A Method of Detecting Human Body Falling Action in a Complex Background

Jia Dongyao *, Viocu Groza **, Liu XiaoHui*, Liu Xu*, Zhu Huaihua*,Qingsheng Zeng**
* Advanced Control Systems Laboratory, School of Electronics and Information Engineering, Beijing Jiaotong University, Beijing, China
** School of Electrical Engineering and Computer Science (EECS), University of Ottawa, Ottawa, Ontario K1N 6N5, Canada
dongyaojia1974@gmail.com, vgroza@uottawa.ca, truelove1974@126.com, 15125103@bjtu.edu.cn, 14120312@bjtu.edu.cn, qingshengzeng03@163.com

***Abstract*—Humans have various complex postures and movements. Considerable attention is given to the problem of recognizing a human fall. However, the recognition rates must be further improved, for practical applications, from that obtained in the previous research. In this paper, a new recognition method, based on the analysis of a human fall, is provided. Furthermore, five eigenvectors that describe a fall are defined i.e. the aspect ratio, effective area ratio, human point margin, body axis angle, and centrifugal rate of the body contour. Then, a support vector machine based on the Gauss radial basis function is trained to obtain a better identification result. The simulation results show that the model, though the combination of the five eigenvectors, has a recognition rate of 94.5%, which is a significant improvement as compared to the previous research.**

I. INTRODUCTION

Due to the aging population, the physical and mental health of the elderly has gradually drawn our attention. Due to their weak balance ability and adaptability, they can easily fall down. If the elderly don't get timely treatment after they fall, they may suffer further disability or even death, which has additional effects on their mental and physical health [1]. Thus, it is necessary to establish a method by which one can analyze an emergency situation quickly in order to provide correct medical treatment for aging citizens. Therefore, a key aspect in the supervision of the elderly is to have a method that can detect a fall. The human behaviors should be analyzed according to the relationship between the characteristics of human behavior. There is a difference between the characteristics of the body when falling down and other usual behaviors. The behavioral characteristics of the former are similar, and vary among different people and environments. Currently, traditional intelligent identification methods must be improved, in both their ability to analyze similar behavioral characteristics and to reduce error in their identifications. Some international journals, such as IJCV, CVIU, and IVC, and important academic conferences, such as CVPR, ICCV, and ICPR, have topic areas devoted to research on intelligent video surveillance technology, especially research using visual analysis on human movement. In addition, IEEE organizes symposiums on human movement analysis regularly. The real-time visual control system W4, developed by Haritaoglu, et al. was not only able to track the position and segmentation of a human subject, but could also track several people at once. In addition, this system could be used to detect and track people in the outdoors. Yu Yingzhuo is proposed a 3D reconstruction method, which he used to recognize squats and falling actions. However, there is no experimental study on continuous video [2]. Xu Liangwu used a machine learning method to identify lateral falling actions, but didn't specifically study behaviors that are easily confusing [3]. A 3-D model of the human head was established by Rougier, et al. that analyzed and tracked the motion of the head to detect whether there is a fall [4]. Hsieh, et al. presented a system that marked a model of the body with its main joint sand could be used to determine posture type through a corresponding matching algorithm. However, none of these studies made a distinction between the falling actions of humans, animals, and objects [5].

In response to the above problems, a method using multiple characteristics is proposed, which can detect a falling action quickly and accurately. From this, the newly defined eigenvector model is made, which contains a total of five categories and eleven eigenvectors. We define five categories, the aspect ratio, the effective area ratio, the human point margin, body axis angle, and centrifugal rate of a body contour. The results indicate that a support vector machine (SVM), trained by the eigenvector model, improves accuracy compared to previous models.

II. SELECTION AND DEFINITION OF FALL CHARACTERISTICS

A. The Human Aspect Ratio

We define the minimum enclosing rectangle and the aspect ratio of the human body (r_1) in this paper. The former is defined as the minimum rectangle that can completely cover the human body. The latter (r_1) is defined as the ratio of width to height of the human body contour within the minimum external rectangle in the image. The height will be longer than the width when standing, thus the aspect ratio is far less than one. The aspect ratio is only longer than one when falling. In most cases, we can determine if the subject has upright posture or not, through the aspect ratio. However, for movements such as squats, push-ups, sit-ups, and other similar fitness movements, it is less effective.

978-1-5090-1216-9/16/$31.00 ©2016 IEEE

$$r_1 = \frac{w}{h}. \quad (1)$$

B. The Effective Area Ratio of Human Body

In this paper, the effective area ratio of human body is defined as the ratio of the contour area in the minimum external rectangle [6] to the entire area of the rectangle:

$$r_2 = \frac{S_p}{S_r}. \quad (2)$$

In the formula above, r_2 is the effective area ratio. S_p is the number of pixels with the value 1 in the minimum external rectangle of the image after binarization. S_r is the total number of pixels in the minimum external rectangle. When the body covers most of the rectangle, the method is considered to be effective. When the body falls down, the aspect ratio is far greater than one. However, this is just one of the detection factors because there are many similar actions that could easily cause faulty judgment, including squats, sit-ups and push-ups, etc. When a person falls down, the area of the minimum external rectangle will become larger, and the effective area of the subject will be relatively smaller, the effective area ratio will also be smaller. Therefore, the effective area ratio can be used as an indicator of a fall.

C. The Point-edge Distance of Human Body

As shown in Fig. 1(c), the kite model is constructed by connecting the centroid [7] of the subject figure with four points that are farthest from the centroid to the top, bottom, left and right. The body point-edge distances are defined as the four lines between the centroid and the points mentioned above.

$$d_i = \sqrt{(x_i - x_m)^2 + (y_i - y_m)^2}, \quad i = 1,2,3,4. \quad (3)$$

The eigenvector is obtained: $D = [d_1, d_2, d_3, d_4]$

After dividing the body contour, the point-edge distance of the subject contains information on the status of the subject's body posture, therefore it is selected as one of the characteristics in this paper, as shown in Fig. 3.

D. The Axis Angle of Human Body

This characteristic is defined as four angles obtained from the axis and the adjacent axes by connecting the centroid with the four points of the kite model. As shown in Fig. 1(d), the eigenvector of the four-star angle is calculated by (4) and $\theta = [\theta_1, \theta_2, \theta_3, \theta_4]$

$$\theta_i = \cos^{-1}\frac{(x_i - x_m)(x_{i-1} - x_c) + (y_i - y_c)(y_{i-1} - y_c)}{d_i \bullet d_{i-1}}, i = 1,2,3,4. \quad (4)$$

With different body postures, the axis angles are different. Because the axes angles contain information that can be used to distinguish falling actions, they are also used as a one of the characteristics in our system.

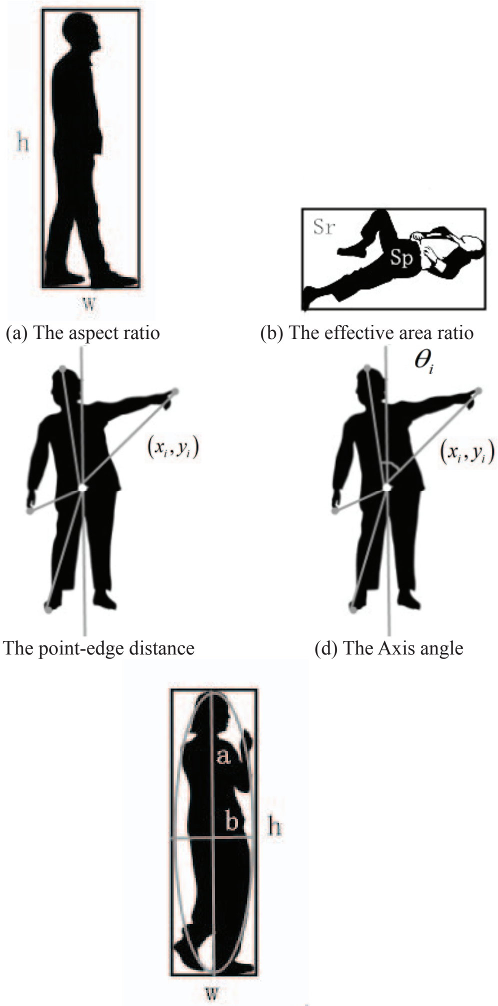
E. The Centrifugal Rate of Human Body

The centrifugal rate of human body is defined as the centrifugal rate of the minimum inscribed ellipse of the minimum external rectangle of body contour. From this measure we can gain more information on a falling action. Therefore, the centrifugal rate of the body contour is used as a shape characteristic, as shown in Fig. 1(e), calculated by (5).

$$e = \sqrt{1 - \frac{b^2}{a^2}}. \quad (5)$$

In this way, the eigenvector with fused characteristics is obtained: $T = [r_1, r_2, d_1, \cdots, d_4, \theta_1, \cdots, \theta_4, e]$

As shown in Fig.1, the five characteristics of the falling action above describe different aspects of the human body form and motion, thus integration of each characteristic gives a better description of a falling action.



(a) The aspect ratio (b) The effective area ratio

(c) The point-edge distance (d) The Axis angle

(e) The centrifugal rate

Fig.1. Five eigenvectors

After the characteristics of the contour are extracted, a support vector machine (SVM) is chosen to classify and build a model defining falling and non-falling behaviors. A kernel function is used to convert the nonlinear feature space to a linear separable space through the SVM. After the transformation, a hyperplane is constructed in the linear feature space to obtain the optimal classification model for distinguishing features so that better robustness is obtained even in the case of a small number of samples [8-9].

III. Falling Detection Based on Radial Basis Function Support Vector Machine

After extracting the characteristics by which we are defining specific types of human behavior, the SVM is chosen to model and classify falling actions. The SVM is used to transform the nonlinear feature space into a linear space. After the transformation, a hyperplane is constructed in the linear feature space to obtain an optimal classification model that makes clear distinctions [10] between different classifications, in order to improve robustness even in the case of a small number of samples.

Fig. 2 shows the mapping of the radial basis function (RBF). It can be seen that the inseparable line in low dimensional space turns into a separable line in high dimensional space.

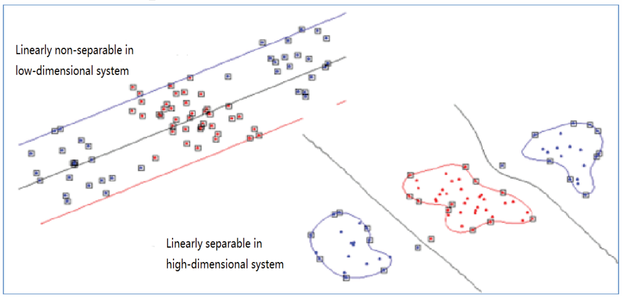


Fig.2. The mapping of RBF

For the RBF, the feature space is an infinitely dimensional Hilbert space [11]. Even the largest interval optimization problem can be solved in an infinite dimension because of its regularity. In order to ensure the accuracy and efficiency of the classifier based on multi-characteristics, an SVM classifier based on the RBF is studied in order to meet the requirements necessary to detect a fall.

Observed human behaviors are divided into two categories for the purpose of this paper: falling and non-falling behaviors. Non-fall behaviors, including running, jumping, squatting, bending, and sitting down, are not considered. According to this paper, the eigenvector includes11 parameters and is defined as $T=[r_1,r_2,d_1,\cdots,d_4,\theta_1,\cdots,\theta_4,e]$. The SVM model process is shown as follows:

Supposing that the training sample is $S=\{S_1,S_2,\cdots,S_k\}$, where S_i is the *i*-th sample.

(1) The training sample, the eigenvector $T=[r_1,r_2,d_1,\cdots,d_4,\theta_1,\cdots,\theta_4,e]$ is used as the learning parameter. The RBF is chosen to transform the nonlinear feature space into the linear space, which is shown as:

$$k(x_i,x_j)=\exp\left(-\frac{\|x_i-x_j\|}{2\sigma^2}\right). \tag{5}$$

(2) Then the RBF is used in Lagrange's dual function.

$$L(w,b,\alpha)=\sum_{i=1}^{N}\alpha_i-\frac{1}{2}\sum_{i,j=1}^{N}\alpha_i\alpha_j\omega_i\omega_j k(x_i,x_j) \tag{6}$$

After a series of mathematical transformation, the discriminant function is

$$f(x)=\sum_{i\in SV}\alpha_i\omega_i k(x_i,x)+b \tag{7}$$

Where α_i must be maximized.

(3) Obtain x_i, which satisfies the requirements of the discriminant function with the maximum α_i. Then save the nonzero support vector α_i and corresponding training vector x_i.

(4)For the classification of mode x, support vector x_i and corresponding weight α_i are used to evaluate the function:

$$f(x)=\sum_{i\in SV}\alpha_i\omega_i k(x_i,x)+b. \tag{8}$$

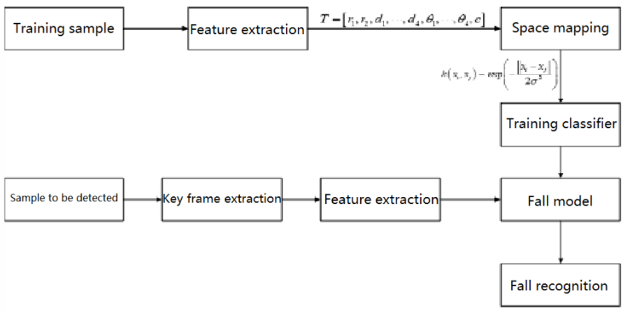


Fig.3. The flow chart of falling detection based on multi feature fusion

The process used to detect human body falling action in a single image is as follows. We first selected 1454 static pictures for training. From these, we extracted the eigenvectors, which included the human aspect ratio as well as the effective area ratio, the point-edge distance, the axis angle and the centrifugal rate of the pictured human. We then mapped the eigenvectors into a higher dimensional space using RBF to separate the characteristics. After the training was completed, the classifier was obtained, which defines the model used to detect falling actions. Finally, we extracted the eigenvectors of the pre-processed samples, and the eigenvectors were input into the model to detect falling actions in test images.

IV. Experimental Results and Analysis

In this paper, the process for detecting a falling action is divided into two steps. The first step is classifier training and the second step is the detection of falling in test images. A single static image or a group of several consecutive images are used in the experiments. A total of 469 positive samples (fall images) and 985 negative samples (non-fall images) were obtained locally and from the Internet and were used as training images. Subsequently, 100 fall images and 100 non-fall images were selected as test images.

Using all eigenvector data, the classifier detecting a falling action was trained according to the method outlined above. After training, 1000 images were selected as test samples, with a 2:1 ratio between positive and negative samples. Among the samples, 735 images were selected from local video images, and the remaining 265 images are from the Internet. Some experimental results are shown as flows: The images which are right detected are shown in Fig.4. The false images of falling detection are shown in Fig.5. The images which are not detected are shown in Fig.6.

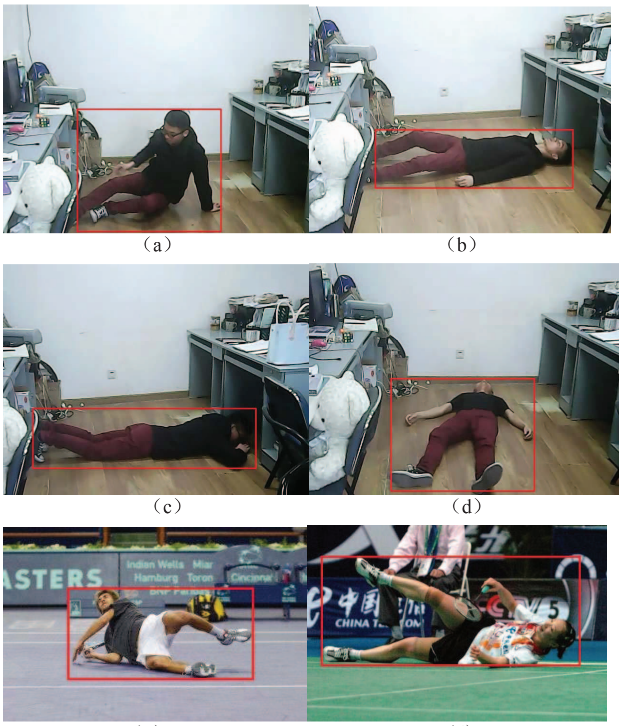


Fig.4. Part of the image that falling detection right

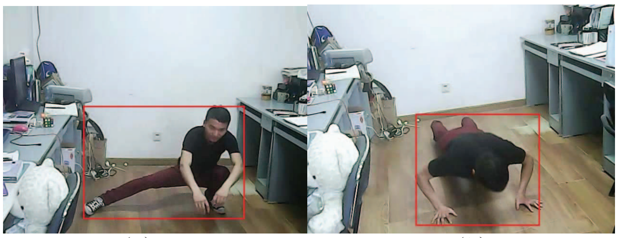


Fig.5. The false images of falling detection

Fig.6. The images which are not detected

From the experimental results above, we can see that the detection of usual falling actions is effective, but some similar images are still not correctly detected, specifically when the human aspect ratio and the effective area ratio are similar to those of falling action. In addition, some external factors, such as significant illumination change and background, can interfere with detection results. Some false positives are images of body-building movements such as squats, sit-ups, and push-ups.

It can also be seen that some falling actions are not detected. This occurs is the beginning and middle of the falling action, but not at the end. Errors also occur when movements are interrupted through human interference, such as when one catches and braces themselves or is held up by another individual.

In this paper, 1000 images are selected as test samples and the trained classifier is used to detect falling action according to the proposed algorithm. The results are as follows:

TABLE I.
RECOGNITION ACCURACY OF THE METHOD IN THIS PAPER

The number of fall images identified	632
The number of non-fall images identified	368
The actual number of fall images	669
The actual number of non-fall images	331
The recognition rate of falls	(632/669)×100%=94.5%

Two methods of detecting a falling action are selected to compare with the method proposed in this paper. One of the methods is based on the Hopfield neural network and the other is template matching method. The same 1000 images are used as test samples to compare the performance of the three methods.

TABLE II.
RECOGNITION ACCURACY BASED ON HOPFIELD NEURAL NETWORK

The number of fall images identified	579
The number of non-fall images identified	421
The actual number of fall images	669
The actual number of non-fall images	331
The recognition rate of falls	(579/669) ×100%=86.5%

TABLE III.
RECOGNITION ACCURACY BASED ONTEMPLATE MATCH

The number of fall images identified	503
The number of non-fall images identified	497
The actual number of fall images	669
The actual number of non-fall images	331
The recognition rate of falls	(503/669) ×100%=75.2%

From the results above, we see that only 503 images are identified by the template matching method, with a recognition rate of 75.2%, the lowest of the three methods. The recognition rate of Hopfield neural network method is 86.5% and 579 images are identified. The recognition rate of the method proposed in this paper is 94.5%, and 632 images are identified.

CONCLUSION

In this paper, we propose a method, based on Bayesian fused multi-characteristics, for the detection of a fall. We define five types of eigenvectors, which can describe the characteristics of a fall. SVM model, based on RBF, is used to detect a falling action. In our experiment, the data obtained from videos and public databases are used to train the classification model. The experimental results show that our model is more accurate in detecting falls than the previous methods.

REFERENCES

[1] Wang Qitian. The quantitative analysis of the development trend of China's population aging [D]. Jilin University, 2011

[2] Yu Yingzhuo, based on the characteristics of human motion detection and gesture recognition [D]. Xi'an: Xi'an Electronic and Science University, 201

[3] Xu Liangwu, based on the SVM's indoor anomaly detection algorithm research [Dl. Nanjing: Nanjing University of Science and Technology, 2008

[4] Zhou Hong, Zu Yuanyuan, Chen Xiao. Standard treatment for 3D head face model of magnetic resonance imaging [J], 2011, 32(3):57-61. DOI:10.3969/j.issn.1003-0158.2011.03.011.

[5] Zhang Chunhong, Lin. Analysis of the bionic technology in the human model of the [J]. machinery 2009,01:1-3+73.

[6] Sidi O,Van Kaick O,Kleiman Y,et al.Unsupervised Co-Segmentation of a Set of Shapes via Descriptor-Space Spectral Clustering.ACM Transactions on Graphics (Proceedings of ACM SIGGRAPH Asia),2011,30(6):126-134

[7] Hu R Z,Fan L B,Liu L G.Co-segmentation of 3D Shapes via Subspace Clustering.Computer Graphics Forum(Proceedings of SGP).2012,31(5):1703-1713

[8] Wuhrer S,Azouz Z B,Shu C.Semi-Automatic Prediction of Landmarks on Human Models in Varying Poses//Proc of the Canadian Conference on Computer and Robot Vision.Ottawa,Canada,2010:136-142

[9] Peng-fei C, Hao-wen Y, Zhen-hui H. An Algorithm for Computing the Minimum Area Bounding Rectangle of an Arbitrary Polygon[J]. Journal of Engineering Graphics, 2008, 29(1):122-126.

[10] Yi-fang F, Liang-chou Z, Zhi-run Y. Research on the Locating of Image Centroid of Human Movements--Locating the Human Centroid[J]. Journal of Guangzhou Physical Education Institute, 2004.

[11] Shalev-Shwartz S, Singer Y, Srebro N, et al. Pegasos: Primal Estimated sub-gradient solver for SVM [J]. Mathematical Programming, 2011, 127(1):3-30.

[12] Duan K, SathiyaKeerthi S. Which is the best multiclass SVM method? An empirical study[C]. Proceedings of the Sixth International Workshop on Multiple Classifier Systems. 2005:278--285.

[13] Zhang B, Chen X, Gao W. Discriminant Analysis Based on Kernelized Decision Boundary for Face Recognition [J]. Lecture Notes in Computer Science, 2005.

[14] Er M J, Wu S, Lu J, et al. Face recognition with radial basis function (RBF) neural networks [J]. Neural Networks IEEE Transactions on, 2002, 13(3):697 - 710.

[15] Liang Luhong, Ai Haizhou, Xiao Xipan et al. Face detection based on template matching and support vector machine [J], 2002, 25(1):22-29. DOI:10.3321/j.issn:0254-4164.2002.01.004.

[16] Wu Xiaojun, Zou Guanghua. High performance template matching algorithm based on edge geometry feature [J], 2013, 34(7):1462-1469. DOI:10.3969/j.issn.0254-3087.2013.07.004.

[17] Yong Hou, A Novel SVM Algorithm and Experiment [J].International Conference on Computer Science and Electronics Engineering, 2012, Vol.2:31-35.

[18] Chin-Wei Hsu, Chih-Chung Chang, Chih-Jen Lin. A Practical Guide to Support Vector Classification[EB\OL].April 15, 2010.

[19] Qidan Zhu, Exploration and improvement of Ostu threshold segmentation algorithm [J].Congress on Intelligent Control and Automation, 2010: 6183 一 6188.

[20] LI Xue-hua， SHU Lan． Fuzzy theory based support vector machines classifier [C]//Proc of the 5th International Conference on Fuzzy Systems and Knowledge Discovery.2008: 600-604．

[21] POLAT K，GUNES S． A novel hybrid intelligent method based on C4.5 decision tree classifier and one-against-all approach for multiclass classification problems [J].Expert Systems with Application，2009，36(2):1587-1592．

[22] MANIKANDAN J， VENKATARAMANI B． Study and evaluation of a multi-class SVM classifier using diminishing learning technique [J].Neurocomputing， 2010， 73(10-12) : 1676-1685．

[23] WIDODO A，YANG B S． Support vector machine in machine condition monitoring and fault diagnosis [J]． Mechanical Systems and Signal Processing，2007，21(6) : 2560-2574．

[24] Hao Peiyi,Chiang J H.Fuzzy regression analysis by support vector learning approach [J].IEEE Trans on Fuzzy Systems,2008,16(2):428-441.

[25] GONZALEZ—ABRIL L，VELASCO F，ANGULO C，et a1． A study on output normalization in muhiclass SVMs [J]． Pattern Recognition Letters，2013，34(3)：344-348．