Real-time Vision-based People Counting System for the Security Door

Jae-Won Kim, Kang-Sun Choi, Byeong-Doo Choi, and Sung-Jea Ko
Department of Electronics Engineering,
Korea University, Anam-dong, Sungbuk-ku, Seoul, 136-701, Korea
Tel. +82-2-3290-3228, Fax.: +82-2-3290-3672
E-mail: sjko@dali.korea.ac.kr

Abstract: This paper describes an implementation method for the people counting system which detects and tracks moving people using a fixed single camera. This system counts the number of moving objects (people) entering the security door. Moreover, the detected objects are tracked by the proposed tracking algorithm before entering the door. The proposed system with Intel Pentium IV operates at an average rate of 10 frames a second on real world scenes where up to 6 persons come into the view of a vertically mounted camera.

1. Introduction

Real-time people flow information is very useful source for security application as well as people management such as pedestrian traffic management, tourists flows estimation. To track and count moving people is considered important for the office security or the marketing research. Many of such measurements are still carried out on manual works of persons. Therefore it is necessary to develop the automatic method of counting the passing people.

Several attempts have been made to track pedestrians. Segen and Pingali [1] introduced a system in which the pedestrian silhouette is extracted and tracked. The system runs in real-time, however, the algorithm is too heavy to track many people simultaneously and can not deal well with temporary occlusion. Masoud and Papanikolopoulos [2] developed a real-time system in which pedestrians were modeled as rectangular patches with a certain dynamic behavior. The system had robustness under partial or full occlusions of pedestrians by estimating pedestrian parameters. Rossi and Bozzoli [3] avoided the occlusion problem by mounting the camera vertically in their system in order to track and count passing people in a corridor, but assumed that people enter the scene along only two directions (top and bottom side of the image). Terada [4] proposed a counting method which segmented the human region and road region by using the three dimensional data obtained from a stereo camera. However, this system also assumed only simple movement of pedestrians.

In this paper, we propose a real-time people counting system with a single camera for security inside the building. The camera is hung from the ceiling of the gate so that the image data of the passing people are not fully overlapped. The implemented system recognizes people movement along various directions. To track people even when their images are partially overlapped, the proposed system estimates and tracks a bounding box enclosing each person in the tracking region. The approximated convex hull of each individual in the tracking area is obtained to provide more accurate tracking information.

This paper is organized as follows: Section 2 describes the system architecture of the proposed people counting system. In Section 3, the people counting and tracking algorithms are given in detail. Section 4 presents the experimental results and conclusion.

2. System Architecture

Fig. 1 shows a scene of the passing people through the security door inside the building. There are incoming and outgoing individuals in the scene. A single camera unit is hung from the ceiling of the gate so that the passing people can be observed and tracked in a tracking area in front of the door. The image captured by the camera are processed and the number of the passing people is calculated.

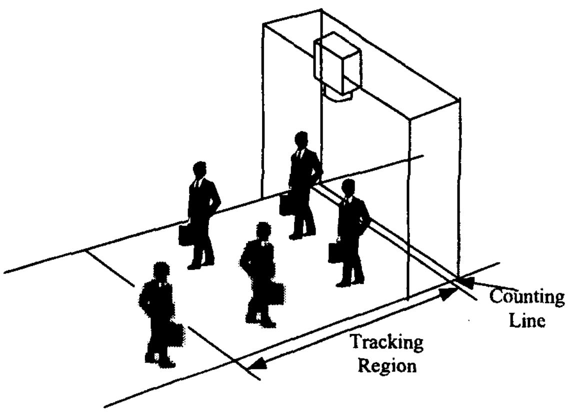

Fig. 1. Scene of the gate.

To cope with inherently dynamic phenomena (people enter the scene, move across the field of view of the camera, and finally cross the counting line), the people tracking and counting problem has been decomposed into the following three steps: [3]

- Determine whether any potentially interesting objects have entered the scene (Alerting phase);
- Track their motion until the counting line is reached (Tracking phase);
- Establish how many people correspond to tracked objects (Interpretation phase).

The proposed system provides the graphic user interface (GUI) to define the alerting area, the tracking area, and the counting line as shown in Fig. 2.

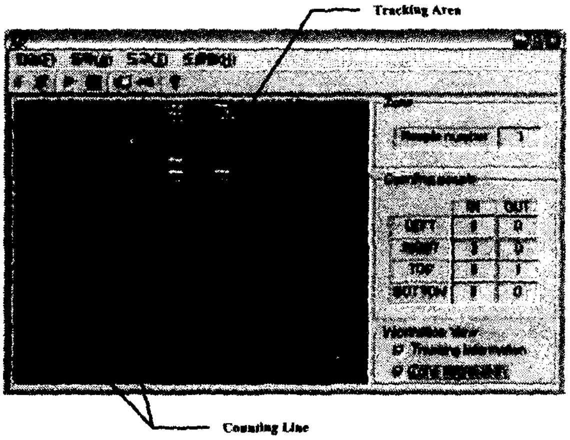

Fig. 2. The graphic user interface of the proposed system.

Fig. 3 illustrates the flow chart of the proposed algorithm. Background subtraction and thresholding are performed to produce difference images. The difference image is preprocessed by a morphological opening operator (erosion followed by dilation) to remove small clusters in the image. Then, each object is matched the corresponding object in the previous captured image by comparing their center positions. This tracking information is used for counting people. As shown in Fig. 2, the proposed system shows the trajectories and traffic information of incoming and outgoing people.

The proposed algorithm uses two types of difference images. These two difference images contain different information of newly incoming objects. The left flow generates a background-subtracted image where moving objects appear. However, if the pattern of moving object is similar to that of background, the moving object can not be distinguished. The other difference image is obtained by subtracting two successive images. This image can provide motion (boundary) information of moving object even when the moving object is similar to the background. However, when the object stays at a position, the motion information does not appear. On the other hand, when the object moves fast, the boundary information is blurred and incorrect.

3. Proposed People Counting and Tracking Algorithm

3.1 Background Estimation

The background estimation method affects the performance of the system. Since lighting conditions vary in time, the background image is updated using a very slow recursive function to capture slow changes in the background. Background subtraction has been used by many to extract moving objects in the scene [1], [5], [6], [7]. To obtain more accurate background, we propose an adaptive background estimation algorithm which has robustness under the change of illumination as follows:

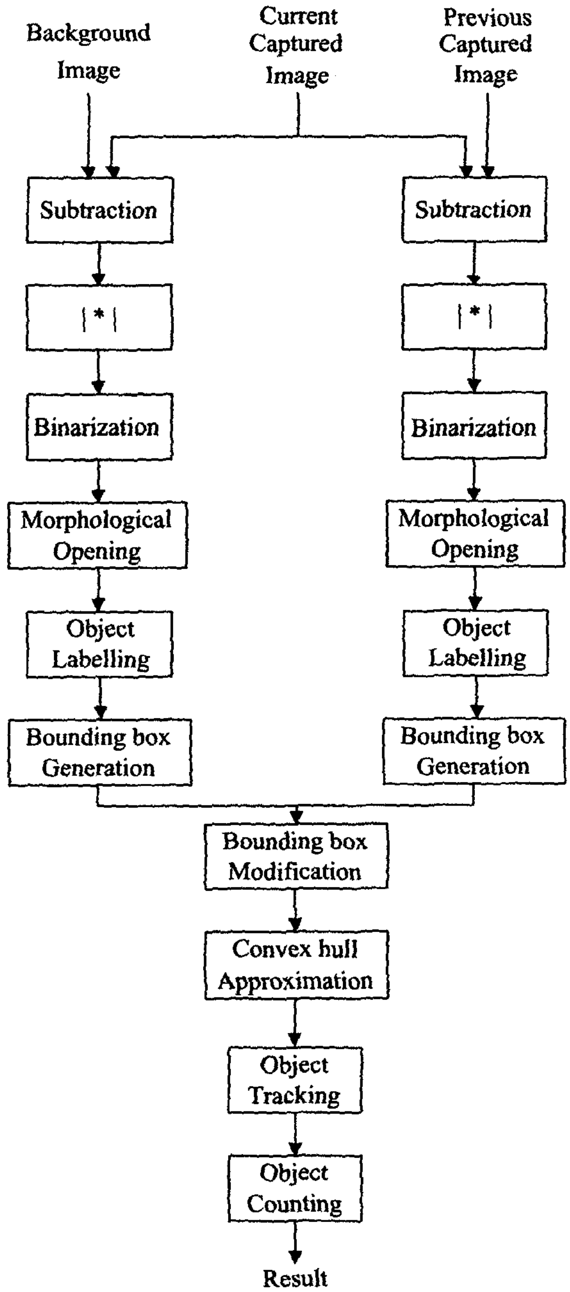

Fig. 3. Main steps in the proposed people counting system.

First, the system determines whether moving objects exist in a current image by comparing the current image with the previous captured image. If there is no moving object, a new background image is obtained by averaging three images including the previous background image, the current background image, and the current captured image.

This background estimation method deals well with the gradual change of illumination, but can not deal with the abrupt large change of illumination occurred in the whole image. To overcome this problem, we compensate the average intensity level of the illumination. From the background image and the current captured image, the average intensity levels of the four blocks each of which is located at the corner of the images are measured. The intensity difference between the corresponding intensity levels of two images is calculated and applied to the current background image to compensate the abrupt change of intensity if moving objects exist in the current image.

3.2 Object Extraction

After object labelling is performed, each object is determined whether it is a main body or a part of the body. The part of the body is merged into the closest main body. Then, the bounding box of each merged object which is the smallest rectangle surrounding the object is obtained. Fig. 4 shows the bounding box surrounding a binarized object mask.

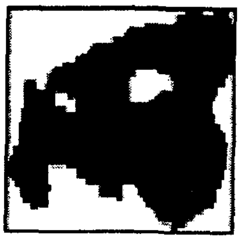

Fig. 4. Mask and its associated bounding box.

The bounding box is used for tracking people since several features of the bounding box such as area, center point, and boundary are less changeable than those features of the object mask, and even when occlusion of people occurs in the captured image, separation of each person can be achieved by using the past tracking information about bounding boxes. To extract more accurate object features, the simple convex hull of the object inside the bounding box is approximated. As shown in Fig. 5, 24 points of approximated convex hull are found by rotating the bounding box at the every angle of 15 degrees.

3.3 Object Tracking

For the approximated convex hull of each merged object, the corresponding convex hull in the previous captured image is searched for by comparing their center positions. In the object tracking algorithm, the center point of convex hull verified as a pedestrian in the previous image is predicted in the current image by analyzing the past velocity information of the moving object as

$$\hat{v}(t+1) = (1-a)v(t) + a\hat{v}(t), \qquad (1)$$

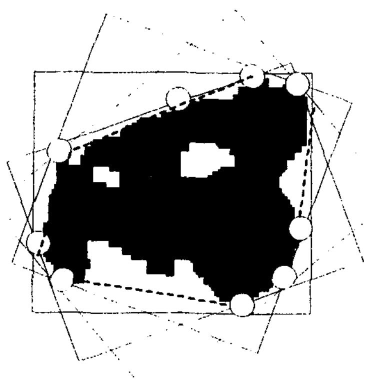

Fig. 5. Convex hull approximation.

where a is a damping coefficient, $v(t)$ and $\hat{v}(t+1)$ are the measured displacement vector at time t and the predicted displacement vector at time $t+1$, respectively.

Fig. 6 shows the procedure of the object tracking method using motion prediction. First, the next position $P(t+1)$ of the current object at time t is estimated by shifting the current actual center point $C(t)$ by $\hat{v}(t+1)$. Then, the system defines the circular search range centered on $P(t+1)$ with the predefined radius and searches for an object whose center point is inside the circular search range. In Fig. 6, $C(t+1)$ represents the center point of the detected object. The tracking information is updated by linking the information of the detected object to the information of the tracked object. From $C(t+1)$, this procedure is iterated.

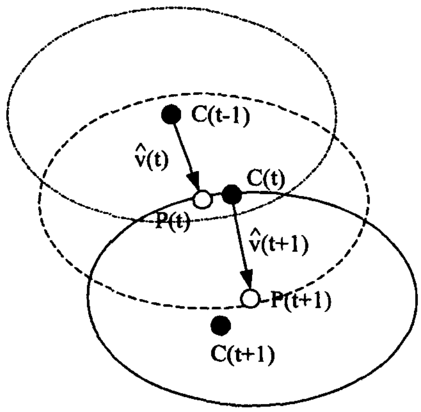

Fig. 6. Object tracking using motion prediction method.

4. Experiments and Conclusion

The proposed people counting system was implemented to operate at about 10 frames per second and use an analog CCD camera. The experiments were performed to demonstrate the performance of the proposed people counting system. These experiments were carried out using several test sequences as shown in Fig. 7. The CCD camera was set 3m above the floor. Although this situation can introduce the false warning, the bounding box representation which allows partial object occlusion reduces the false warning. Correct people counting of 96% was obtained.

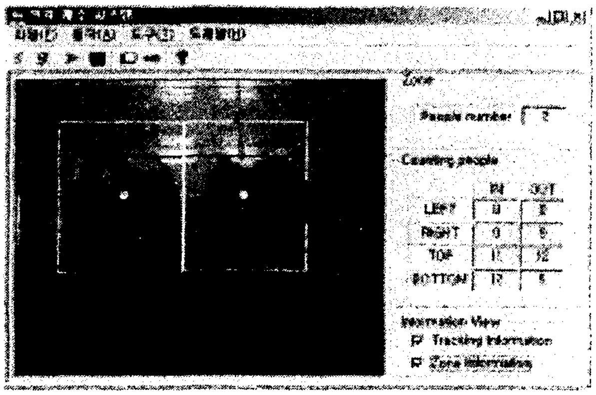

Fig. 7. The proposed people counting system.

In this paper, we have described an implementation method for the people counting system which detects and tracks moving people using a fixed single camera. To track people even when their images are partially overlapped, the proposed system estimates a bounding box enclosing each person and tracks the bounding box. The simple convex hull of each person in the tracking area is approximated to provide more accurate tracking information.

References

[1] J. Segen and S. Pingali, "A Camera-Based System for Tracking People in Real Time," *IEEE Proc. of Int. Conf. Pattern Recognition*, Vol.3, pp.63-67, 1996.

[2] O. Masoud and N. P. Papanikolopoulos, "A novel method for tracking and counting pedestrians in real-time using a single camera," *IEEE Trans. on Vehicular Tech.*, Vol. 50, No. 5, pp.1267-1278, 2001.

[3] M. Rossi and A. Bozzoli, "Tracking and Counting Moving People," *IEEE Proc. of Int.Conf. Image Processing*, Vol. 3, pp.212-216, 1994.

[4] K. Terada, D. Yoshida, S. Oe, J. Yamaguchi, "A counting method of the number of passing people using a stereo camera," *IEEE Proc. of Industrial Electronics Conf.*, Vol. 3, pp.1318-1323, 1999.

[5] A. Baumberg and D. Hogg, "Learning flexible models from image sequences," *Proc. Eur. Conf. Computer Vision*, vol. 1, Berlin, Germany, pp.229-308, 1994.

[6] C. Smith, C. Richards, S. A. Brandt and N. P. Papanikolopoulos, "Visual tracking for intelligent vehicle-highway systems," *IEEE Trans. on Veh. Technol.*, Vol.45, pp.744-759, 1996.

[7] Q. Cai and J. K. Aggarwal, "Tracking human motion using multiple cameras," *Proc. 13th Int. Conf. Pattern Recognition*, Los Alamitos, CA, pp.68-72, 1996.

[8] D. Murray and A. Basu, "Motion tracking with an active camera," *IEEE Trans. on Pattern Analysis and Macine Intelligence*, Vol. 16, pp.449-459, 1994.

[9] C. Gu and M. Lee, "Semiautomatic segmentation and tracking of semantic video objects," *IEEE Trans. on Circuits and Sytems for Video Technology*, Vol. 8, pp.572-584, 1998.

[10] M. K. Leung and Y. H. Yang, "First Sight: A Human Body Outline Labelling System," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol.17, no.4, pp.359-377, 1995.

[11] J. Rehg and T. Kanade, "Model-Based Tracking of Self-Occluding Articulated Objects," *Proc. Int. Conf. Computer Vision*, pp.612-617, 1995.

[12] K. Rohr, "Towards Model-Based Recognition of Human Movements in Image Sequences," *CVGIP: Image Understanding*, Vol.59, No.1, pp.94-115, 1994.

[13] J. O'Rourke and N. J. Badler, "Model-based image analsys of human motion using constraint propagation," *IEEE Trans. on Pattern Analysis and Machine Inteligence*, Vol.2, pp.522-536, 1980.