People Counting System for Getting In/Out of a Bus Based on Video Processing

[1]Chao-Ho (Thou-Ho) Chen, [1]Yin-Chan Chang, [1]Tsong-Yi Chen and [2]Da-Jinn Wang
[1]Department of Electronic Engineering, National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan 807, R.O.C
[2]Department of Information Management, National Kaohsiung Marine University, Kaohsiung Taiwan 811, R.O.C
E-mail: thouho@cc.kuas.edu.tw

Abstract

This paper presents an automatic people counting system for getting in/out of a bus based on video processing. The basic scheme is to set a zenithal camera in the bus for capturing the passenger flow bi-directionally. The captured frame is firstly divided into many blocks and each block will be classified according to its motion vector. If the block quantity of similar motion vectors is more than a threshold, those blocks are regarded as belonging to the same moving object. As a result, the number of such moving objects is counted to be the passenger number of getting in or out of a bus. can be segmented for counting. Experimental results show that the proposed bus-passenger counting algorithm can provide a high count accuracy of 92% on average.

1. Introduction

Recently, the intelligent transportation surveillance becomes the major trend of the development in traffic control. The information of the number of passengers getting in/out of a transporter is very important for control and management of the mass transport system. Generally, an early automatic counting approach of using light beams had suffered one intractable problem: they could not count the passengers accurately unless there is only one person passing through the base-line at one time. To solve this problem, the image-processing based approach in various applications [1][2] is hence motivated and may provide a real-time automatic counting for passengers getting in/out of a transporter by analyzing a series of images captured with a video camera.

In many reported image-processing based methods for counting people passing through a door or gate [3]-[9], the main problems include occlusion, overlapping, merge-split, and shadow effect. To avoid the occlusion problem, Rossi and Bozzoli [3] and Sexton et al. [4] mounted the camera vertically with respect to the floor plane and set the optical axis of the camera in such a way that the passing people could be observed from just overhead. To increase the count of passing people through a gate at one time, Terada et al. [5] used the stereo images captured by a pair of cameras to cope with both problems of the crowd counting and direction recognition of the passing people. Based on cost-effective consideration, a single camera with a tracking algorithm may be the better solution and thus Masoud and Papanikolopoulos [6] developed a rectangular model-based recognition of the pedestrian with human motion analysis to achieve a reliable people count. By setting a fixed single camera hung from the ceiling of the gate, Kim et al. [7] proposed a real-time scheme to detect and track the people moving in various directions with a bounding box enclosing each person. Also using a single zenithal camera, Bescos et al. [8] introduced a DCT based segmentation, which can efficiently consider both lighting and texture information, to cope with some problems, such as shadows, sudden changes in background illumination and sporadic camera motion due to vibration, in order to count people crossing an entrance to a big store. To overcome both inherent problems of people-image overlapping and merge-split, a two-stage counting strategy using features of area and color information is proposed for counting people passing through a gate or door [9].

Basically, a zenithal camera setting scheme can avoid occlusion problem and obtain a cost-effective solution for counting passengers getting in/out of a transporter [1][2]. However, to accurately count passengers getting in and out of a bus, the major

978-0-7695-3382-7/08 $25.00 © 2008 IEEE
DOI 10.1109/ISDA.2008.335

difficulty is how to cope with both problems of camera shaking and variation of illumination that will seriously affect the frame-difference techniques for moving-object segmentation. To solve the above problems, this research adopts a zenithal camera setting scheme and uses block motion features for provide a cost-effective people counting system for getting in/out a bus.

2. The proposed bus-passenger counting system

The proposed bus-passenger counting system adopts a zenithal camera setting scheme in which it will capture the downward-vertical view of image while passenger getting in/out of a bus. Figure 1 shows the flowchart of the counting scheme which is composed of three models: passenger detection and segmentation, collection of passenger's feature and passenger analysis and counting.

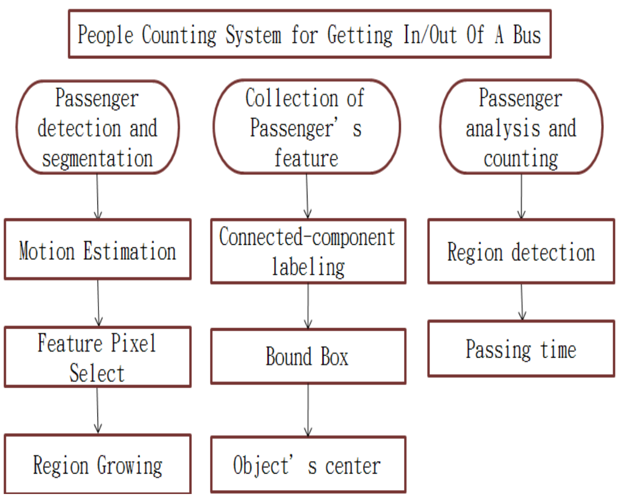

Figure 1 The proposed bus-passenger counting system

2.1. Passenger detection and segmentation

The first step of this system is to extract moving pixels because the major feature of passenger is forward motion. Nevertheless, the conventional technology of frame difference can't segment the passenger owing to shaking and illumination variation in the door of a bus. To cope with such problems, the dynamic features of passenger are exploited for distinguishing the opening action of door and shaking and extracting the moving passengers. Among dynamic features, the moving direction of a passenger is always toward the door and thereby the motion vector can be employed to distinguish the passenger from the other moving objects. At first, an image frame is divided into many k*k blocks and then the motion vector of each block is calculated. When the quantity of similar motion vectors of vertical-bias direction (toward the door) is more than a threshold, such a heap of blocks is regarded as a passenger. Figure 2 shows the result of motion estimation using block-size of 16*16, in which red arrow means the estimated motion vectors. For more detailed motion estimation, each 16*16 block is divided into many 4*4 blocks and the search for each 4*4 block is limited in a range of [-1, 1]. Figure 3 demonstrates the result of motion estimation using block-size of 4*4, in which color pixels indicate the selected blocks.

By selecting feature pixels, Figure 3 demonstrates the result of motion estimation using block-size of 4*4, in which color pixels indicate the selected blocks. From the figure, there are still many non-passenger blocks to be detected owing to false motion estimation resulting from lens shaking and light changing. Anyway, this problem of false motion estimation can be removed by the following method, called "feature pixel selection".

After motion estimation, the feature pixel selection will search the pixels of most significant feature in the selected block. The feature value M is calculated by differential matrix at the neighborhood of (x, y), as described in the following equation (1), where S(p) is the neighborhood of size nxn at (x, y) and I is the graylevel value of pixel. The pixel with more significant feature means that its pixel value varies heavily around its neighborhood, i.e., the pixel seems to belong to the moving object. In Figure 3, non-black pixels denote the selected block by motion estimation and then such pixels are permuted in order of feature value, followed by selecting the front k (k=150) pixels indicated by yellow circle. Obviously, those pixels with greater feature value are likely located on the moving object.

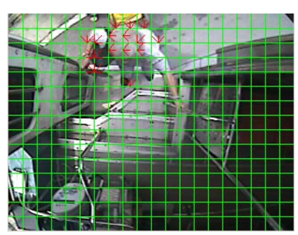

Figure 2. Result of motion estimation using block-size of 16*16.

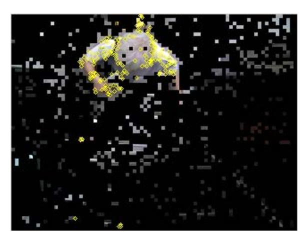

Figure 3. Result of feature pixel selection through motion estimation using block-size of 4*4.

$$M = \begin{bmatrix} \sum_{s(p)} \left(\frac{dI}{dX} \right)^2 & \sum_{s(p)} \left(\frac{dI}{dX} \cdot \frac{dI}{dY} \right) \\ \sum_{s(p)} \left(\frac{dI}{dX} \cdot \frac{dI}{dY} \right) & \sum_{s(p)} \left(\frac{dI}{dY} \right)^2 \end{bmatrix} \quad (1)$$

To refine the above segmentation result, region growing technique is employed for gathering pixels of moving object and removing trivial pixels. This procedure takes the feature pixels derived above for seeds and then grows at an eight-connection search, according to the following condition. The result of region growing is shown in Figure 4, in which all the wrong pixels selected by motion estimation are removed.

$$\begin{aligned} P(x',y')_R - Low \le P(x,y)_R \le P(x',y')_R + Up \\ P(x',y')_G - Low \le P(x,y)_G \le P(x',y')_G + Up \\ P(x',y')_B - Low \le P(x,y)_B \le P(x',y')_B + Up \end{aligned} \quad (2)$$

where $P(x, y)$ is seed pixel, $P(x', y')$ is the neighboring pixel, and low and up mean thresholds.

2.2. Collection of passenger's features

After extracting the passenger's pixels by region growing, as shown in Figure 4, such pixels need to be merged into an object according its feature and then record its direction for passenger counting. The connected-component labeling technique [11] is introduced to merge those pixels of similar feature into the same object (i.e., passenger). Figure 5 plots the flowchart of the connected-component labeling process. In the labeling process, firstly each object is labeled by a different mark that is to code the small objects for later classification and then it starts to search the similar object. If it finds the similar object, the mark of this object is replaced by the original mark. After the labeling process is performed through all objects, the trivial small objects will be regarded as the part of background and hence need to be removed from the passenger. Finally, the remained pixels can be viewed as the passenger's image.

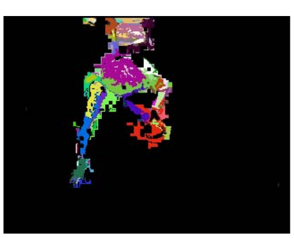

Figure 4. Result of region growing.

Figure 6. Two virtual base-lines.

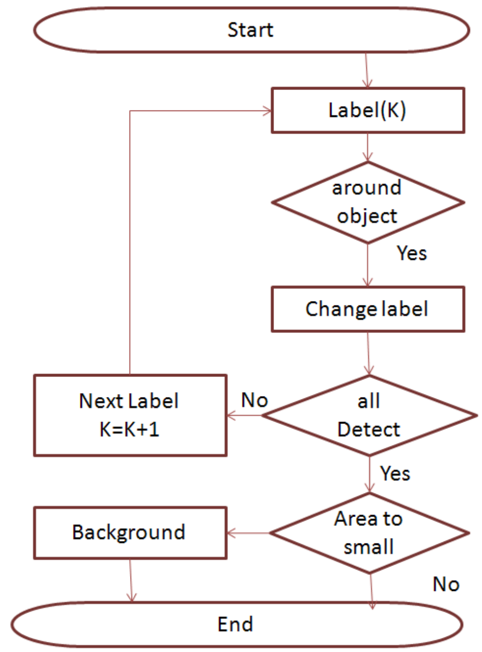

Figure 5. Flowchart of the connected-component labeling

After deriving the objects of the same mark, formula (3) is used to find the bounding-box of that moving object and then calculate the center of its gravity, (x_0, y_0). Therefore, the passenger can be tracked with its position and moving direction.

$$\begin{aligned} & x_0 = \frac{\sum_{(x,y\in R)} \sum x}{\sum_{(x,y\in R)} \sum 1}, y_0 = \frac{\sum_{(x,y\in R)} \sum xy}{\sum_{(x,y\in R)} \sum 1} \\ & top = \min_y \{(x,y) \mid (x,y) \in R\} \\ & bottom = \max_y \{(x,y) \mid (x,y) \in R\} \\ & left = \min_x \{(x,y) \mid (x,y) \in R\} \\ & right = \max_x \{(x,y) \mid (x,y) \in R\} \end{aligned} \quad (3)$$

2.3. Passenger analysis and counting

Based on tracking the moving object's position and moving direction, the counting direction can be determined according to the order of moving across two base-lines or the base-line from which the people go away, during the tracking process. This direction-orienting strategy will require the least computational cost. In such two virtual base-lines, one is on the upper stair and the other one is on the lower stair, as shown in Figure 6. When the moving object crosses two base-lines, it is sure that the passenger will stand on the stair and then it is counted. The counting algorithm is described in Figure 7.

For getting in/out of a bus, only one passenger can pass through the entrance door each time, in which there are three steps at the entrance stair. So, it is necessary to firstly detect if there is any passenger stands on the stair region. If the bounding-box of a passenger crosses the two base-lines, it is regarded as that the passenger is passing through and hence it is counted. Figure 8 shows that two passengers are getting on the bus, one is on the entrance stair and the other stands outside the bus.

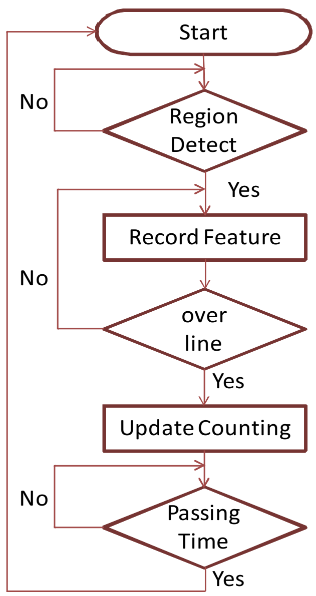

Figure 7. The counting algorithm

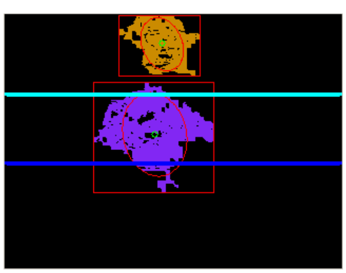

Figure 8. Passenger counting.

To avoid the repeated counting, the bounding-box will be tracked completely because sometimes it may stay on the stair for a long time. For this reason, the time of passing through the stair is set to 1.5 second and no count is performed in such a time period. In the conventional tracking techniques, the intersection check of the bounding-box pair for person-identification may be not allowable since the passengers are usually crowded. Besides, some passenger entering situations with the intentional disturbance will be not considered in this research.

3. Experimental results and discussions

A theoretical analysis about the proposed bus-passenger counting system has been given in the above section, but the implementation of such a counting system with the practical image sequences can provide a realistic and interesting evaluation. To realize the proposed counting system, a color video camera using SONY CAD-2017 CMOS sensor is set 245cm above the first stair at the front door of Kaohsiung City bus, with a surveillance area of 320 × 240 pixels (i.e., frame size). The capture rate of the camera is 30 (frames/second) on Intel Core 2 CPU T5500 NB computer. For the deep evaluation, three various types of test videos all involve passengers getting in/out of a bus at one stop, i.e. recording passenger flow from door-opening to door-closing. The counting accuracy is defined according to equation (4) and those accuracies of simulations are tabulated in Table 1. In the equation, C_{in}, R_{in}, C_{out}, and R_{out} denote the count and real number for getting in a bus and out of a bus, respectively, and the selected partial accuracy (C/R or R/C) will be lower than 100%.

$$\begin{aligned} &Accuracy \\ &= \frac{1}{2}\times\big[(Cin/Rin \text{ or } Rin/Cin) \\ &\quad +(Cout/Rou \text{ or } Rout/Cout)\big]\times 100\% \end{aligned} \tag{4}$$

Table 1. Accuracy of counting passenger-flow for getting in/out of a bus

Test samples	Real passengers		Counted passengers		Accuracy
	In	Out	In	Out	
video1(52sec)	10	1	9	1	95%
video2(65sec)	17	4	15	4	94%
video3(82sec)	29	6	33	5	86%
Average					92%

From Table 1, it can be observed that the accuracies of both video1 and video2 are higher than that of video3. This is because in video3 the passenger flow of getting into the bus is crowded and, besides, there are some passengers are also getting out of the bus at the same time. The passenger flow is so crowded that someone person may stay on the stair for longer time as to be counted twice. However, the proposed bus-passenger counting algorithm can provide a high count accuracy of 92% on average.

4. Conclusions

This paper proposes a cost-effective automatic passenger counting system for getting in/out of a bus based on video processing. By a zenithal camera setting in the bus, block-based motion feature is exploited for classifying each block according to its motion vector and those blocks with the similar motion vector are collected to form a passenger object for counting. Thus, the inherent problems of camera shaking and variation of illumination in the bus can be overcome. Experimental results manifests that the proposed bus-passenger counting system with 92% accuracy will be more attractive for counting passengers getting in/out of a bus.

5. References

[1] F. Bartolini, V. Cappellini and A. Mecocci, "Counting people getting in and out of a bus by real time image-sequence processing", *Image and Vision Computing*, vol. 12, no. 1, Jan. 1994, pp. 36-41.

[2] A. Albiol, I. Mora and V. Naranjo, "Real-time high density people counter using morphological tools", *IEEE Trans. Intelligent Transportation Systems*, vol. 2, no. 4, Dec. 2001, pp. 204-218.

[3] M. Rossi and A. Bozzoli, "Tracking and counting moving people", *IEEE International Conference on Image Processing (ICIP)*, vol. 3, 1994, pp. 212-216.

[4] G. Sexton, X. Zhang and G. Redpath, "Advances in automatic counting of pedestrians", *1995 European Convention on Security and Detection (ECSD95)*, 1995, pp. 106-110.

[5] K. Terada, D. Yoshida, S. Oe and J. Yamaguchi, "A method of counting the passing people by using the stereo images", *IEEE International Conference on Image Processing (ICIP)*, vol. 2, 1999, pp. 338-342.

[6] O. Masoud and N. P. Papanikolopoulos, "A novel method for tracking and counting pedestrians in real-time using a single camera", *IEEE Trans. Vehicular Technology*, vol. 50, no. 5, Sep. 2001, pp. 1267-1278.

[7] J. W. Kim, K. S. Choi, B. D. Choi and S. J. Ko, "Real-time vision-based people counting system for security door", *International Technical Conference on Circuits/Systems Computers and Communications*, 2002, pp. 1416-1419.

[8] J. Bescos, J. M. Menendez and N. Garcia, "DCT based segmentation applied to a scalable zenithal people counter", *IEEE International Conference on Image Processing (ICIP)*, vol. 3, 2003, pp. 1005-1008.

[9] T. H. Chen and C. W. Hsu, "An automatic bi-Directional passing-people counting method based on color-image processing", *37th IEEE International Carnahan Conference on Security Technology*, Taiwan, Oct. 2003, pp. 200-207.

[10] N. Ikonomatakis, K. N. Plataniotis, M. Zervakis, and A. N. Venetsanopoulos, "Region growing and region merging image segmentation", *International Conference on Digital Signal Processing*, vol. 1, Jul. 1997, p.299-302.

[11] R. C. Gonzaleez and R. E. Woods, Digital Image Processing, 2 ed., New Jersey: Prentice Hall, 2002.