Image Processing Techniques to Determine Bicycle Flow

Cao Li，Liu Jie，Jia HaiTao

（Dept.of Automation, Tsinghua University, Beijing, 100084）

（ E-MAIL: caoli@tsinghua.edu.cn）

Abstract - This paper describes parameter detection of bicycle flow based on video processing. The binary 'Temporal-Spatial image' (TS image) is used for counting the number of people. Skin-color model is introduced to segment the connected blocks in TS image precisely. And a shadow suppression algorithm based on the HSV color space is used to reduce the connected blocks. Finally, experiments with some video samples under different circumstances are carried out. The results are evaluated and discussed.

KeyWords - Bicycle flow, counter, skin-color, shadow suppression

1 Introduction

Bicycles and pedestrians are important factors for transportation ability evaluation in China. Video-based detection system has been developed quickly for real-time traffic data collection. [1] shows one popular image processing technique, which locates small detection areas in the captured image. When moving vehicles pass, the variation in these areas between frames can be utilized to detect velocity, flow and vehicle length. The other common method shown in [2] is to define and search the feature points in moving objects. Then, these points can be used for tracking or counting. However, these researches normally focus on vehicles and have little relationship with bicycles and pedestrians. One research about bicycle flow is in [3], which used Hough Transformation to detect the wheel of the bicycle from the side-view. The other related research is in [4], which used a camera vertically above the persons to count the number.

This paper shows a system that outputs the number of the people in one bicycle path by capturing the video image as shown in Fig.1, from a stationary camera. It aims on determining the usage of one bicycle path and giving the bicycle flow volume to the next intersection. Some image processing technologies are used in this paper including morphology operation, skin-color detection and shadow suppression. Results of several experiments with outdoor scenes are given.

Fig.1 Image of bicycle flow and detection line location

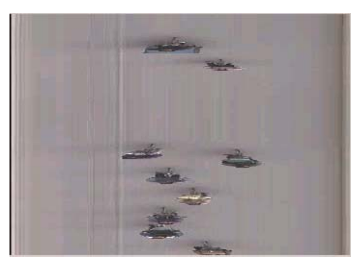


Fig.2 One TS color image

2 Temporal-Spatial Image

In highway, cars and other vehicles travel on regular paths. However, bicycles are not the same and they share one bicycle path, as shown in Fig.1. In this paper, the detection area is the single line shown in Fig.1.

Pixels in this line of every frame are what we will process.

For every frame of image sequence, the pixels on the detection line are recorded. Then we display these recorded lines of pixels from top to bottom in the sequence of time. Thus the temporal-spatial image (TS image) is generated. The horizontal axis represents the spatial information of the image and the vertical axis represents the temporal information of the image sequence. Fig.2 is one example of TS color image.

In Fig.2, the positions of moving persons passing across the detection line and the width of every moving person can be found. Moreover, the time that every moving person spent on passing the line is also can be seen. From Fig.2, we already could count the number of people approximately from those distinct parts on this image.

As the steps to get the counting results, we change the TS color image to TS binary image. For every pixel of the detection area, we use (1) to get the intensity from RGB. Because the camera is static, temporal intensity changes can be related mainly to motion. We specify the detection line as the background and subtract this background from every detection line in every frame of the image sequence. By giving a threshold of this difference, we get the TS binary image, as shown in Fig.3.

$$Y = 0.299 * R + 0.587 * G + 0.144 * B \quad (1)$$

Fig.3 Difference image and binary TS image

From Fig.3, we can see the original binary image has some noises. We prefilter the original by using median filter of size 3. After this step, the binary image can be used for the next processing.

If only a little number of moving persons passed across the line, that is to say, in the binary TS image, one 1-pixel block represents only one bicycle. Under this situation, we just need to count the number of 1-pixel blocks directly. While under the traffic congestion, different 1-pixel blocks may connect with each other in the binary TS image. Hence segmentation methods should be applied when in traffic congestion for precise counting results.

3 Horizontal Segmentation

Fig.4 shows several kinds of blocks in the binary TS image. Block No.1 usually represents one person riding a bicycle passing across the detection line at a normal speed. Block No.2 represents two bicycles passing side by side. Block No.3 represents two bicycles passing one after another. Block No.4 represents one person passing the detection line at a lower speed and Block No.5 represents one moving object (wider than one single person) passing across.

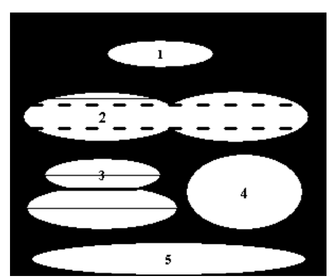


Fig. 4 Some common blocks in binary TS image

We have used an algorithm based on width function of TS binary image to do the horizontal segmentation. That is to segment blocks that are like block No.2 and not to segment blocks that are like block No.5. In our application, much in the same way as [5], we try to find two special rows. These two special rows must exist synchronously, shown with the two dashed line shown in the block No.2, and the absolute value of the width gradient of 1-pixel block on these two rows must above one certain threshold.

Once we find these two lines, we can finish the horizontal segmentation. Fig.5 shows an example. The segmentation result is shown in Fig.5 (b), corresponding to the connection shown in (a).

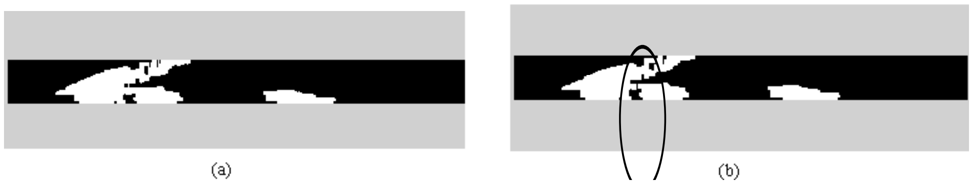


Fig.5 Horizontal segmentation result

We also tested the vertical segmentation method described in [5] in our application, to find maxima and minima of width in block No.3. But unfortunately, it is not suitable, because our camera is far from the bicycles, thus the 1-pixel blocks in binary TS images are relatively small. We use skin-color pixels as the feature points to detect the people and to do the vertical segmentation.

4 Skin-Color Detection for Vertical Segmentation

4.1 Skin-Color Model

Skin-color is an useful cue for locating and tracking faces [6]. In our system, this feature is used for finding persons in the bicycle path. The camera is fixed above the bicycle path and people faces can be seen.

We used the skin-color model described in [7]. The RGB representation of the image sequence obtained from camera is converted to 2D representation.

$$r = R/(R+G+B) \quad , \qquad g = G/(R+G+B) \qquad (2), \ (3)$$

Fig.6(a) is a picture of human face and it is used for computing (r, g) values of every pixel. The distributions of these pixels in (R, G, B) and in (r, g) are also shown in Fig.6(b) and (c).

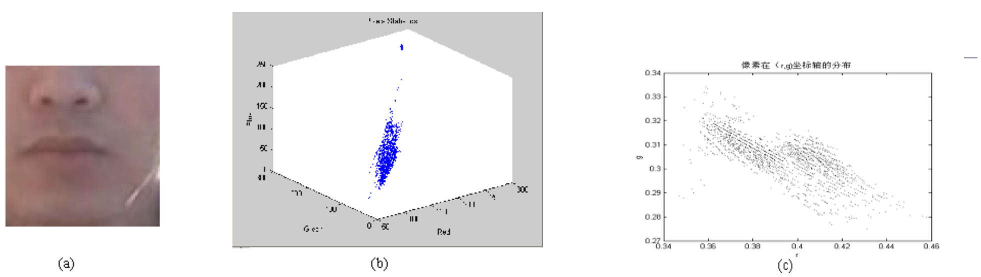


Fig.6 Skin-color distribution in (R,G, B) and (r,g)

The skin-color distribution in (r, g) can be described approximately as an independent Gaussian model $N(\mu,\Sigma^2)$, where $\mu = (\bar{r},\bar{g})$, $\bar{r} = \frac{1}{N}\sum_{i=1}^{N} r_i$, $\sigma_{rr} = \frac{1}{N}(r_i - \bar{r})^2$, $\bar{g} = \frac{1}{N}\sum_{i=1}^{N} g_i$, $\sigma_{gg} = \frac{1}{N}(g_i - \bar{g})^2$.

The parameters including the above mean and variance of r and g, need to be estimated. And this model is used to detect skin-color pixels in RGB images by giving a threshold value of (4).

$$P(r,g) = P(r)P(g) = e^{-\frac{(r_i-\bar{r})^2}{2\sigma_r^2}} e^{-\frac{(g_i-\bar{g})^2}{2\sigma_g^2}} \qquad (4)$$

Skin-color is easy to be affected by the outdoor sunshine. We captured one skin-color image every hour from 8:00 am to 6:00 pm. and computed corresponding parameters of the temporal skin-color model. Fig.7 shows one statistical figure of $\bar{r}$ with solid line and $\bar{g}$ with dotted line.

In Fig.7, these two parameters did not change drastically except on 18:00, when the sunshine was very weak. Fig.8 shows the results that we used skin-color model parameters of 8:00am to detect skin-color pixels in the images of other moments. All skin-color pixels were almost detected by this detection model of 8:00am. This result can prove that skin-color model parameters vary little with the variation of outdoor sunshine.

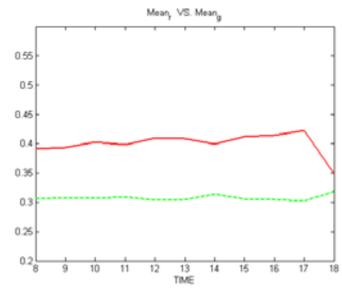
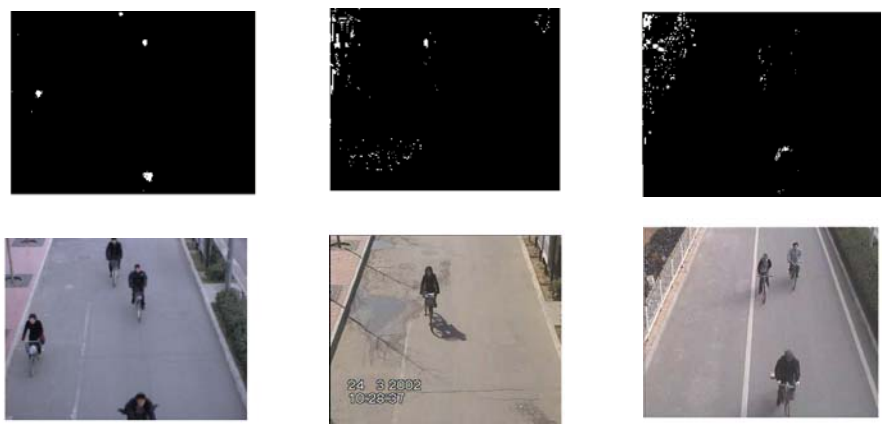


Fig.7 Mean values of r and g at different day time



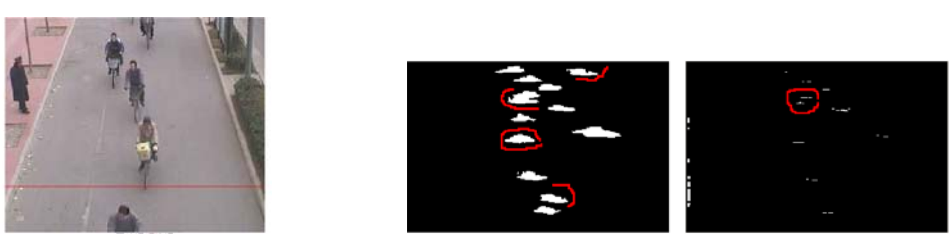
(a) at 16:00 (b) at 11:00 with shadow (C) at 9:00

Fig. 8 Skin-color detected images at different day time

4.2 Vertical Segmentation

By detecting skin-color pixels in detection line in every frame of the image sequence, we can get another binary TS image, which is skin-color detected TS image shown in Fig.9 (c), corresponding to the binary TS image of difference image shown in Fig.9(b).

When a 1-pixel block in binary TS image is longer in vertical axis than a certain distance, we examine the corresponding region on the skin-color detected TS image and count the number of skin-color blocks. If more than one skin-color blocks exist and the distance between them is above one certain threshold, this block is considered as representing more than one persons and it will be segmented. This can be seen in Fig.9 (b) and Fig.9 (c).



(a) Original video (b) Binary TS image (c) Skin-color detected TS image

Fig. 9 Skin-color detection combined for vertical segmentation

5 Shadow Suppression

Under the shadow condition, deformations will appear in the binary TS image compared with non-shadow condition, which may bring mistakes into the final counting results.

The shadow suppression method in [8], which is based on HSV color space, is used in our application. We choose one shadow region and record the pixels as I (x, y). When the moving shadow disappears in the region, we record this background region as B (x, y). For every pixel, we compute v, s and h based on (5)-(7).

$$v = \frac{I_k^V(x,y)}{B_k^V(x,y)}, \quad s = I_k^S(x,y) - B_k^S(x,y), \quad h = \left| I_k^H(x,y) - B_k^H(x,y) \right| \qquad (5), (6), (7)$$

By computing the mean and variation of v, s, h, we can get four parameters for shadow pixels recognition, as shown below.

$$\alpha = Mean(v) - 3 * \sigma_v, \qquad \beta = Mean(v) + 3 * \sigma_v \qquad (8), (9)$$

$$\tau_s = Mean(s) + 3 * \sigma_s, \qquad \tau_h = Mean(h) + 3 * \sigma_h \qquad (10), (11)$$

Then (12) is used for judging whether it is a shadow pixel.

$$\begin{cases} shadow & (\alpha \le v \le \beta) \,\&\, (s \le \tau_S) \,\&\, (h \le \tau_H) \\ non-shadow & otherwise \end{cases} \qquad (12)$$

Fig.10(a) shows the binary TS image before shadow suppression. Fig.10 (b) is the result that after the shadow suppression. Median filter is used to remove those isolated points. Then, we use one open operation to extract the thicker parts from the 1-pixel blocks. Those thicker parts normally are the reflections of real moving persons. Fig.10(c) is the final result and from which we count the number of moving persons.

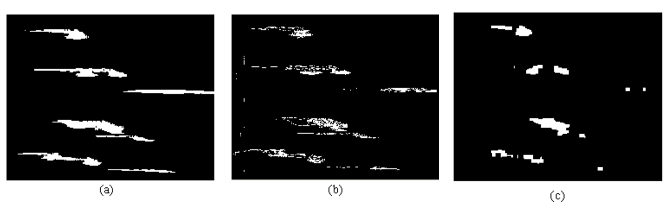


Fig.10 Shadow suppression results

6 Experiment Result

Experiments have been carried out using several image sequences. When moving persons pass across the detection line, we record the binary TS image and skin-color detected TS image. Once there is no any moving person on the detection line, the two TS images will be processed immediately and synchronously. Then the final counting results can be obtained from the fusion of these two images.

Table1 shows the experiment results. Sample1 to sample4 are under the non-shadow condition. Sample 5 and Sample 6 are under the shadow condition. The error shown in Table1 is about 10%. However, the error is much more under the shadow condition than non-shadow condition.

Fig.11 and Fig.12 show some scenes that may bring counting mistakes in our experiment, including the cases of high people density, car passing, object color near to skin-color, object passing across the line at a faster speed, two persons passing side by side closely and so on.

Table 1 Experiment results

	Sample1	Sample2	Sample3	Sample4	Sample5 (Shadow)	Sample6 (Shadow)
Real Number	96	115	81	48	46	112
Counted Result	88	110	77	45	41	90
Error (%)	-8.33	-4.35	-4.94	-6.25	-10.8	-19.8

Fig.11 Some scenes without shadow

Fig.12 Some scenes with shadow

7 Conclusion and Future Work

One integrated method to do people counting in bicycle flow is proposed in this system. The binary TS image is used to segment horizontally and the skin-color detected TS image is used to do the vertical segmentation. Some acceptable results are obtained. However, there still have a certain number of false counting, especially in shadow conditions. And the parameters used in the shadow suppression algorithm were chosen not automatically but manually now.

References

[1] Zhigang Zhu, Guangyou Xu, etc “A real-time vision system for automatic traffic monitoring”, Image and Vision Computing, 18 (2000). P 781-794

[2] Benjamin.etc “A real-time computer vision system for vehicle tracking and traffic surveillance”, Transportation Research Part C 6 (1998). P 271-288

[3] Scott.Rogers etc“Counting Bicycles Using Computer Vision”, 2000 IEEE ITS Conference Proceedings, P 33-38

[4] M.Rossi, A. Bozzoli “Tracking And Counting Moving People”, Proc.2nd IEEE Int. Conf. Image Processing, pp.212-216, Nov.1994

[5]. Antonio Albiol, Inmaculada, Mora and Valery Naranjo, "Real-Time High Density People Counter Using Morphological Tools", IEEE Transactions on ITS, Vol.2, No.4, December 2001: 204-218

[6] Darrell, T.; Gordon, G.; Harville, M.; Woodfill, J “Integrated person tracking using stereo, color, and pattern detection ”Computer Vision and Pattern Recognition, 1998 Proceedings 1998 IEEE Computer Society Conference on, 1998 Page(s): 601 -608

[7] Jie Yang, Alex Waibel “A Real-Time Face Tracker” Applications of Computer Vision, 1996 WACV '96, Proceedings 3rd IEEE Workshop on, 1996 Page(s): 142 –147

[8] R. Cucchiara, C.Grana, M. Piccardi, and A. Prati, “Detecting objects, shadows and ghosts in video streams by exploiting color and motion information”, Proceedings of the IEEE International Conference on Image Analysis and Processing, 2001. 360-365

中文摘要 - 本文研究了基于视频图像处理的自行车流量检测方法，主要方法为利用二值化的“时空图像”计算 1 像素块儿的个数。对于拥挤的自行车流，为提高 1 像素块儿分割以及行人计数的准确率，提出了结合皮肤颜色检测的方法。另外，对于阴影可能造成的误计数结果，采用了基于 HSV 空

间的阴影去除算法来减小误差。文中给出了各种方法对自行车流量检测的实验结果。

关键词 - 自行车流，人数统计，皮肤颜色，阴影去除