Safety at the platform

M. Miyachi\textsuperscript{1}, Y. Sasaki\textsuperscript{1}, K. Morisaki\textsuperscript{2}, K. Onda\textsuperscript{2} & M. Miyaoka\textsuperscript{2}

\textsuperscript{1}Safety Research Laboratory, East Japan Railway Company, Japan

\textsuperscript{2}Corporate Engineering Division, Matsushita Communication Industrial Co. Ltd., Japan

Abstract

On the Yamate Line and the other main Lines in the Tokyo metropolitan area, fall-detection mats are widely used to stop trains immediately in cases when passengers fall onto the tracks from the platform. However, these can’t detect them in the track.

Therefore, we have developed a high precision fall-detection system which utilizes image processing and other advanced technologies. This system adopts stereo cameras to monitor the fixed area of the track three-dimensionally. In cases when passengers fall onto the track, the system will automatically detect them and work to stop trains ensuring safety.

This system annunciates “passengers’ falls” within a second when passengers may fall onto the track from the platform.

Using the stereo image processing, the influence of shadows and reflected lights over the surrounding circumstances can be easily eliminated, which has been supposed to be difficult by mono-camera image processing [1].

Since one stereo camera covers the range of detection within 40 m, only six cameras are required to monitor the whole of the track when they are equipped in each platform of the Yamate Line in the Tokyo metropolitan area.

This system analyzes one frame every 0.1 seconds. If passengers are falling onto the fixed area, this system annunciates not immediately, but more than three frames of passengers’ fall are observed consecutively. This is so the system is not influenced by the noises generated by accident.

The outline of the total system and the processing will be presented in the paper.
1 Introduction

The stereo image processing has the advantage of detecting an object in three dimensions, by omitting effects of the surroundings such as shadows and reflection, which is difficult with mono-camera image processing. This technology has been employed for a new system to detect a fallen passenger wherever he/she falls on the track. With stereo cameras equipped at the roof of the platform to monitor the whole track obliquely downward, it is now possible to monitor the range of 40 meters. The system would work as follows: when a passenger falls on to the track in a pre-set area for monitoring, an image processing unit installed at a station staff office immediately detects the passenger to report the fact to an observation monitor.

![Image](https://example.com/stereo-camera.png)

**Figure 1:** Fall-detection system using stereo image processing.

2 Principle of stereo image processing

2.1 Detection of the distance to an object

Figure 2 indicates an example of the pictures of an object taken from the stereo cameras located on both sides. The system measures a degree of difference between the two pictures, using the one from the right side as a standard on a board dedicated for image processing. This degree of the gap is called a disparity and Distance $Z$ is then calculated geometrically as expressed in formula (1).

$$Z = \frac{B \times f}{d} \quad (1)$$

where

- $B$: distance between two cameras
- $f$: focal length
- $d$: disparity
2.2 Processing of detected three-dimensional objects

Figure 3 explains how to calculate the height of the object point ‘P’. X denotes a point where a line that goes above the object point ‘P’ from one of the cameras would cross with the ground.

The ratio of distance \( a \), from the camera to the ground, to distance \( b \), from the camera to the object point ‘P’, would be equal to the one of distance \( Z_0 \) to distance \( Z \), both of which is on optical axis. In this case, the height \( h \) of the object point ‘P’ would be obtained from formula (2) according to the conditions of similarity in triangles.

\[
h = H_c \times \left( 1 - \frac{b}{a} \right)
\]

\[
h = H_c \times \left( 1 - \frac{Z}{Z_0} \right)
\]
2.3 Merits of stereo image processing

Compared with mono-camera image processing, stereo camera image processing has the following advantages:
- Easier measurement of the distance to the object, and
- Better performance against outdoor obstacles including shadows and reflection.

Since it has the ability to detect the distance to and the height of the object, it is able to differentiate a person from a small object with ease, which has been difficult with a mono-camera.

Furthermore, the system perceives the height of the shadows cast on the track as 0 meter. Since the system detects a fall within a range of a pre-set height, the effect of the shadows would be easily removed.

3 Fall detection system

The system consists of a stereo camera section with 6 cameras equipped with on at the roof of the platform and a fall-detection processing section at a station staff office. Figure 4 illustrates a configuration of a test sample system for Shinjuku station.

Figure 4: Fall-detection system configuration using image processing.
An example of camera set-up positions is shown in Figure 5. Taking a construction gauge into consideration, it is necessary to locate the center of the cameras on the track side on the extended line from the end of the platform to install a camera at the height of 4.5 meters above the ground. In this case, there would be unfavorable dead zone of 15 cm under the platform. However, since it is unlikely for a passenger to fit in this space due to the size of a fallen passenger, this space is determined to be out of the detection zone.

Figure 5: Set-up position of stereo cameras.

While setting a point of 24.3 m, being ahead from the point right below the camera position, as 0 m, detection area of stereo cameras is determined at 40 m from the standard point. Six cameras are required to cover the track of Yamate line under this condition, as shown in Figure 6. The entire length of the Yamate line platform covered is 240 m: 20 m * 11 cars plus 20 m (10 m spare for both sides of the platform).

Figure 6: Arrangement of stereo cameras.
In installing the cameras, cameras for the areas from No. 2 to No. 6 are set to capture the train from the back to avoid effect by headlights of an approaching train. In terms of the camera for area No. 1, it is set toward the train since the area without a roof requires a new pole to be constructed to set the camera backward as in the other areas.

3.1 Fall detection processing section

Fall detection processing section includes image processing, LAN control, and monitoring PC sections. It is designed to detect a passenger fallen from a platform within one second. Image processing board of the image processing section provides feedback controls for lenses, based on the brightness of each image taken from both sides from the stereo cameras. At the same time, it extracts disparity map out of images from both sides at the processing speed of 100 ms. Besides, CPU board attached to the section detects a fallen passenger from the disparity map sent from the image processing board at the processing speed of 54 ms.

For stable detection of the fallen passenger, it is necessary to monitor the effect of noises, the speed and the direction of the fall. Hence, the system would activate detection when it perceives the fallen passenger in more than 3 frames consecutively to ensure stability. Since it takes approximately 0.4 second for a passenger to fall onto the track, it was necessary to set the processing time per one frame to be less than 133 ms. Therefore, the Pipeline processing described in Figure 7 is adopted to allow to process the whole image processing section within 100 ms cycle, which is 10 frames per second.

To ensure this high speed processing, one image processing section is assigned to each stereo camera section. Image of the fallen passenger detected at the image processing section is automatically transmitted from the image processing section to monitoring PC section via LAN control section. The monitoring PC section uses generally marketed personal computers to display settings of detection area and processing status of each camera. Figure 8 shows an example of a monitor screen when a fallen passenger is detected.

This is a case when No. 5 camera detects a fallen passenger. Camera status sign on the left side of the Figure shows that No. 5 camera detects a fallen passenger. The detected image is then to be shown in the upper right part of the screen. This image is saved even after the fallen passenger is removed from the
monitored area, which makes it easier to confirm how a station staff member realized the fall. Twenty-one frames in total, including 10 frames each before and after the incident was first detected by a hard disk at the PC control section, is to be saved for immediate play back.

Figure 8: Example of monitor screens when a fallen passenger is detected.

Stereo camera and fall-detection processing sections are connected with high speed and strong noise resistance optical fiber cables. Though one optical fiber cable is currently assigned to each set of stereo camera and image processing sections to obtain accurate signal data for the test sample system, arterial cables or radio would be considered to reduce installation cost upon introduction of a completed system.

4 Fall-detection algorithm

Figure 9 describes outline (A) to (C) of the detection when a passenger falls down to the pre-set area for the system on the track.

At first, the system extracts high frequency pictures out of input stereo images taken from the right and the left sides, respectively, then divides this picture into three categories to obtain a feature image (A). Next, using this feature image, the system carries out stereo matching (B) between the left and the right pictures to obtain a disparity map. The system extracts a fallen passenger when the following conditions are satisfied: a fallen object with a size of more than 5 blocks (i.e., a size of 30 cm cubed box 40 m away) is being detected within a pre-designated area (i.e., in the height range of above 15 cm but below 100 cm) for more than 3 consecutive frames (C). Process (A) and (B) is to be conducted on image processing board illustrated in Figure 4. The judgment described in Process (C) is carried out with CPU board.
5 Performance of fall detection

5.1 Current performance

To ensure the stability of this system, long-term capacity tests have been currently conducted at Yamate line Shinjuku Station. The trial fall detection tests are performed at night to evaluate detection capacities. Table 1 indicates a relation between the distance from the 0-meter point and detection capacities assuming 24.3 m away from a point below a camera as 0 meter. A separate basic test during daytime shows that the system is capable to percept a box of 30 cm cubed at 40-m point. At the nighttime trial test, however, the system sometimes failed to percept a person in black laid down at a distance of over 35-m.

One of the reasons is assumed as follows: During the daytime, brightness is approximately 100,000 lx under the sunshine and approximately 10,000 lx even...
in the shadows. During the nighttime, however, since the brightness at the center of the track is only 150 lx, a passenger is assimilated with the surroundings and the system could not extract features shown in (A) of Figure 9, and fails to obtain disparity.

Taking the results of the fall detection trial test, the following measures are considered to ensure certain extraction even in the nighttime:
1) To control optimal lens iris based on the brightness of the track,
2) To improve sensitivity of a feature extraction filter, and
3) To utilize data on cases when a person falls.

Table 1: Performance of fall detection system.

<table>
<thead>
<tr>
<th></th>
<th>0m</th>
<th>20m</th>
<th>25m</th>
<th>30m</th>
<th>35m</th>
<th>40m</th>
</tr>
</thead>
<tbody>
<tr>
<td>Falling person</td>
<td>○</td>
<td>⋯</td>
<td>○</td>
<td>○</td>
<td>○</td>
<td>○</td>
</tr>
<tr>
<td>Decumbent person</td>
<td>○</td>
<td>⋯</td>
<td>○</td>
<td>○</td>
<td>△</td>
<td>△</td>
</tr>
</tbody>
</table>

△: sometime can’t detest

5.2 Occurrence of false alarms

To make an image recognition system fit for practical use, we must reduce the number of false alarm.

Figure 10 shows the samples of false alarm.

In addition, table 2 shows the number of false alarm for about two weeks (2002.6.14-2002.7.1). For these false alarms, we now are trying to distinguish the person from the other small obstacles (polythene bag, news paper, cat, etc.) by using the height of the object.

Figure 11 shows the distribution map of the object height. Black marks are the data of person, and white marks are the data of the other small obstacles. If we set the threshold to 0.6m, we will distinguish the person from the other small obstacles.

Figure 10: Samples of false alarm.
Table 2: The number of false alarms.

<table>
<thead>
<tr>
<th>Distance (m)</th>
<th>(a) polythene bag</th>
<th>(b) news paper</th>
<th>(c) cat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Camera 1</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Camera 2</td>
<td>0</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>Camera 3</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Camera 4</td>
<td>14</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Camera 5</td>
<td>2</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Camera 6</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>16</td>
<td>11</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 11: Distribution map of the object height.

6 Conclusion

Though it is verified that there is no problem in detecting a fall during daytime, problems do exist to ensure certain detection of a person in black during nighttime, for which the authors would take step in future. In addition, the authors have plans to improve the algorithm of reducing the false alarms, and to confirm various effects on fall detection system, which would be caused by factors such as snow.

References