An Intelligent People-Flow Counting Method for Passing Through a Gate

Thou-Ho (Chao-Ho) Chen, Tsong-Yi Chen and Zhi-Xian Chen
Department of Electronic Engineering
National Kaohsiung University of Applied Sciences
Kaohsiung, Taiwan, R. O. C.
thouho@cc.kuas.edu.tw, chentso@cc.kuas.edu.tw, 1093320119@cc.kuas.edu.tw

***Abstract*—Based on area and color analyses, a cost-effective bi-directional people counter dedicated to the pedestrian flow passing through a gate or a door is proposed. Firstly, the passing people are roughly counted with the area of people projected on an image captured by a zenithal video camera. The moving direction of the pedestrian can be recognized by tracking each people-pattern with an analysis of its HSI histogram. To improve the accuracy of counting, the color vector extracted from the quantized histograms of intensity or hue is introduced to refine the early counting. Besides, the inherent problems of both people touching together and merge/split phenomenon can be overcome. Experimental results show that an 100% accuracy of bi-directional counting can be achieved if the people number of a people-touching pattern is less than six.**

***Keywords*—people count, color image processing, segmentation, tracking**

I. INTRODUCTION

An accurate automatic counting of pedestrian flow through a gate is very attractive for the entry control and access surveillance of the important military, building security and commercial applications. Without losing the generality, the early automatic counting approaches, such as turn stiles, rotary bar, and light beams, had suffered one intractable problem: they could not count the passing people accurately unless there is only one pedestrian through the gate at one time. To solve this problem, many image-processing based approaches with various applications [1-11] are proposed.

For the transportation applications, Bartolini et al. [1] and Albiol et al. [2] addressed the problems of determining the number of people getting into and out of a bus and train, respectively. 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. Though, the system [3] based on template motion-estimation tracking may be very time-consuming due to the computation complexity. Focused on dynamic backgrounds, Zhang and Sexton [5] developed an automatic pedestrian counting method on an escalator or a moving walkway by using a model-specified directional filter to detect object candidate locations followed by a novel matching process to identify the pedestrian head positions in an image even with complicated contents. With the graylevel-based head analysis, the method will suffer from the following situations: a low contrast of the head image with the background and hair styles or wearing various hats for pedestrians. To increase the count of passing people through a gate at one time, Terada et al. [6] 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. Anyway, the setting of the stereo camera is complicated and the measurement will be seriously sensitive to any shift of camera. To avoid limiting the setting position of the camera and counting several times for someone people as they move around, multiple cameras located over the region of interest will be the allowable solution [7, 8]. Based on the cost- effective consideration, a single camera with a tracking algorithm may be the better solution and thus Masoud and Papanikolopoulos [9] developed a rectangular model-based recognition of the pedestrian with a 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. [10] proposed a real-time scheme to detect and track the people moving in various directions with a bounding box enclosing each person.

Although some of the above people-counting methods can solve the problem of real-time counting for the crowded pedestrians. However, those methods haven't dealt with the merge-split case that people walk sometimes touching with one another and sometimes separating from others. To overcome the above problems, the area and color informations based approach is proposed in [11] but some problems of tracking needs to be improved for increasing the counting accuracy, especially when the pedestrian walks fast.

This paper addresses the problem of determining the number of pedestrians passing bi-directionally through a gate (or door). Based on analyses of area and color information in a segmented people-image, the inherent people-touching overlapping problem can be overcome. By a two-stage counting strategy, the area of pedestrian image segmented form the background is firstly used to count the passing people and then its color information is utilized for refining the initial count. The following section will describe the proposed people-counting system including the system setting and counting algorithm. Then, experimental results and analyses

1-4244-0025-2/06/$20.00 ©2006 IEEE RAM 2006

will be discussed in Section III and conclusions will be made in the final section.

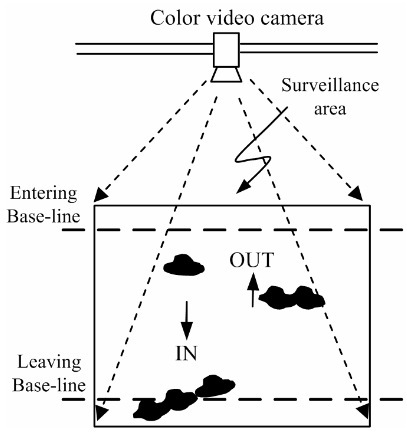

Fig. 1. Setting of the proposed counting system.

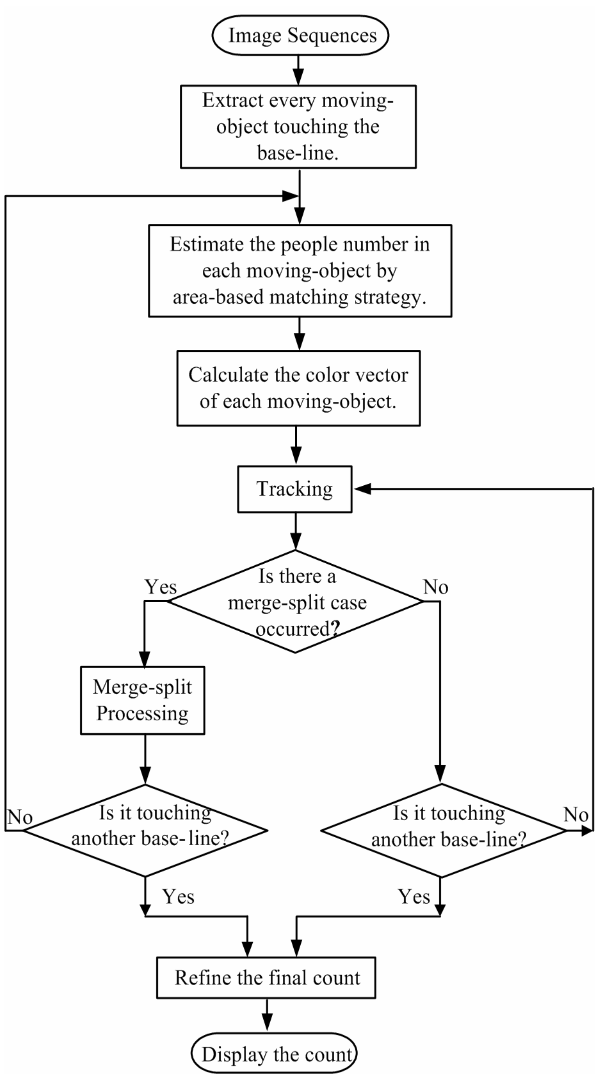

Fig. 2. The proposed counting algorithm for the pedestrians passing through a gate.

II. THE PROPOSED PEOPLE-COUNTING SYSTEM

A. Conting Algorithm

In the Proposed scheme, one color video camera is set on the ceiling of the gate with a directly downward view so that the passing people will be observed just overhead, as shown in Fig. 1. To achieve an automatic bi-directional counting for the pedestrian passing through a gate (or door), the proposed algorithm is described in Fig. 2. Based on the motion analysis, the background subtract method is applied to extract pedestrian pattern from an image. If there is a higher flow-rate of pedestrians, one view captured by the camera may contain several people-images, where each people-image may consist of one or more persons.

After obscuring each moving object, the quantity of people existed in a moving object is firstly estimated through an area-based strategy followed by a chromatic analysis to track the first count. In the tracking process, the merge-split case may be encountered, where the merge-split phenomenon means that one people-image is split into multiple people-images and/or multiple people-images are merged into one. Based on the HSI histogram analysis, the distinguishing vector, also called “color vector”, is extracted for recognizing each moving object against the merge-split problem. It should be pointed out that in practice, the merge-split problem usually happens because the pedestrian’s body may touch with one another when walking and thus it will incur a merging case in an image captured. Also, a splitting case will appear when someone is leaving from a moving multi-people heap.

B. Area-Based People Counting

In the proposed counting scheme, the area of people-image pattern is considered to be able to provide a good first estimation of how many people being contained in that pattern. Furtherly, it can support informations for tracking and discriminating between multi-person and single-person patterns when encountering a merge-split case.

At first, the typical area size of a single-person image is defined as N_P pixels per people, where the typical value means a minimal. Owing to clothing or wearing, a threshold value of allowable maximal quantity of pixels for a single-person image is given as 1.4 times of N_P. Practically, the area of n-person image will be smaller than total area of n single-person images due to the overlap resulted form people-touching overlapping situation mentioned previously. From the statistical data of experiments [11], the decision rule of how many people being existed in a multi-person image can be deduced as follows:

$$PN_{PI} = 1,\ if \quad N_P \leqq N_{PI} < 1.4N_P \tag{1}$$

$$PN_{PI} = 2,\ if\ 1.4N_P \leqq N_{PI} < 2.6N_P \tag{2}$$

$$PN_{PI} = 3,\ if\ 2.6N_P \leqq N_{PI} < 3.6N_P \tag{3}$$

$$PN_{PI} = 4,\ if\ 3.6N_P \leqq N_{PI} < 4.8N_P \tag{4}$$

$$PN_{PI} = 5,\ if\ 4.8N_P \leqq N_{PI} < 5.8N_P \tag{5}$$

where PN_{PI} and N_{PI} denote the people number and pixel number for a people-image, respectively. However, the above rules are based on the evaluation of the oriental human and

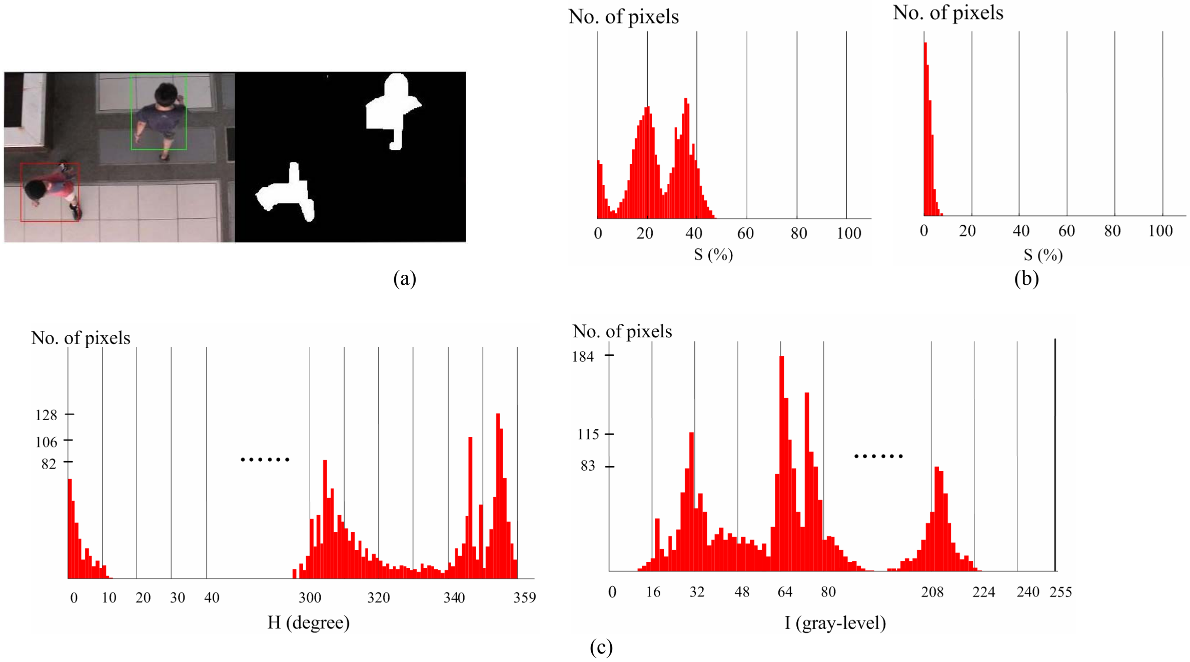

Fig.3. Analysis of two single-pedestrians walking in the opposite direction: (a) the segmented patterns have an area of 2449 pixels for the right one and 1942 pixels for the left one; (b) the S-histograms of the segmented patterns; (c) the I-histogram with a color vector CV_I= (65, 30, 210) for the right pattern and the H-histogram with a color vector CV_H= (354, 346, 304) for the left pattern.

light clothing effective when people walk in a normal way. If the people number in a people-image is more than five, it is not sure that a reliable decision rule can be derived. Although, it is very seldom to see that six or more people walk together in touch with one another. So, the above five conditions are sufficient for most of the general situations of walking together. Fig. 3(a) show the two single-persons (walking in the opposite direction each other) and their segmented patterns. The segmented the two single- pedestrian patterns in Fig. 3(a) have 1942 pixels for the left pattern and 2449 pixels for the right pattern.

C. *Analysis of HSI Histogram*

Among various such color models, HSI color model is very suitable for providing a more people-oriented way of describing the colors, because the hue component is intimately related to the way in which human beings perceive color. For converting colors from RGB to HSI, given an image in RGB color format, the H, S and I components of each RGB pixel can be deduced by the following equations [12]:

$$I = \frac{1}{3}(R+G+B), 0 \le I \le 1 \tag{6}$$

$$S = 1 - \frac{3}{(R+G+B)}\left[\min(R,G,B)\right], \ 0 \le S \le 1 \tag{7}$$

$$H = \begin{cases} \theta \ \ , if \ B \le G \\ 360-\theta, \ if \ B > G \end{cases},$$
$$\theta = \cos^{-1}\left\{ \frac{\frac{1}{2}\left[(R-G)+(R+B)\right]}{\left[(R-G)^2+(R-B)(G-B)\right]^{\frac{1}{2}}} \right\}, \ \ 0° \le H \le 360° \tag{8}$$

Loosely speaking, a histogram can give an estimate of the appearance probability of someone pixel-value in an image. Thus, the histogram of hue or intensity can be used to discriminate one pedestrian from another. At first, suppose that the range of hue, i.e., from 0° to 360° is divided by a slicing factor k into (360/k) slices, where a slice means a k-degree interval of hue. The level of each slice is defined as the hue value which has a maximal pixel-number within that interval (i.e., slice). For numerical representation, let HS_i represent the hue level of the *i-th* slice where i = 1, 2,…, 360/k. To reduce the matching complexity and enhance the discrimination, a moderate quantity of larger-level hue slices are extracted for forming a distinguishing vector, also called “color vector”, denoted by CV. Therefore, for person identification, an n-dimensional color vector composed of n larger hue levels (i.e., HS_i) is used to label each pedestrian. Such a hue-based color vector is formed as

$$CV_H = \max{}_n\left\{HS_i \ , i = 1, 2, ..., \frac{360}{k}\right\} \\ = (HS_{\max-1}, HS_{\max-2}, ..., HS_{\max-n}) \tag{9}$$

where the operation of $\max_n\{\}$ is defined to select n significant values of HS_i, denoted by HS_{max-1}, HS_{max-2},……, HS_{max-n}, in order of pixel-number. Nevertheless, it should be noted that the angle definition of H (hue) will be meaningless if S (saturation) approaches to zero. In such a situation, the I (intensity) component will be introduced to form the color vector, CV_I. By the same deduction of CV_H, we can obtain

$$\begin{aligned} CV_I &= \max{}_n \left\{ IS_i, i = 1,2,\ldots,\frac{256}{k} \right\} \\ &= \left(IS_{\max-1}, IS_{\max-2}, \ldots, IS_{\max-n} \right) \end{aligned} \tag{10}$$

where the range of I (intensity) is supposed from 0 to 255, IS_i denotes the level of the *i-th* slice of the intensity histogram and IS_{max-1}, IS_{max-2},…, IS_{max-n} denote n significant ones of IS_i in order of pixel-number.

When most pixels are distributed on zero saturation or so, the I-histogram, replacing H-histogram, will be analyzed for providing the color vector, as shown in Fig. 3(b) and (c). The right person in Fig. 3(a) has the dark-grayish clothing and hence the segmented pattern adopts the I-histogram to form a three-dimensional color vector CV_I = (65, 30, 210), as illustrated in Fig. 3(c), in which the slice-factor k=16 is used to divide the I-histogram into 16 slices. The left segmented pattern in Fig. 3(a) has a non-zero saturation distribution in the S-histogram, so the chromatic feature is used to form a color vector of CV_H = (354, 346, 304) with a slice-factor k = 10.

In theory, a larger k can reduce the above problem of matching tolerance but it may have a risk of mismatching. Relatively, a small k can provide a more precise person-identification but it will incur a complicated identification. On the other hand, n, the number of elements included in the color vector can also affect the correctness of person identification. According to statistical data of experiments, k = 10 for the quantization of hue histogram, k = 16 for the quantization of intensity histogram and n = 3 for the dimension of color vector are very suitable in the proposed counting system. Besides, a moderate matching tolerance is still necessary for accelerating the matching process to provide an acceptable result.

D. People Tracking

From the chromatic analysis of people's clothing discussed in the above subsection, each people-pattern will be labeled with a color vector for the purpose of tracking. However, the color vector will become ineffective in identification if each pedestrian's clothing has the same color, e.g. a certain type of clothing which is worn by all the members of a group or organization. Hence, a bounding-box is introduced to trail the people-pattern, as illustrated in Fig. 4, in which an intersectional case is described in the subfigure (a) and disjointed case in the subfigure (b). If any two bounding- boxes of the identical people-pattern in consecutive images captured have an intersection, such two people-patterns enclosed with a box in adjacent images (e.g. the current and next frames) will be judged that the both patterns are generated from the same person. On the contrary, the two people-patterns may not represent the identical person if their bounding-boxes have no intersection. It is noted that the value of time interval τ will be adjustable and dependent on the frame-sampling rate.

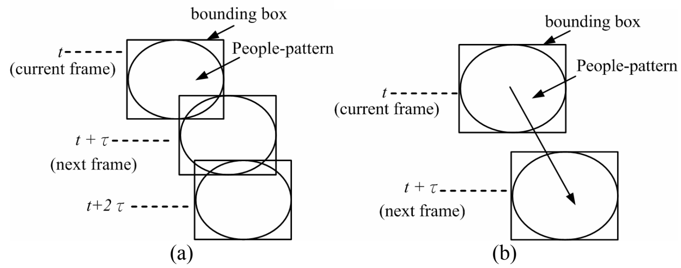

Fig. 4. Trailing of a bounding-box for (a) intersectional case and (b) disjointed case.

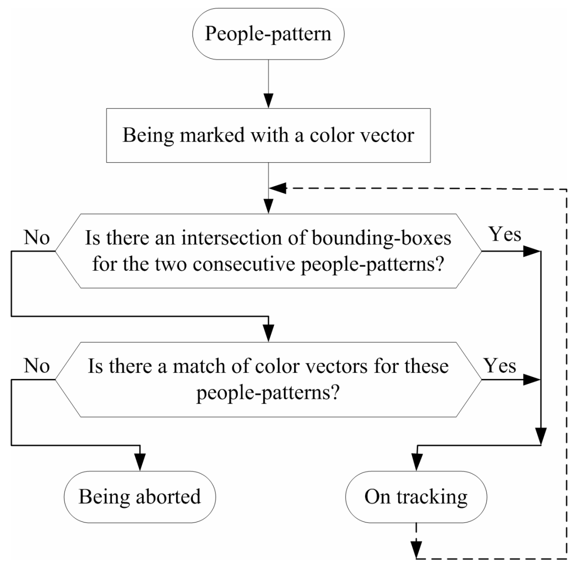

Fig. 5. The tracking procedure.

In respect of tracking, the intersection check of the bounding-box pair for person-identification will last throughout the whole tracking process but the color vector of each people-pattern is sometimes calculated. The color vector is adopted for person-identification only when there is no intersection of the two bounding-boxes located in two successive frames. This can reduce a large amount of calculations required for color vectors. The above discussions about the tracking procedure can be described in Fig. 5.

It can be observed that the counting direction can be oriented by the order of moving across both base-lines or the base-line from which the people go away, during the tracking process. This direction-orienting strategy requires the least computational cost, unlike many previous systems needing additional operations.

E. Merge-Split Processing

In general, people may walk alone or together with their friends and this will result in one single-people pattern, several single-people patterns, one multi-people pattern, several multi-people patterns, or their mixed patterns in a surveillance area captured. As some reasons mentioned previously, each moving object may happen to be split or merged within a range of the captured surveillance area and this merge-split problem may

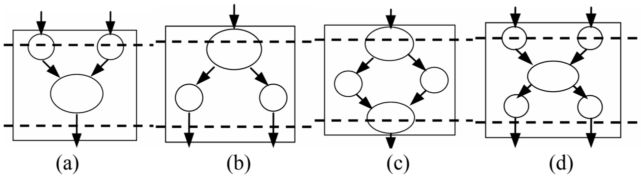

Fig. 6. Four basic cases of merge-split: (a) 2-to-1 merging case; (b) 1-to-2 splitting case; (c) 1-to-2-to-1 split-merge case; (d) 2-to-1-to-2

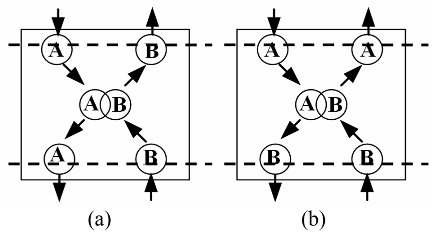

Fig. 7. Inverse-directionauching touching situations. (a) Normal case.(b) Abnormal case.

confuse the counting. Basically, the merge-split phenomenon can be classified into four cases, as shown in Fig. 6. A 2-to-1 merging case of Fig. 6(a) means that two moving patterns are merged into a single moving pattern, where each moving pattern may be composed of one or more persons. By the same deduction, 1-to-2 splitting case of Fig. 6(b), 1-to-2-to-1 split-merge case of Fig. 6(c), and 2-to-1-to-2 merge-split case of Fig. 6(d) mean that one moving pattern is split into two partial ones, one moving pattern is firstly split into two partial ones and finally merged into a large one, and two small moving patterns are firstly merged into a large one and finally split into two small ones, respectively.

From the above analysis for merge-split, to check if there is a merging or splitting case to happen, the area change of the moving pattern can give a fundamental judgment. A merging case will be detected if two separate moving patterns move in the current frame and then are combined into a larger-area pattern in the next frame. Thus, the people number of the new merged pattern needs to be calculated according to those area-based decision rules mentioned above in order to refine the initial count. Besides, the color vector also requires to be calculated for labeling the new moving object in the following tracking. Based on the opposite reasoning, a judgment of splitting case can be also deduced. If there is no intersection between two successive bounding-boxes, for both merging and splitting cases, each moving object, can be also identified by use of color vector for the tracking purpose. But if both conditions of no intersection and no match of color vectors are met, it is difficult to distinguish the new generated patterns from their neighboring patterns, i.e., we can't recognize the patterns which form the new patterns. If there is the similarity of color but no intersection of successive bounding-boxes for both moving patterns, the track of each moving object will be confused. In the normal case (Fig. 7(a)), the count is still correct in spite of a lose of track. But, it will make on erroneous count in the abnormal case (Fig. 7(b)). Furthermore, the above cases may be mixed with one another to produce other hybrid cases of merge-split, in which some cases are intractable to handle. It should be pointed out that the situation of splitting and merging of more than two blobs can be done with the single blob described above.

Fig.8. Real setting of the proposed counting system

III. Experimental Result And Analysis

A theoretical analysis about the proposed pedestrian counting system has been given in the above section, but the implementation of such a counting system with the practical sequences can provide a realistic and interesting evaluation. To the realize the proposed pedestrian-counting system, a color video camera (DynaColor D7722) is set 4.2m above the floor, as shown in Fig. 8, and the surveillance area is 320 × 240 pixels (i.e., frame size). The capture rate of the camera is 30 (frames/second) but the processed frames per second are distributed from 10 to 30, depending on the quantity of people-images extracted, on Intel Pentium IV 3.0 GHz.

For the detailed evaluation, five various types of moving objects containing from one-people pattern to five-people pattern with the walking dynamics, such as the moving direction (uni- or bi-direction), moving speed (normal, fast or abrupt, where the threshold between the normal and fast speeds is 100 meters per minute) and merge-split action (basic cases or the mix), are used for the input to the proposed counting algorithm. Such evaluations are tabulated from Table I to Table IV. In these tables, each test is performed with many sequences in which people are clothed in various types and colors. In addition, some situations that people walk with a package, a handbag or knapsack (or rucksack) are also included. From these four tables, it is observed that the accuracy of 100% can be achieved if the walking speed of pedestrians in an uni-direction is normal, i.e., below 100 (meters/minute), with basic merge-split cases, under the assumption that the people number of a single segmented moving-pattern is less than six. Besides, an 100% accuracy of bi-directional counting can be also obtained in the case of one-/two-people patterns, excluding the abrupt moving case, as shown in Table 1. It should be pointed out that 100% accuracy is surely attained if there is no merge-split situation to occur. However, the merge-split phenomenon frequently appears in most of real-word situations while people walk through a gate.

As the people number contained in a moving pattern increases above two, some situations of fast moving will make the counting accuracy to be slightly decreased, especially for a mix of fast bi-directional merge-split cases, as shown in Table

TABLE I. Evaluation for counting a crowd composed of one- and/or two-people patterns.

Test	Direction	Merge-Split	Speed	No. of People	Count	Accuracy (%)
1	uni-	Basic case	normal	132	132	100
2			fast	116	116	100
3			abrupt	108	105	97.2
4	bi-	mix	normal	123	123	100
5			fast	110	110	100
6			abrupt	102	96	94.1

TABLE II. Evaluation for counting a crowd containing the three-people pattern.

Test	Direction	Merge-Split	Speed	No. of People	Count	Accuracy (%)
1	uni-	Basic case	normal	126	126	100
2			fast	115	112	97.3
3	bi-	mix	normal	138	131	94.9
4			fast	104	95	91.3

TABLE III. Evaluation for counting a crowd containing the four-pattern.

Test	Direction	Merge-Split	Speed	No. of People	Count	Accuracy (%)
1	uni-	Basic case	normal	96	96	100
2			fast	95	90	94.7
3	bi-	mix	normal	90	83	92.2
4			fast	88	78	88.6

TABLE IV. Evaluation for counting a crowd containing the five-pattern

Test	Direction	Merge-Split	Speed	No. of People	Count	Accuracy (%)
1	uni-	Basic case	normal	94	94	100
2			fast	92	86	93.4
3	bi-	mix	normal	84	76	90.4
4			fast	84	72	85.7

II, III and IV. The greater the people number of a moving pattern is, the more the counting accuracy is reduced. For the fast uni-directional case, sometimes we can't recognize which people-patterns forming the new pattern after merging or splitting, due to similar clothing in chromaticity, and thus the count may be confused. In regard to the bi-directional case, the inverse-directional touching is the major reason of count-accuracy reduction. Basically, the inverse-directional touching, as shown in Fig. 8, will slightly affect the correctness of counting when the people quantity of a moving pattern is increased substantially. For this reason, the proposed algorithm can't correctly count in a case that two moving patterns touch with one another and then are merged into a people-pattern composed of more than five persons, as revealed in the bi-directional cases of Table II, III and IV. Though the counting accuracy of a mixed fast merge-split case is reduced even below 90%, as indicated in Table IV, such a situation is seldom to happen. It is noted that, generally speaking, the abrupt case will hardly occur in a situation that the people number of a fast moving crowd is more than two, except for an intentional behavior.

From an overall analysis on the above experimental results, it implies that a people-pattern with a fast moving speed seems to be difficult for tracking and it is the most significant reason of confusing the counting. In addition, a mix of bi-directional merge-split cases, in which people wear the similar-color clothes, will puzzle the tracking for each person, while the count may be correct. The situation of abrupt moving may also interrupt the tracking process. There are still other factors which influence the counting accuracy, such as a crowded situation that the segmented people-pattern is composed of more than five persons, and some intentional actions. However, the tracking ability can be improved by a complicated template-based approach if a higher implementation cost is considered.

IV. CONCLUSIONS

This paper presents an automatic bi-directional people-counting method dedicated to passing through a gate or door. An area-based estimation is adopted for deriving an early count of pedestrians. Then, the color information of the pedestrian's clothing or wear is used to label each person followed by a bounding-box intersection based strategy, to track the people. Hence, two important inherent problems, including that people walk touching with each other and the merge-split situation, can be overcome. Besides, the accuracy of the proposed counting algorithm can be further improved by increasing the frame-rate processed. Based on the experimental results, it has been shown that this research will provide a cost-effective people counting technique and it will be more attractive than others.

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, pp. 36-41, Jan. 1994.

[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, pp. 204-218, Dec. 2001.

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

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

[5] X. Zhang and G. Sexton, "Automatic human head location for pedestrian counting," IEE 6th International Conference on Image Processing and its Applications, pp. 535-540, 1997.

[6] 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, pp. 338-342, 1999.

[7] L. Khoudour, J. P. Deparis and L. Durvieubourg, "Linear image sequence analysis for passengers counting in public transport," IEE International Conference on Public Transport Electronic Systems, pp. 100-104, 1996.

[8] V. Kettnaker and R. Zabih, "Counting people from multiple cameras, IEEE International Conference on Multimedia Computing and Systems," vol. 2, pp. 267-271, 1999.

[9] 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, pp. 1267-1278, Sep. 2001.

[10] 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, pp. 1416-1419, 2002.

[11] 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, pp. 200-207, Oct. 2003.

[12] R. C. Gonzalez and R. E. Wood, Digital Image Processing, 2Ed., New Jersey: Prentice Hall, 2002.