Chapter 8
Dynamic Analysis of Crowd Behavior$^{1, 2, 3, 4}$

8.1 Behavior Analysis of Individuals in Crowds

In many cases, we cannot obtain the background image easily, and if the occlusion is serious, then it is difficult to segment masses of blobs into single blobs. We therefore use optical flow to detect abnormal behavior in crowded environments $^{[57]}$, $^{[58]}$.

The velocity of each pixel is calculated using optical flow and then filtered using a $3 \times 3$ template. A predetermined threshold is used to obtain the binary images. If the value of the filtered velocity is larger than the threshold, then we set the gray value of this pixel to 255; otherwise, it is set to 0.

The image on the screen is distorted because of the perspective rule, and thus a camera model $^{[164]}$ is used and compensation algorithms are employed so that the length and speed are roughly independent of the depth of a blob from the camera.

After the binary images are obtained, ordinary methods are used to detect abnormal behavior in a very crowded environment.

In this experiment, we aim to detect two types of abnormal behavior using optical flow: (1) a person running in a shopping mall, and (2) a person waving a hand in a crowd.

$^{1}$ Portions reprinted, with permission, from Ning Ding, Yongquan Chen, Zhi Zhong, Yangsheng Xu, Energy-based surveillance systems for ATM machines, 2010 8th World Congress on Intelligent Control and Automation (WCICA). ©[2010] IEEE.


$^{3}$ Portions reprinted, with permission, from Zhi Zhong, Weizhong Ye, Ming Yang, and Yangsheng Xu, Crowd Energy and Feature Analysis, Proceedings of the 2007 IEEE International Conference on Integration Technology. ©[2007] IEEE.

The detection of a person running in a shopping mall is very difficult using the ordinary foreground subtraction method. It is very difficult to obtain the background image in a crowded shopping mall, and even if the background is available, segmentation is extremely difficult.

However, using the optical flow method, we can obtain the velocity of each pixel, and then the revised algorithm is used to compensate for the distortion of the velocity image. It is very important to compensate for the perspective distortion of the velocity image. In the camera image, the people near the camera will be bigger than the people further away from it, and thus the velocity of the people near the camera will be faster than their real velocity if the distortion is not revised.

The velocity image is then filtered by a $3 \times 3$ template, and a threshold is set to obtain the binary velocity image. To remove the noise, the pixels of the binary image are grouped into blobs. Blobs with small areas are removed. Figure 8.1 shows how the system detects a person running in a shopping mall. The red circles in the pictures hereinafter are used to highlight the abnormal region.

We also use the optical flow method to detect a person waving a hand in a crowd. Figure 8.2 shows the three original consecutive images and three binary images obtained by the optical flow method. Periodicity analysis [162] is applied to the images shown in Figure 8.3, and the hand of the person is detected after two periods.

In very crowded regions in which extreme overlapping occurs, obtaining the background and successful segmentation are difficult. In such cases, the optical flow approach is effective for the detection of abnormal behavior.

### 8.2 Energy-based Behavior Analysis of Groups in Crowds

We have elaborated the estimation of crowd density in the previous chapter. However, density estimation is a static issue, whereas the detection of crowd abnormalities is a dynamic one. The latter involves factors such as direction, velocity, and acceleration, which are too important to be ignored.

In this section, we address two energy methods, which are based on intensity variation and motion features, respectively. The results of wavelet analysis of the energy curves show that both methods can satisfactorily deal with crowd modeling and real-time surveillance. A comparison of the two methods is then made in a real environment using a metro surveillance system.

To tackle crowd density problems, researchers have tried to establish precise models to build virtual environments in which crowd situations are simulated [134][147][154]. However, pedestrian kinetics in a virtual environment is completely different from that in the real world, and the image or video quality generated in a virtual environment is always much better than that obtained from a real one. Because of these limitations and the tremendous amount of computation involved, such a method is not useful in actual video surveillance systems. The second method is tracking, in which the system tries to acquire people’s trajectories by keeping track of individuals as they enter the monitored scene to detect abnormalities [133][137]. This method has been proven effective for the surveillance of abnormal individuals.