Research on elevator passenger fall detection based on machine vision

Shaofeng Liu¹, Ziliang An²*, Ning Wang¹, Dingsong Bai¹, Xintong Yu¹

¹School of Railway Transportation, Shanghai Institute of Technology, 201418 Shanghai, China
²School of electromechanical engineering and information, Shanghai Urban Construction Vocational College, 201415 Shanghai, China
*email: anziliang@succ.edu.cn

Abstract. Aiming at the problem that the current elevator monitoring system cannot detect the accidental fall of passengers, this paper proposes a fall detection method based on machine vision and multi-feature fusion. First, moving targets were extracted by ViBe algorithm, and then the human body was marked with an external rectangle. Three characteristic parameters, namely the aspect ratio, effective area ratio and centroid acceleration of the human body, were calculated. At last, thresholds were set and SVM classification training was conducted to judge whether there was a fall event. Experimental results show that the algorithm has high accuracy and good stability. It can effectively reduce the injury caused by the elderly falling down in the elevator.

1. Introduction.
As an important vertical travel tool for people, elevator is closely related to people's travel safety. Therefore, it is of great social significance to develop an elevator passenger fall detection algorithm[1]. At present, the methods of elevator passenger fall detection are mainly divided into three kinds.

1) Fall detection based on wearable devices. Wearable sensor based fall detection methods mainly depend on sensory data gathered from wearable accelerometer and gyroscope, Based on the sensory data to determine whether there is a fall event. Wearable devices, however, introduce a wearability burden and are not very suitable for elevator passengers[2-3].

2) Environment-based fall detection: this method is to install all kinds of sensors in advance in the elevator, such as infrared sensor, pressure sensor, sound sensor. This kind of equipment is used to obtain the information of human body state and posture, so as to judge the fall[5]. The disadvantage is that elevator vibration and noise may affect the test results.

3) Fall detection based on machine vision: this method collects human body image information through the camera, analyzes and processes it with image processing technology, and finally determines whether there is a fall event. Compared with the first two types of detection methods, the detection method based on computer vision will not produce wearability burden on the elderly, and will not misdiagnose due to the influence of environmental noise[4]. Therefore, there are more studies on fall detection based on computer vision at present.

In this paper, a fall detection algorithm based on machine vision is proposed. the camera mounted on the elevator as the only image source, based on these data, the system extracts the moving target through the ViBe algorithm, and integrates the three features of aspect ratio, effective area ratio and
centroid acceleration of the human body to determine whether a fall event has occurred, which has good operability and economy.

2. Moving target detection.
ViBe algorithm is used to detect moving objects. It is an algorithm based on pixel value comparison and update. The algorithm uses neighborhood pixels to create the background model, and detects the foreground by comparing the background model with the current input pixel value. Which is mainly divided into three steps. First, the background model of each pixel in a single frame image is initialized. Each pixel model of Tahtis is represented by the pixel of its domain. Then, foreground target is extracted from the following image sequences by segmentation. Finally, each background model needs to be updated after use to remove the shadow and error caused by the subtle jitter of the captured video. After testing, the algorithm can effectively process video with a frame rate of 60, and can detect targets accurately and in real time.

3. Target feature extraction.
After the moving target is extracted, the image features of the target need to be extracted. The quality of feature extraction directly affects the accuracy of subsequent fall detection. In most cases, we use rectangle or ellipse to describe human behavior. In this paper, the object is circled by the way of enclosing rectangle, and on this basis, the characteristics of the object are analyzed. It is easy to generates false detection if we only chose one feature as detection standard, and Rather than makes comprehensive judgment by combining various features. Because of the complexity of passengers' behavior when taking the escalator. Therefore, this paper adopts the method of multi-feature fusion to analyze the features that change significantly and are easy to extract when elevator passengers fall, such as the aspect ratio feature of the outer rectangle, the acceleration feature of the center of mass, the effective area ratio feature, and carries out experiments for different behaviors.

3.1. Aspect ratio
The calculation method of the aspect ratio of the surrounding rectangle is defined as Formula (1), where H is the height of the surrounding rectangle and W is the width of the surrounding rectangle.

\[ \text{Rate} = \frac{H}{W} \]  

(1)

The body's outer rectangle changes dramatically during a fall. When a person is standing, his height must be bigger than the width of the minimum circumscribed rectangle, falls down, the width must be larger than the height, according to this characteristic, we can set threshold ratio is 1.00, when the minimum circumscribed rectangle ratio greater than 1.00, identified as normal events, when the minimum circumscribed rectangle ratio less than 1.00, identified as fall event. The following Figure 1 show the minimum outer rectangle experimental figure of human body foreground detection in different states, which basically covers the common behavior of elevator passengers when taking the elevator.

![Comparison of aspect ratio under different behaviors](image)

Figure 1. Comparison of aspect ratio under different behaviors

3.2. Effective area ratio
In most cases, the feature of aspect ratio can be used to effectively judge whether a passenger falls down or not. However, when a passenger makes some special movements in the elevator, for example,
when people open their arms, the aspect ratio is likely to be greater than 1.00, as shown in Figure 2. According to the threshold analysis, this method can easily lead to misjudgment.

![Figure 2](image)

**a)** Stand Up (Rate=2.23)  **b)** stretch (rate=1.03)

In order to reduce the misjudgment rate, the feature of effective area ratio should be introduced to supplement the discrimination and improve the accuracy of fall detection. Effective area Ratio (EA-ratio) is defined as the Ratio between the pixel area occupied by human body and the pixel area occupied by the minimum external matrix.

The definition is

$$EA - Ratio = \frac{S_{people}}{S_{rectangle}}$$  \hspace{1cm} (2)

Where, $S_{people}$ represents the pixel area occupied by the human body, and $S_{rectangle}$ represents the area of the minimum enclosing rectangle. In the experiment, we set the threshold value of the effective area ratio as 0.8. When the effective area ratio is greater than 0.8, it is highly likely that the person fell down.

3.3. Centroid acceleration

The center of mass is the diagonal intersection point of the smallest outer rectangle, which represents the central position of the human body. In the process of falling, the center of mass of the human body will drop rapidly along the Y axis, and the different behaviors of the human body can be effectively distinguished by calculating the acceleration of the center of mass falling. The acceleration of falling center of mass is defined as shown in Formula (3).

$$a = \frac{2(y_i - y_{i-1})}{\Delta t^2}$$  \hspace{1cm} (3)

Where, $a$ is the centroid acceleration, $y_i$ and $y_{i-1}$ are the Y value between two adjacent frames of images, and $\Delta t$ is the time interval between two adjacent frames of images.

3.4. Algorithm detection process

The detection process in this paper is based on three extracted eigenvalues, namely, the aspect ratio of the outer rectangle, the effective area ratio and the centroid acceleration. The data sample is based on the public video training set and self-shot images, including walking, falling, bending, sitting, squatting and other daily behaviors. After extracting the above three features, it is imported into the SVM trainer for classified training. Finally, it is comprehensively judged whether it is a fall event. Specific detection steps are shown in Figure 3.
First, the image captured by the camera is transmitted to the computer for preprocessing. After the image is grayscale and segmented, the moving target and the minimum outer rectangle in the image are extracted based on the VIBE algorithm. After the moving target is extracted, the aspect ratio of the outer rectangle is calculated. If the aspect ratio is greater than the threshold value (the threshold value is 1), it is judged as a normal event and the next frame of image is read back. If the aspect ratio is less than the threshold value, the effective area ratio and centroid acceleration will be analyzed with the same method. Finally, the fall event will be judged after SVM classification training.

4. Analysis of experimental results

4.1 Recognition and detection of moving objects
In this paper, VIBE algorithm is used to recognize and detect moving targets, in which Figure 4(a) is the original image containing the moving target, and Figure 4(b) is the image after extracting the moving target. According to the experimental results, VIBE algorithm can obtain the moving target more completely, so as to prepare for subsequent fall detection.
4.2 Analysis of fall detection experimental results

In this paper, the detection method is designed to detect the fall of elevator passengers in the elevator car. Mainly comes from the training dataset Le2i fall data set, the data set included 221 fall and fall a video, each section of the video can extract multiple samples for SVM training, the algorithm is based on 1000 training samples as, in elevator capsules for experimental scenario, simulate passengers by ladder of common action, the effectiveness of the detection algorithm, the test results as shown in figure 5 and figure 6.

![Fig 5](image)

Figure 5. Results of daily movement experiment

![Fig 6](image)

Figure 6. Fitting curves of daily movements
In Figure 5, when the target is circled by a green rectangle, it means that the target is in a non-fall state; when the target is circled by a red rectangle, it means that the target is in a fall state. Combined with the fitting curve comparison in Figure 6, when the threshold of aspect ratio is set as 1 and the effective area ratio is set as 0.8, the algorithm can clearly distinguish the fall accidents of elevator passengers.

In order to further test the accuracy of the algorithm, group experiments were carried out for walking, sitting, bending, squatting and other movements that are prone to misjudgment, and each movement was tested for 150 times. Finally, the recognition rate was counted. The results are shown in Table 1. The experimental results show that the algorithm also has a high detection rate for behaviors that are easy to be misjudged.

| Daily behavior | Number of experiments | Fall detected | No fall detected | Recognition rate |
|----------------|-----------------------|---------------|------------------|------------------|
| Walk           | 150                   | 3             | 147              | 98%              |
| Bend           | 150                   | 11            | 139              | 92.67%           |
| Sit Down       | 150                   | 14            | 136              | 90.67%           |
| Squat Down     | 150                   | 7             | 143              | 95.33%           |
| Fall Down      | 150                   | 141           | 9                | 94%              |

5. Conclusion
In order to reduce the injury caused by the elderly falling down in the elevator, this paper designed a fall detection based on ViBe algorithm and multiple features fusion method. Three characteristic parameters, namely aspect ratio feature, effective area ratio and centroid acceleration, are used to comprehensively judge whether there is a fall event. Through many experiments, this method has been proved to be of good real-time performance, high precision and good popularization value.

Reference
[1] NIE Jing. Discussion on Design of Elevator Safety Supervision System Based on Internet of Things. China Equipment Engineering, 2019(23):118-119.
[2] Liu Peng, Lu Tancheng, Lu Yuexing, Deng Yongli, Lu Qiyong. Fall Detection Based on MEMS Three-axis Acceleration Sensor. Chinese Journal of Sensors and Actuators, 2014, 27(04): 570-574.
[3] Tang Yinsheng, Xie Nan, He Jianqiang. Design and Implementation of Fall Detection Algorithm for the Elderly Based on Three-axis Acceleration Sensor. Microcomputer Applications, 2019, 35(02):42-44.
[4] Song Fei. Design and Implementation of ZigBee and OpenCV-based Intelligent Monitoring System for the Elderly. Shanghai Jiao Tong University, 2015.
[5] Yin Zhicheng. Research on Human Abnormal Behavior Recognition Technology Based on Video. Changchun University of Science and Technology, 2020.
[6] Wang pingye. Video fall detection algorithm based on Yolo. Computer programming skills and maintenance, 2019 (11): 137-139
[7] Shi Liuhai. Study and Application of Fall Detection Mechanism Based on Multi-sensor Technology. Zhejiang University of Technology, 2017.