Fall Detection Algorithm Based on Human Posture Recognition

Kewei Zhao[1,2,3]*, Kebin Jia[1,2,3]* and Pengyu Liu[1,2,3]*

1 Beijing Laboratory of Advanced Information Networks, Beijing, China
2 College of Electronic Information and Control Engineering, Beijing University of Technology, Beijing, China
kebinj@bjut.edu.cn

Abstract. In the background of global aging, more attention should be paid to the elders' health and the equality of their life. Nowadays falls became one of greatest danger for old people. Almost 62% of injury-related hospitalizations for the old are the result of it. In this paper, we propose a new method to detect fall based on judging human's moving posture from the video. It consists of three main parts, detecting the moving object, extracting the feature and recognizing the pattern of behavior. To improve the precision and increase the speed of the detection, we adopt two layers codebook background modeling and codebook fragmentation training. Two level SVM method to recognize the behavior: In the first level of the SVM classifier, we distinguish the standing posture and other posture by the feature of moving object, such as the ratio of the major and minor axis of the ellipse. In the second level of the SVM classifier, angle of the ellipse and head moving trajectory to judge the falls and squat. The experimental results indicate that our system can detect fall effectively.

Key words: fall detection, two layers codebook background modeling, codebook fragmentation training, two level SVM classifier

1 Introduction

Based on the statistic census from the National Bureau of statistic in 2015 shows that the population of people over 60 in china is 221.82 million, 16.15 of the total population, and the population over 65 years old is 143.74 million, 10.47% of that. Compared with the sixth national census, the proportion have increased 2.89% and 1.60%, china is stepping into the population aging society. Facing the serious tendency of the population aging, government should establish a lot of now information system to ensure the safety of the old people. According to the statistic, more than 35% old people who is older than 65 years old have fall down before, 63% of the old people died by falls, it even reached 70% among the old people over 75 years old[1]. In conclusion, it is especially important to detect the falls for the protecting the safety of old people.

J.-S. Pan et al. (eds.), *Advances in Intelligent Information Hiding and Multimedia Signal Processing*, Smart Innovation, Systems and Technologies 64,
DOI 10.1007/978-3-319-50212-0_15

There are lots of techniques to detect fall. They can divided roughly into three categories by different kinds channel of signal[2]: one of them is based on some wearable devices, the problem of such detectors is that old people often forget to wear or charge devices; another one is based on the environment devices, setting up the equipment is so complex and expensive; the third one is based on the video, it can not only perform a 24-hour monitoring, but also can avoid the danger cause by forgetting to wear. The paper proposed a method which is based on judge the posture of the old people, it is totally implemented by the image processing techniques, it can detect the posture of fall effectively, and also satisfy the require of real-time processing on the hardware with low computing capability. The basic principle show below: it recognize the moving object by the modified codebook, Then, human characteristic matrices are constructed based on the information of human body posture extracted from human silhouette and are used as features to train SVM classifier for fall detection.

2 Related Work

Some method to detect fall have been purposed in recent research. Accelerate sensors can collect values and direction information, it is widely use in the fall detection systems based on wearable sensor. Dai et al. get the accelerator values and latitude to detect fall with the accelerator sensor in the mobile phone. After fall, it can directly alarm the family, it is easy to carry and alarm rapidly. The false-negative is 2.67% and false-positive is 8.7%[3]. Bourke et al. designed a fall detection system based on two-link gyroscope sensor. It can be bind on the chest and detects fall by the angular acceleration and velocity. The system has 100% sensitivity、specificity and accuracy[4].

Scott et al. designed a matters fall detection system, it consists of lots of airmattress and pressure sensors, by the changes of pressure on the airmattress guesses the pressure changes of human. In this way, we can get the posture changes of the human. It is used for detection while people is sleeping or scribing, it is comfortable and convenient[5]. Litvak et al. designed a detection system based on the analyzing the shake of the ground and the sound. It combines the accelerator sensor and microphone, it doesn't need to wear anything and affects people move less. It's sensitivity is 97.5% and specificity is 98.6%[6].

In the video frame fall detection system, Jean Menunier et al. proposed an advance detection system based on MHI[7]. They extracted the human in the foreground by cutting the background, it computes the ratio of the major and minor axis、obliquity and MHI statistic of the ellipse to detect the fall. The camera always be set at the roof to have a wide view avoiding occlusion.

3 Technical Details

In this paper, we propose a method to detect the fall of the old people. It can divided into three part: moving object detection, feature extraction and behavior recognization(Fig.1). We use a modified codebook foreground detection algorithm in making object detection, compared with the old algorithm, it improves the precision of the foreground extraction and makes the extraction more real-time. We use two level SVM classifier to recognize the fall and squat in object behavior recognization.

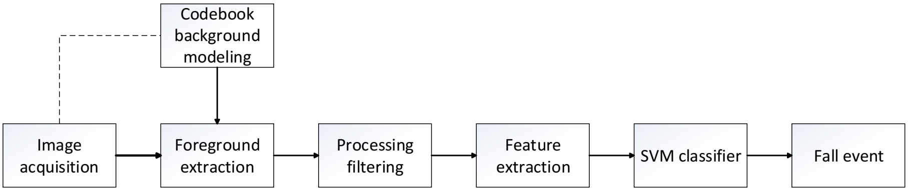


Fig.1. Flow diagram of the overall fall detection procedure.

3.1 Foreground extraction

Codebook algorithm is to establish the background model in the long-term observation sequence with quantization technique. The algorithm will make every pixel a codebook model. During the initial algorithm training, $c = \{x_1, x_2, \cdots, x_n\}$ represent the training sequence of one pixel, which is made up of RGB vectors, $\xi=\{c_1,c_2,...,c_L\}$ represents a pixel codebook which consists of L codeword, It has different numbers of code word depending on the variation of the pixel. Each codeword $c_i (i = 1 \ldots . l)$ consist of a RGB vector $v_i = (\overline{R}_i, \overline{G}_i, \overline{B}_i)$ and $aux_i = \langle \check{I}_i, \hat{I}_i, f_i, \lambda_i, p_i, q_i \rangle$.

We match the currently codeword c_m with each instance x_t, then we use the matching codeword as the estimation of the instance. We generally using colordist and brightness to evaluate whether a codeword is in a best matching status.

Foreground-background segmentation

（1）Let $L \leftarrow 0$； $Z \leftarrow \emptyset$

（2）When t = 1....N

1. $x_t = (R, G, B), I \leftarrow \sqrt{R^2 + G^2 + B^2}$
2. If a. $colordist(x_t, v_m) \le \varepsilon_1$

   b. $brightness(I, \langle \check{I}_m, \hat{I}_m \rangle) = true$

   Searching the codeword c_i in $\xi=\{c_1,c_2,...,c_L\}$to match x_t.
3. If $Z \leftarrow \emptyset$ or the no matching codeword having been found, then $L \leftarrow L + 1$， and a new codeword c_L will be greated：

   a. $v_L \leftarrow (R, G, B)$

b. $\mathrm{aux}_L \leftarrow \langle I, I, 1, t-1, t, t \rangle$

4. otherwise refresh the matching codeword c_i, including $v_m = (\bar{R}_m, \bar{G}_m, \bar{B}_m)$ and $\mathrm{aux}_m = \langle \check{I}_m, \hat{I}_m, f_m, \lambda_m, p_m, q_m \rangle$:

a. $v_m \leftarrow \left(\frac{f_m R_m + R}{f_m + 1}, \frac{f_m G_m + G}{f_m + 1}, \frac{f_m B_m + B}{f_m + 1}\right)$

b. $\mathrm{aux}_m \leftarrow \langle \min\{I, \check{I}_m\}, \max\{I, \check{I}_m\}, f_m + 1, \max\{\lambda_m, t - q_m\}, p_m, t \rangle$

(3) We compute the maximum time interval of each codeword $c_i (i = 1 \ldots . l)$ between with every pixel having been matched again

$\lambda_i \leftarrow \max\{\lambda_i, (N - q_i + p_i - 1)\}$

For pixel $x_t = (R, G, B)$ and a codeword c_i, with $v_i = (\bar{R}_i, \bar{G}_i, \bar{B}_i)$, existing:

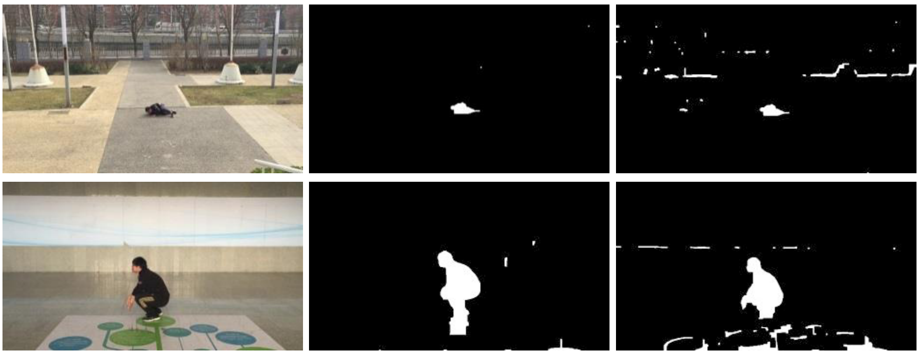
$$\|x_t\|^2 = R^2 + G^2 + B^2$$
$$\|v_i\|^2 = \bar{R}_i^2 + \bar{G}_i^2 + \bar{B}_i^2 \quad (1)$$
$$(x_t, v_t)^2 = (\bar{R}_i R + \bar{G}_i G + \bar{B}_i B)^2$$

$\mathrm{colordist}(x_t, v_i)$ can be calculated by the formula.

$$p^2 = \|x_t\|^2 \cos^2\theta = \frac{(x_t, v_i)^2}{\|v_i\|^2} \quad (2)$$

$$\mathrm{colordist}(x_t, v_i) = \delta = \sqrt{\|x_t\|^2 - p^2} \quad (3)$$

This paper modified the model built by codebook in following parts: we build a module consist 30*30 pixels, each module one or several codewords will be in the codebook. The modified codebook method can prevent failing of foreground extraction when camera shake and it can reduce memory pressure, foreground extraction show below Fig.2.



Original Image 30*30 pixels a pixels

Fig.2. foreground extraction result

The twice codebook matching mentioned in this paper, background subtraction library show below Fig.3. The codeword in codebook B come from codebook A When $\lambda_i < \lambda_0$ (λ_0 is threshold).

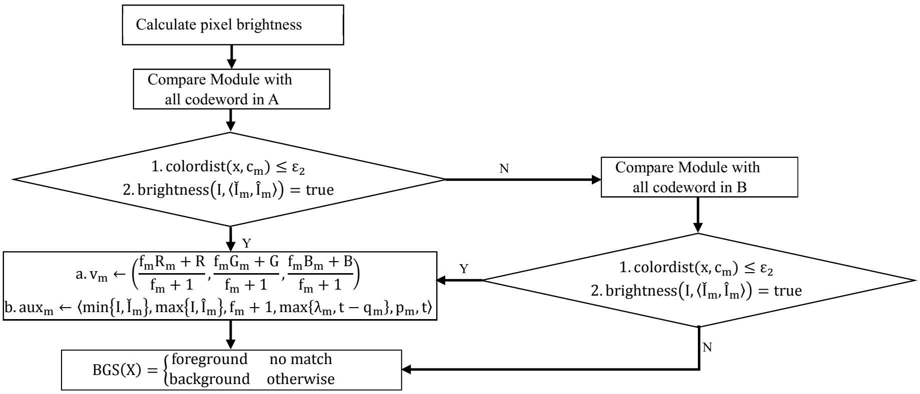


Fig.3. background subtraction library procedure

3.2 Feature Extraction And Posture Classification

Behavior recognition is a complex problem and have been researched for a long time. In paper[8], humans behavior has been modeled by two layer HMM(hidden markov model), while in paper[9],it use RBF(radial basis neural network) to model the behavior. We method concentrate on fall, so we use SVM, it is a good binary classifier. We designed a two level SVM to judge the old people's posture. The first level classifier judge whether the old man is standing, when the result shows the old man is not standing, we put it into the second level classifier. In this level, we can judge whether the old man fall, when we find the old man fall, the system will alarm, the two level SVM use RBF kernel and confirm the optimal parameters g=0.03125, penalty factor c=4 by the method of gridding search.

The first level classifier need C_{motion}, which represent the intense of the MHI(motion history image), show below Fig.4, the ratio of the major and minor axis of the ellipse, the height of the old man to train and judge.

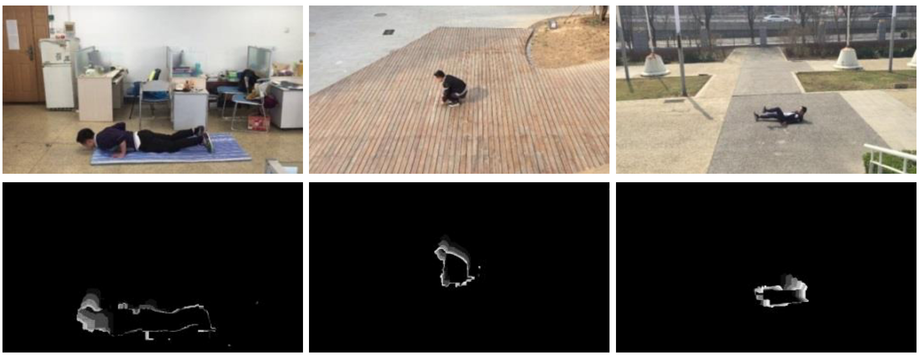


Fig.4. motion history image

1. the ratio of the major and minor axis of the ellipse

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

$$RateY2X = I_{max}/I_{min} \tag{5}$$

2. the height feature of the moving object

We defined HeightRate to represent the changing of the old people's height.

$$HeightRate = \frac{(Y_{max}-Y_{min})}{ManHeight} \tag{6}$$

The ManHeight is the old people's height. It is computed by the average formula

$$ManHeight = \frac{Height_1+Height_2+\cdots+Height_n}{n} \tag{7}$$

In the second level classifier. We process the frames that was judged not standing by first level classifier. Because squat and fall have similar features, the second level classifier mainly distinguish squat and fall. The second level classifier uses four features: the ratio of the major and minor axis of the ellipse, the project area of the moving object, the inclination angle of ellipse, the variation tendency of the major and minor axis.

1. the project area of the moving object

The projection area represent the sum of non-zero moving pixels in the moving object after binaryzation. When squatting, the old man will carried up his body and the project area will be small, when fall, the body will stretch, the area will be larger. So we can use it to distinguish the squat and fall.

2. the inclination angle of ellipse

The changing of the inclination angle can distinguish squat and fall well, when the old people squats, the angle will not change a lot, but when the old man fall, the body change will vertical to horizontal, the angle will change a lot. Formula is as follow:

$$\theta = \begin{cases} \tan^{-1}\left\{\frac{e_{21}}{e_{11}}\right\} & \text{if } \lambda_M = \lambda_1 \\ \tan^{-1}\left\{\frac{e_{22}}{e_{12}}\right\} & \text{if } \lambda_N = \lambda_2 \end{cases} \tag{8}$$

4 Experimental Results

To evaluate the accuracy of the algorithm, we invite 10 college students to do experiment, we choose three outside scene and three inside scene. The postures include: falling but not lying, falling and lying, squat and fall to the left/right(Fig.5). we designed some postures in table1. Experimenters will act some postures with fall to accomplish the experiments. Video data will be collect by HD camera. Each experimenter will do 20 experiments, all the experimenters will accomplish 200 experiments.

The statistic results shows in table1. A_L refers to whether alarm. N_A is number of actions. N_C is number of correctly detected events. N_F is number of falsely detected events. N_T is number of no alarm. R is the recognition rate. Falling down1 is falling down forward but not lying. Falling down2 is falling down backward but not lying. Falling down3 is falling down forward and lying. Falling down4 is falling down backward and lying. Falling down5 is falling on the left/right side. From the results, we know that our method have high validity, it can get most of falls, but it has misdescription when the experimenter fall but not lie and keep balance after fall.

Table 1 RECOGNITION RATE FOR VARIOUS EVENTS

Events	A_L	N_A	N_C	N_F	N_T	R
Stand	N	50	0	0	50	100
Walk	N	50	0	2	48	96.0
Run	N	30	0	1	29	96.7
Falling down1	Y	60	55	4	1	91.7
Falling down2	Y	60	57	1	2	95.0
Falling down3	Y	60	54	4	2	90.0
Falling down4	Y	60	54	3	3	90.0
Falling down5	Y	60	53	5	2	88.3
squat	Y	60	51	4	5	85.0
ALL		**490**	**451**	**24**	**15**	**92.1**

Inside1 squat | Inside2 before fall | Inside3 fall after

Outside1 before fall | Outside2 squat | Outside3 before fall

Fig.5. posture video data

5 Conclusions and Future Work

This paper introduce a fall detection method based on video, the method consists three parts: moving object detection, feature extraction, behavior recognization. Combining modified codebook algorithm and denoising method in OpenCV, we implement the moving object detection. We use two level SVM classifier to recognize the behavior, it is easy to implement. Through the designed well experiments, we get 90.27% accuracy. It shows that the method is practical in some degree and lag the foundation of applying in our daily life. But the algorithm is easily influenced by the complex background, for example, the system will have misdescription when detect lots of people; codebook algorithm can't distinguish the similar color well, this makes foreground extraction has more misdescription. So the method should combined with other method to raise the accuracy.

Acknowledgment

This paper is supported by the Project for the Key Project of Beijing Municipal Education Commission under Grant No. KZ201610005007, Beijing Postdoctoral Research Foundation under Grant No.2015ZZ-23, China Postdoctoral Research Foundation under Grant No.2015M580029, 2016T90022, the National Natural Science Foundation of China under Grant No.61672064 and Computational Intelligence and Intelligent System of Beijing Key Laboratory Research Foundation under Grant No.002000546615004.

References

1. Zhu L, Zhou P, Pan A, et al. A Survey of Fall Detection Algorithm for Elderly Health Monitoring[C]// IEEE Fifth International Conference on Big Data and Cloud Computing. IEEE, 2015:270-274.
2. Mubashir M, Shao L, Seed L. A survey on fall detection: Principles and approaches[J]. Neurocomputing, 2013, 100(2):144-152.
3. Dai J, Bai X, Yang Z, et al. PerFallD: A Pervasive Fall Detection System Using Mobile Phones[C]// Eigth IEEE International Conference on Pervasive Computing and Communications, PERCOM 2010, March 29 - April 2, 2010, Mannheim, Germany, Workshop Proceedings. 2010:292-297.
4. Bourke A K, Lyons G M. A threshold-based fall-detection algorithm using a bi -axial gyroscope sensor [J]. Medical Engineering and Physics, 2008, 30(1): 84-90.
5. Schuman Sr. R J, Collins W F. Indicator apparatus for healthcare communication system: EP, US8384526[P]. 2013.
6. Zigel Y, Litvak D, Gannot I. A method for automatic fall detection of elderly people using floor vibrations and sound--proof of concept on human mimicking doll falls.[J]. IEEE Transactions on Biomedical Engineering, 2010, 56(12):2858-2867.
7. Mubashir M, Shao L, Seed L. A survey on fall detection: Principles and approaches[J]. Neurocomputing, 2013, 100(2):144-152.
8. Awaida S M, Mahmoud S A. Automatic Check Digits Recognition for Arabic Using Multi-Scale Features, HMM and SVM Classifiers[J]. British Journal of Mathematics & Computer Science, 2014, 4(17):2521-2535.
9. Zhuo K Y. Recognition of automobile types based on improved RBF neural network[J]. Journal of Computer Applications, 2011.