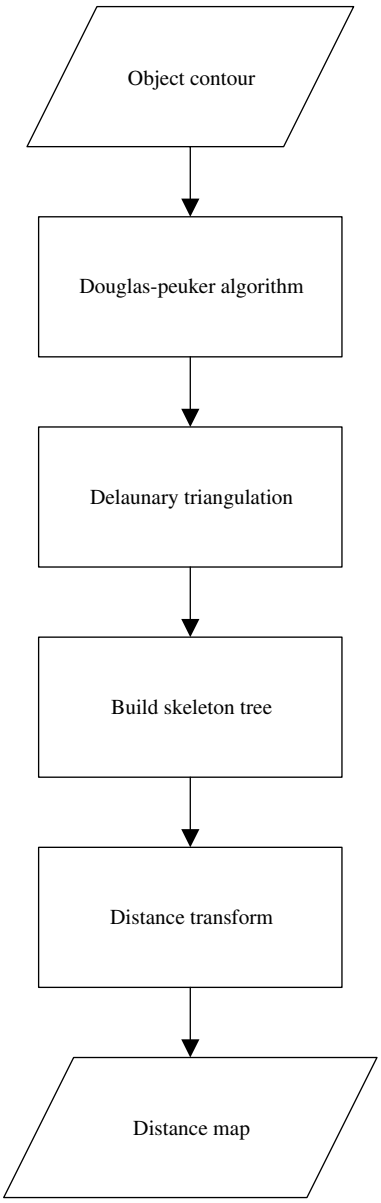

Fig. 5 The flowchart of the human skeleton extraction

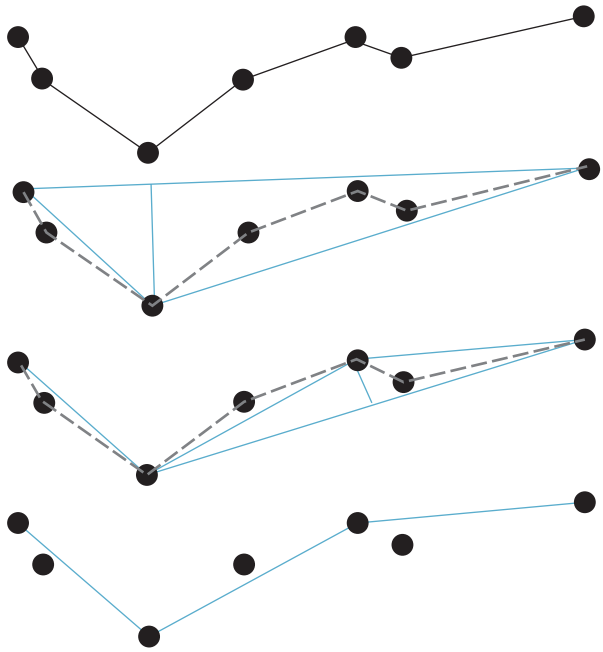

Fig. 6 An example of the Douglas-Peuker algorithm

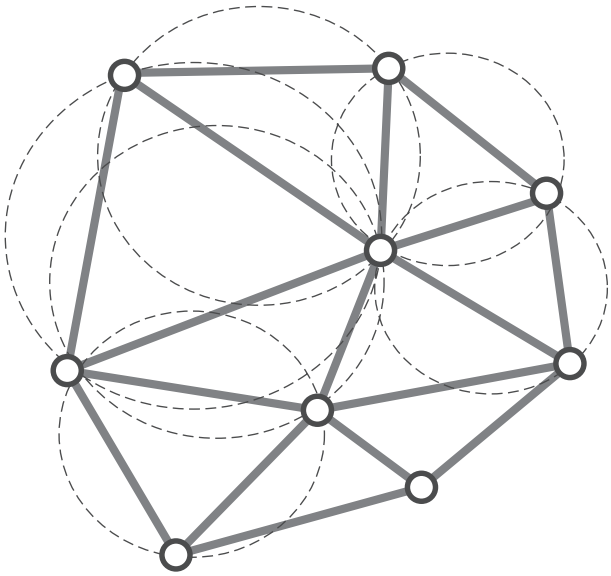

Fig. 7 The Delaunay Triangulation

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.

Definition: Let G be a straight-line planar graph. A triangulation T is a constrained Delaunay triangulation of G if each edge of G is an edge of T and for each remaining edge e of T, there exists a circle C with the following properties.

(1) endpoints of edge e are on the boundary of C, and
(2) if any vertex v of G is in the interior of C, then it cannot be "seen" from at least one of the end points of e.

After simplifying the human contour, we use Delaunay triangulation, well-studied in computational geometry, to partition foreground objects into triangular meshes. Fig. 7 illustrates the Delaunay Triangulation. We can obtain the Delaunay meshes of a human object as follows.

(1) Select boundary nodes of Delaunay meshes:
A polygon is used to approximate the boundary

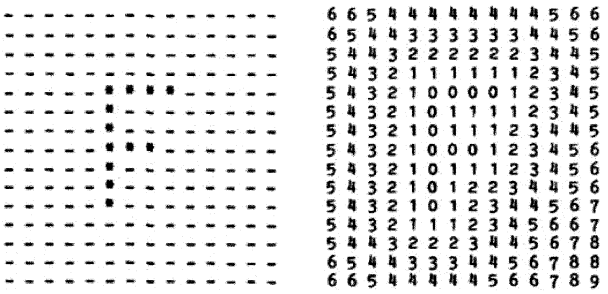

Fig. 8 An example of the distance map

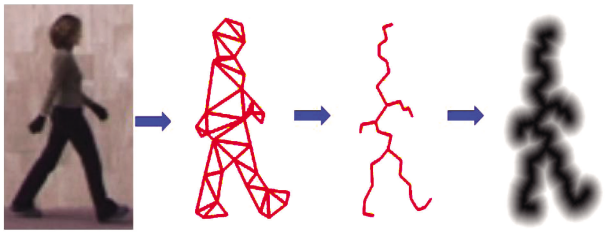

Fig. 9 An example of the Delaunay triangulation, the human skeleton extraction, and the distance map of a skeleton

of a human object. The polygon vertices are the boundary nodes of the Delaunay meshes. We use a simple heuristic to select boundary points with high curvatures as the boundary nodes of the Delaunay meshes.

(2) Choose the interior nodes:

The edge points or corners within the object's boundary are chosen as the interior nodes of the Delaunay meshes.

(3) Perform Delaunay triangulation:

A constrained Delaunay triangulation is performed on the boundary nodes and on the interior nodes. With the bounding polygon as a constraint, the triangulation uses line segments to connect consecutive boundary nodes as edges and form triangles only within the boundary.

The details of the Delaunay triangulation algorithm can be found in (Chew, 1987).

(iii) 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 (Hsu *et al*, 2008). The details of the human skeleton extraction algorithm can be described as follows

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.

(iv) Distance Map of a Human Skeleton

The distance map, also known as the distance transform (Borgefors, 1986), is employed to compute the distance between 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:

$$DM_{S_1}(p) = \min_{q \in S_1} Dist(p, q), \quad (3)$$

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:

$$Dist(S_1, S_2) = \frac{1}{|DM|} \sum_p \left| DM_{S_1}(p) - DM_{S_2}(p) \right|, \quad (4)$$

where $|DM|$ represents the image size of the distance map. Fig. 8 illustrates the distance map, and Fig. 9 illustrates the Delaunay triangulation, the human skeleton extraction, and the distance map of a skeleton.

2. Human Shape Analysis

We use an ellipse to approximate the human shape instead of a bounding box. An ellipse can provide more precise shape information than a bounding box, especially, when a human carries an object. The comparison between an ellipse and a bounding box in shape analysis is shown in Fig. 10.

(i) Ellipse Fitting

We can obtain the contour of a human and approximate the human shape with an ellipse using moments (Pratt, 2007). The ellipse is defined by its center $(\bar{x}, \bar{y})$, the orientation θ and the length I_a and

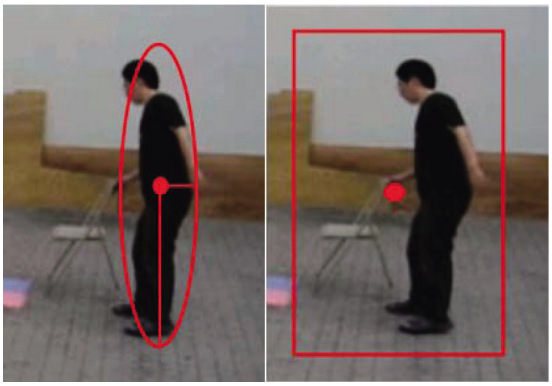

Fig. 10 The comparison between an ellipse and a bounding box in the shape analysis

I_b of its major and minor semi-axes.

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

$$m_{p,q} = \iint x^p y^q f(x, y) dxdy \quad p, q = 0, 1, 2. \tag{5}$$

The center of gravity is obtained by computing the coordinates of the center of mass with the first and spatial moments of zero-order: $\bar{x} = m_{10} / m_{00}$, $\bar{y} = m_{01} / m_{00}$

The center $(\bar{x}, \bar{y})$ is used to compute the central moment as follows:

$$\mu_{p,q} = \iint (x - \bar{x})^p (y - \bar{y})^q f(x, y) dxdy . \tag{6}$$

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:

$$\theta = \frac{1}{2} \arctan\left(\frac{2\mu_{11}}{\mu_{20} - \mu_{02}}\right) . \tag{7}$$

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

$$\boldsymbol{J} = \begin{pmatrix} \mu_{2,0} \mu_{1,1} \\ \mu_{1,1} \mu_{0,2} \end{pmatrix} \tag{8}$$

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

$$I_{min} = \frac{\mu_{2,0} + \mu_{0,2} - \sqrt{4\mu^2_{1,1} + (\mu_{2,0} - \mu_{0,2})^2}}{2} , \tag{9}$$

$$I_{max} = \frac{\mu_{2,0} + \mu_{0,2} + \sqrt{4\mu^2_{1,1} + (\mu_{2,0} - \mu_{0,2})^2}}{2} . \tag{10}$$

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

$$I_a = \left(\frac{4}{\pi}\right)^{\frac{1}{4}} \left(\frac{I^3_{\max}}{I_{\min}}\right)^{\frac{1}{8}} , \tag{11}$$

$$I_b = \left(\frac{4}{\pi}\right)^{\frac{1}{4}} \left(\frac{I^3_{\min}}{I_{\max}}\right)^{\frac{1}{8}} . \tag{12}$$

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

$$\gamma = \frac{I_a}{I_b} . \tag{13}$$

(ii) Ellipse Features for Fall-down Detection

Two features, derived from the orientation of the ellipse, θ, and the ratio of major and minor ellipse semi-axes, γ, are used to discriminate a falling incident from normal daily 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 ellipse semi-axes, P_γ, are then given in Eqs. (14) and (15) respectively:

$$R_\theta = \frac{\left| \theta(SW) - \theta(n) \right|}{\theta(SW)} , \tag{14}$$

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:

$$R_\gamma = \frac{\left| \gamma(SW) - \gamma(n) \right|}{\gamma(SW)} , \tag{15}$$

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.

If a person falls at an angle to the camera optical axis, the change in orientation of the ellipse will be extremely large and the change ratio in the orientation, R_θ, becomes relatively high. If a fall-down incident does not occur, R_θ is relatively low. On the other hand, if a person falls in line with the camera optical axis, the change ratio of the proportion of the ellipse's length and width, R_γ, becomes relatively high. If it is not a fall-down incident, R_γ will be relatively low. To conform with the definition of a fall-down incident, we set the period of the time window (SW) from 0.4 second to 0.8 second in the calculation of R_θ and R_γ.

3. Fall-Down Confirmation

We check the following two conditions to confirm a fall-down.

Table 1 the confusion matrix

	Detected falling	Detected normal
Actual falling	True Positive (TP)	False Negative (FN)
Actual normal	False Positive (FP)	True Negative (TN)

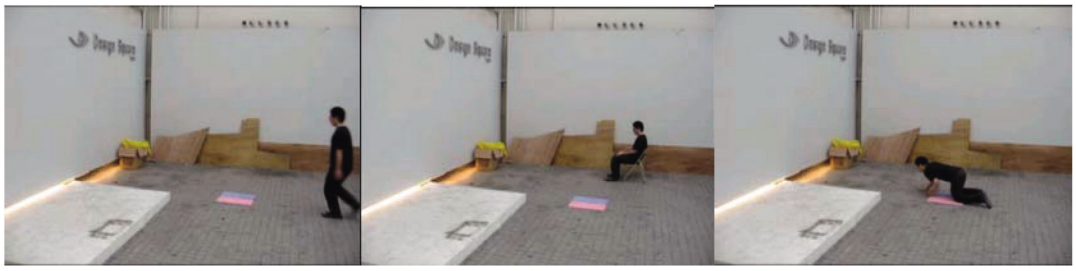

Fig. 11 Parts of the testing videos in experiments: walking, sit-down, and fall

(1) The motion of the human object is smaller than the threshold for a period of time.
(2) Both R_θ in Eq. (14) and R_γ in Eq. (15) are smaller than the threshold for a period of time.

The first condition assures that the person is inactive for a period of time after a possible fall, whereas, the second condition guarantees that both the change rate in the orientation of the ellipse, which approximates the human shape, and the ratio of major and minor ellipse semi-axes approach zero within a fixed time interval, i.e., the human shape becomes immobile after a possible fall.

III. EXPERIMENTAL RESULTS

We implemented the proposed skeleton-based fall detection system on Intel's OpenCV library (Bradski *et al.*, 2008), and decoded the compressed video by means of the FFMPEG library. All test videos were acquired from a single stationary and uncalibrated camera in the MPEG-1 format with 320 × 240 pixels resolution, and 30 frames per second. Human activities in testing videos include fall incidents such as forward falls, backward falls, and sideways falls, and daily activities such as walking, running, and squatting as shown in Fig. 11. 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 can be presented in a confusion matrix, as in Table 1, 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. The detection rate, the fraction of human fall events that are correctly detected, is defined by

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

The false alarm rate, the fraction of non-fall events that are incorrectly predicted, can be defined by

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

1. Experimental Results with Different Thresholds

In the proposed fall detection system, thresholds in human fall detection such as skeleton change detection, ellipse orientation change detection, ellipse aspect ratio change detection and inactivity detection can influence the detection rate and the false alarm rate. To tune the performance of the proposed fall detection system, we conducted a series of experiments under different skeleton distance thresholds as shown in Table 2. In a trade-off between the detection rate and the false alarm rate, the skeleton distance = 0.056 can tune the system to a condition with a detection rate = 90.91% and a false alarm rate = 9.09%. Table 3 shows major system thresholds for the proposed fall detection system in our experiments.

2. Performance Comparison Between Different Approaches

In our experiments, we compared the proposed fall detection system with three different fall detection approaches:

- the skeleton match (Hsu *et al.*, 2008),
- the posture analysis, and
- the shape analysis (Rougier *et al.*, 2007).

In the posture analysis, a distance map of two sampling human skeletons is calculated every 0.4

Table 2 The experimental results of the proposed fall detection with different skeleton distance thresholds

Skeleton distance	0.054	0.055	0.056	0.057	0.058	0.059	0.060
TP	22	20	20	20	19	18	17
FN	0	2	2	2	2	4	5
TN	27	30	30	30	30	31	31
FP	5	2	2	2	2	1	1
Detection rate	100%	90.91%	90.91%	90.91%	90.91%	81.82%	77.27%
False alarm rate	15.63%	6.25%	6.25%	6.25%	6.25%	3.12%	3.12%

Table 3 system thresholds in fall detection experiments

Parameter	Threshold
Change ratio of orientation,(R_θ)	0.28
Change ratio of proportion, (R_γ)	0.3
Skeleton distance, (Dist(S_1, S_2))	0.056
Unmoving time	3 sec

Table 4 The experimental results of the skeleton match

Event	No. of videos	Detected falling	Detected non-falling	TP	TN	FP	FN
Falling	22	16	6	16	0	0	6
Sit-down	8	1	7	0	7	1	0
Squat	8	0	8	0	8	0	0
Walking	8	0	8	0	8	0	0
Running	8	0	8	0	8	0	0

Table 5 The experimental results of the posture analysis

Event	No. of videos	Detected falling	Detected non-falling	TP	TN	FP	FN
Falling	22	22	0	22	0	0	0
Sit-down	8	8	0	0	0	8	0
Squat	8	0	8	0	8	0	0
Walking	8	3	5	0	5	3	0
Running	8	0	8	0	8	0	0

Table 6 The experimental results of the shape analysis

Event	No. of videos	Detected falling	Detected non-falling	TP	TN	FP	FN
Falling	22	20	2	20	0	0	2
Sit-down	8	5	3	0	3	5	0
Squat	8	0	8	0	8	0	0
Walking	8	0	8	0	8	0	0
Running	8	0	8	0	8	0	0

second, and a human fall is detected if this distance map is larger than a threshold, i.e., only a drastic change in the human posture as discussed in Section II.1 is accepted as showing human falling. On the other hand, the shape analysis 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, i.e., only a drastic change in the human shape as discussed in Section II.2 is accepted as indicating human falling. Tables 4 and 5 show experimental results for the human skeleton match and the posture analysis respectively. The experimental results for shape analysis, utilizing the change ratio of two ellipse features, are tabulated in Table 6. From

Table 7 The experimental results of the proposed fall detection with the skeleton distance = 0.056

Event	No. of videos	Detected falling	Detected non-falling	TP	TN	FP	FN
Falling	22	20	2	20	0	0	2
Sit-down	8	2	6	0	6	2	0
Squat	8	0	8	0	8	0	0
Walking	8	0	8	0	8	0	0
Running	8	0	8	0	8	0	0

Table 8 The comparisons of four fall detection schemes

Fall detection approach	Detection rate	False alarm rate
The skeleton match	0.727	0.031
The posture analysis	1	0.34
The shape analysis	0.909	0.16
The proposed scheme	0.909	0.06

Table 9 Comparison of the proposed scheme and the shape analysis in terms of the execution time, detection rate and false alarm rate

Performance metric	The proposed detection 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

Tables 4 to 6, we can observe that the fall detection systems based on a single approach yield a high false positive rate and low detection accuracy because they cannot differentiate a sit-down from a fall-down.

The experimental results for the proposed falling detection system are shown in Table 7. Two fall-down incidents were not detected because they were slow-speed fall incidents, which did not register as fast posture changes in the first step of the proposed approach. On the other hand, two sit-down activities were flagged fall incidents because fast sit-down activities trigger a posture change step in the proposed approach. Table 8 summarizes the performance of the four human fall detection approaches in terms of the detection rate and the false alarm rate. These results demonstrate that an intelligent combination of different fall detection approaches can provide reliable fall detection. Table 9 compares the proposed hybrid human detection approach with the shape analysis approach (Rougier *et al*, 2007) in terms of the execution time, the detection rate and the false alarm rate. We can observe from Tables 8 and 9 that our approach can achieve high detection accuracy and a lower false alarm rate than other systems within a reasonable execution time.

IV. CONCLUSIONS

Since the global population is aging rapidly, fall detection for aging people has become an important issue in smart homes. The major contribution of this paper is to propose a novel real-time fall detection approach for elderly people, which is an intelligent combination of the skeleton change and the shape change detection scheme while still satisfying the real-time constraint. A human skeleton is first extracted from a human posture. The distance between two sampling skeletons beyond a threshold flags a posture change. We then use 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 human shape change. Finally, we confirm a human fall incident by monitoring the inactivity of a person for a period of time. Experimental results indicate that the proposed hybrid human fall detection system can achieve a high detection rate and a low false alarm rate with reasonable computational costs.

NOMENCLATURE

$\mathrm{Dist}(S_1, S_2)$	the distance of the two skeletons, S_1, and S_2
DM_{S_1}	the distance map of the binary image S_1
I_a	the major semi-axes of the ellipse
I_b	the minor semi-axes of the ellipse
I_{max}	the maximum eigenvalue.
I_{min}	the minimum eigenvalue

R_θ	the change rate in the orientation of the ellipse
R_γ	the change rate in the ratio of major and minor ellipse semi-axes
$V(x)$	a vertical projection histogram
θ	the orientation of the ellipse
γ	the ratio of the major and the minor semi-axes of the ellipse

REFERENCE

Bobick, A., and Davis, J., 2001, "The Recognition of Human Movement Using Temporal Template," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 23, No. 3, pp. 257-267.

Borgefors Gunilla, 1986, "Distance Transformations in Digital Images," *Computer Vision, Graphics, and Image Processing*, Vol. 34, No. 3, pp. 344-371.

Bradski, G. R., and Kaeher, A., 2008, *Learning OpenCV: Computer Vision with OpenCV Library*, O'Reilly Media, Inc., Sebastopol, CA, USA.

Chew, L. P., 1987, "Constrained Delaunay Triangulations," *ACM Proceedings of Third Annual Symposium on Computational Geometry*, Waterloo, Ontario, Canada, pp. 215-222.

Douglas David, and Peucker Thomas, 1973, "Algorithms for The Reduction of The Number of Points Required to Represent A Digitized Line or Its Caricature," *The Canadian Cartographe*r, Vol.10, No. 2, pp. 112-122.

Hs, Y. T., Liao, H. Y. M., Chen, C. C., and Hsieh, J. W., 2008, "Video-based Human Movement Analysis and Its Application to Surveillance Systems," *IEEE Transactions on Multimedia*, Vol. 10, No. 3, pp. 372-392.

Huang, W., Gu, I. Y. H., and Li, Q., Tian, L., 2003, "Foreground Object Detection from Videos Containing Complex Background," *Proceedings of the Eleventh ACM International Conference on Multimedia*, Berkeley, CA, USA, pp. 2-10.

Lin, Chia-Wen, Ling, Zhi-Hong, Cheng-Yeng, and Kuo, Chung J., 2007, "Compressed-domain Fall Incident Detection for Intelligent Homecare," *Journal of VLSI Signal Processing System*, Vol. 49, No. 3, pp. 393-408.

Lindemann, U., Hock, A., and Stuber, M., 2005, "Evaluation of A Fall Detector Based on Accelerometers: A Pilot Study," *Medical and Biological Engineering and Computing*, Vol. 43, No. 5, pp. 548-551.

Pratt Willam, 2007, *Digital Image Processing*, 4th ed., John Wiley and Sons, Ltd., NY, USA.

Rougier Caroline, Meunier Jean, St-Arnaud Alain, and Rousseau Jacqueline, 2007, "Fall Detection from Human Shape and Motion History Using Video Surveillance," *Proceedings of the 21st International Conference on Advanced Information Networking and Applications Workshops*, Niagara Falls, Canada, pp. 875-880.

Rougier Caroline, Meunier Jean, St-Arnaud Alain, and Rousseau Jacqueline, 2008, "Procrustes Shape Analysis for Fall Detection," *The Eighth International Workshop on Visual Surveillance*, Marseille, France, pp. 50-62.

Töreyin, B. U., Dedeoglu, Y., and Cetin, A. E., 2005, "HMM-based Falling Person Detection Using Both Audio and Video," *Lecture Notes in Computer Science*, Vol. 3766, Springer-Verlog, Berlin, pp. 211-220.

Vishwakarma Vinay, Mandal Chittaranjan, and Sural Shamik, 2007, "Automatic Detection of Human Fall in Video," *International Conference on Pattern Recognition and Machine Intelligent , Lecture Notes in Computer Science*, Kolkata, India, pp. 612-613.

William, A., Ganesan, D., and Hanson, A., 2007, "Aging in Place: Fall Detection and Localization in Distributed Smart Camera Network," *Proceedings of the 15th International Conference on Multimedia*, Augsburg, Germany, pp. 892-901.

Zhang, T., Wang, J., and Xu, L., 2006, "Using Wearable Sensor and NMF Algorithm to Realize Ambulatory Fall Detection," *International Conference on Advances in Natural Computation, Lecture Notes in Computer Science*, Xi'an, China, pp. 488-491.

Manuscript Received: Dec. 10, 2009
Revision Received: Mar. 10, 2010
and Accepted: Apr. 10, 2010