Vehicle Detection in Video Based on the Framework of Kernel Density Estimation

Dan Yang[1], Yantao Chen[1], Richen Liu[2]
[1]Institute of Image and Graphics, Sichuan University , Chengdu, China
[2]College of Computer Science, Sichuan University, Chengdu, China
yangdanjk@126.com, chenyantao@scu.edu.cn, liurichen@gmail.com,

***Abstract*— This paper introduced a method of background subtraction model for detecting vehicles in video, the model was based on the framework of the kernel density estimation. To compute the adjacent frame subtraction on the spatio-temporal by the median filter, when the median value is 0, keep the background pixel value of the point stable, and still use the pixel value of the previous frame of the background to speed up the processing, when the value is not 0, using kernel density estimation to compute the probability, and determine whether update the background pixel value or not. Using histogram to show the statistics of the probability distribution, and find out more accurate threshold of the background adaptively. To solve the problem of deadlock, update the background automatically in a certain period of time. The experiments show that this method can segment the moving vehicles rapidly and accurately from the video.**

***Keywords*— Background model; Vehicle detection; Kernel density estimation; Intelligent transportation system**

I. INTRODUCTION

With the popularity of cars, many cities faced the problems that the degradation of the traffic environment, traffic jams and accidents. The intelligent traffic system has been introduced in the research of image-based intelligent traffic parameters detection model. In [1] the intelligent traffic system provides more effectively analytical tools, and it will be applied widely. This paper focuses on the first stage of intelligent transportation-moving object detecting, it aims to detect vehicle traces from the vehicle video.

The main contents are as follows: The second part elaborates the three main vehicle video detection methods; the third part introduces the algorithm, presents the steps of the method; the fourth part presents the experiment results and analysis process; the fifth part concludes the paper and presents the inadequacies of this article objectively.

II. THREE MAIN METHOD IN VEHICLE DETECTION

A. Frame Difference

In [2, 3] fame difference method is based on the motion of two adjacent frames in video sequence, they have a strong correlation. Subtracting the corresponding pixel value of the two or three consecutive images, when the illumination of the environmental changed slightly, if the difference of the corresponding pixel value is small, we confirm that it was a static pixel, and marked it for the background. If the difference is greater, that it was caused by the motion of moving object, then marked it as moving object pixels.

B. Optical Flow

In [4, 5] optical flow method detect moving object by calculating the optical flow of the image. Its basic idea is: the movement can be described by moving flow, and in an image plane, the motion of the object is often embodied by the change of the different pixel value distribution in the image sequence. Thus, the moving flow in the space can transfer into the image to denote the optical flow field. Optical flow field reflects the change trend of the pixel value at each point in image.

C. Background Subtraction

In [6] background subtraction method use the subtraction of the current frame and background image subtraction in the sequence of images to detecting moving objects. The pixel is considered to be moving object if the difference is large, otherwise, consider it as background.

The comparison of the three methods is shown in “Table 1.” For the detection of vehicles while camera was fixed, we can use the static information of the background to detect vehicles. For background modeling, commonly used methods are: average background in [7], non-parameter model based on kernel density estimation in [8], the background model based on Gaussian distribution in [9, 10]. Non-parameter model did not make any assumption of the background model. It used the previous pixel values obtained to estimate the probability of the current pixel as the background. This paper used the non-parametric model for background subtraction based on the kernel density estimation to detect vehicles in the video.

TABLE I. Com parison of the Three Methods

Characteristic	Detection Method		
	Frame difference	***Optical flow***	***Background subtraction***
Advantages	Speed, not sensitive to illumination	Can be used while camera is not static	Easy to achieve
Disadvantages	Can not detect static objects, resulting in cavity	Computation complexity, time-consuming	Can not run adaptively well

III. ALGORITHM DESCRIPTION

This paper is based on the framework of the kernel density estimation in [8], the algorithm of 3.1 specific seen in [8]. To compute the adjacent frame subtraction on the spatio-temporal by the median filter, dealing with the value 0, and solved the problem of deadlock. The flowchart of the system is shown in "Figure 1."

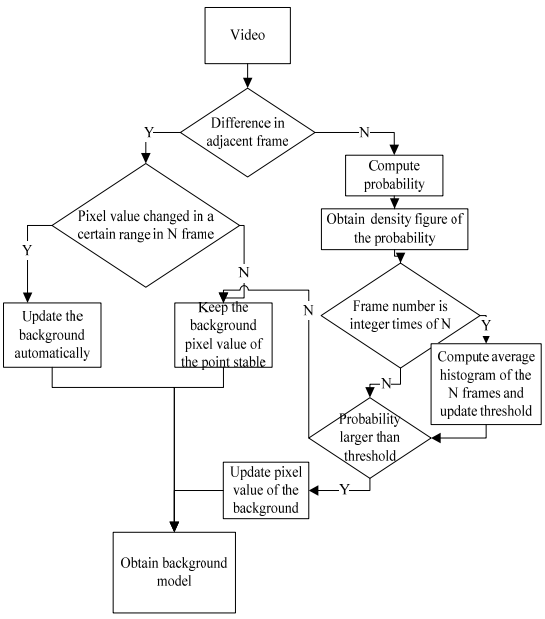


Figure 1. Flow chart of the system processing

A. Kernel Density Estimation

Suppose the video frame width W, height H, then the number of pixels in video frames are W*H. Selecting N frames as the samples of the background, then each pixel in the image has N samples. At time t, the value of the pixel is x_t, x_1, x_2, …, x_N are the samples of the point, its probability density function can be estimated using the following kernel function K without parameters.

$$Pr(x_t) = \frac{1}{N}\sum_{i=1}^{N} K(x_t - x_i) \tag{1}$$

As the value of pixels is Gaussian distribution, so choose Gaussian function as kernel function. The probability density function can be expressed as:

$$Pr(x_t) = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma^2} e^{-\frac{(x_t - x_i)^2}{2\sigma^2}} \tag{2}$$

Where N is number of the samples, x_i is the value of the pixels i, σ is the bandwidth of the kernel function which can be calculated by the following formula:

$$\sigma = \frac{m}{0.68\sqrt{2}} \tag{3}$$

Where m is the median, of each consecutive pair in the samples, compute the absolute value of the difference.

If the probability of the formula (2) meets:

$$Pr(x_t) > T \tag{4}$$

Then the pixel is considered to be a background point, T is the global threshold.

B. Compute m by Median Filter in Spatio-temporal

To compute the adjacent frame subtraction on the spatio-temporal by the median filter and obtained the median, with the value of m obtained to calculate the kernel bandwidth in formula (2). For each pair (x_i, x_{i+1}), there is $|x_i - x_{i+1}|$, sort all samples, m is the median.

When m is 0, the kernel bandwidth is 0, we can not use the kernel density to estimate the probability of the point. In this case, the pixel value was not updated in more than half of the samples, so, it is either the background point or moving object point which does not appear in most of the frame in sample point. At this point, we do not ensure that it is part of the background or the occasional movement object, so, we do not update the point of the background.

$$m \begin{cases} = 0 & \texttt{not update} \\ \neq 0 & \texttt{compute } Pr \end{cases} \tag{5}$$

If m = 0, we do not update the background pixel value, or calculate the probability Pr, and then determine whether to update the background pixel value or not.

In order to observe the distribution of the median, we transform the median value m of each pixel in the current image into binary image, while m is 0, the pixel value in binary image obtained is 255, otherwise 0.

$$MImg = \begin{cases} 255 & m = 0 \\ 0 & m \neq 0 \end{cases} \tag{6}$$

The binary image obtained shown in Figure 2, we can see for a frame in vehicle video, in the spatio-temporal, the median is 0 in most point of the image. We can see, by using formula (6) can accelerate the processing speed greatly.

Figure 2. The binary distribution of the median m

C. Adaptive Threshold

When illumination or environment changes, a fixed global threshold will make error in the pixel of background model, the adaptive threshold can be more reasonable to obtain background model. The probability density of each pixel in current image is mapped to 0 to 255, mapping as follows:

$$PrImg = \frac{255 - 0}{MaxPr - MinPr}(Pr - MinPr) \tag{7}$$

When m is 0, according to the formula (5), the probability density does not exists, it is either the background point or moving object point which does not appear in most of the frame in sample point. We set the pixel value to be 255 in the probability density image. The probability density map shown in Figure 3.

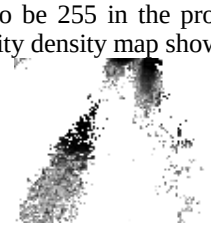


Figure 3. Probability density

Probability density reflects the change of the pixel value, if that point changed frequently in a period of time, its probability value of the point is small, There is more possibility to be moving objects, if changes less, or not changed , then the probability will be larger, the possibility to be background is large. Probability density map is a form of the probability distribution, then the threshold *T* can be transformed into the problem of the division of probability density map.

Each frame has a probability density map, and each probability density map corresponds to a histogram, although divide the probability histogram for each frame will make the background model more accurate, but processing like this will cause large computation. We can average the probability histogram of consecutive N frames. We use the threshold obtained as a global threshold value to process the next N frames, and update the threshold continuously in this way.

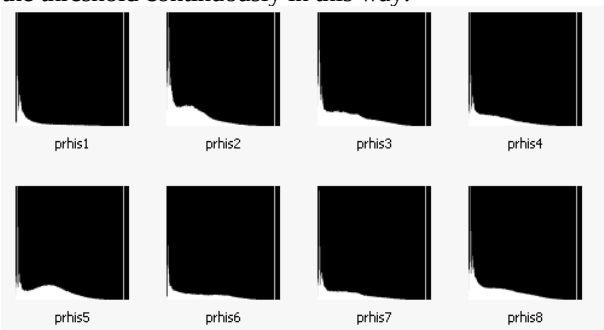


Figure 4. The average histogram of the probability

Figure 4 is the average probability histogram for each N frames, its size is 270 * 256. All the average histograms of the probability density histogram have a similar distribution. The probability is small while gray value is small in the left part of the histogram, it is the moving object, and it is background pixel while the gray value is large, thus, we can make use of the probability histogram to find a suitable threshold to determine whether the pixel is belong to the background. The value of most pixels in the histogram is 255. In this case, the probability density of the point dose not exist, we do not need to update the pixel value of the background. Next is the part that its gray value is very small, it is belong to the moving object.

There is an obvious peak in the histogram when the gray value is small in the probability density map, we can divide the crest to obtain the threshold pixel value, then use it to compute formula (7), obtain the probability, and use the probability as the threshold T of formula (4).

D. Deadlock Avoidance in Background Update

When the background model was built, we need to update it constantly, so make it close to the true background. If the background model did not update in a long period of time, it can lead to inaccurate detecting results.

When m is 0, the operation we taken is not update the pixel value in the background. If the speed of vehicle is too slow or there is a car in background when initialization, then it will cause the car to be a part of background, the car still exists in the area of the background model even if the car left the background area. Using background subtraction method, subtracting the current frame and background model, then transform it into binary image, the car will still exist in the foreground image, resulting in detecting error. And if the environment and illumination do not change, and no vehicle pass through the region, so the region will never be updated, resulting in deadlock. If the pixel value of the background changes within a certain range in certain frames, then make it update mandatory to avoid deadlock. In the spatio-temporal, at the point m = 0, we take the following actions to the background model:

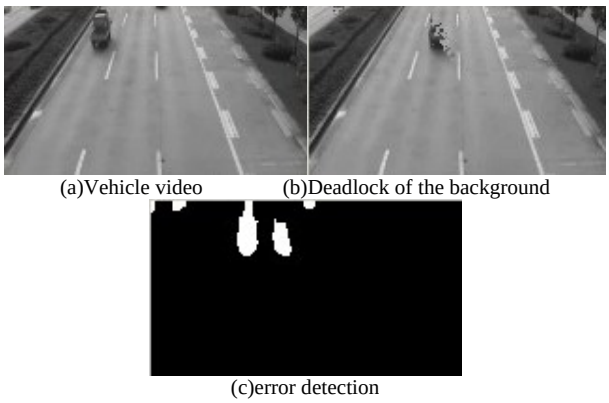
$$\left(x_i \in \left(x_1, x_2, \cdots, x_n\right)\right)\begin{cases} \left|x_i - x_t\right| < T_1 & \text{update} \\ else & \text{not update} \end{cases} \quad (8)$$

There is random noise exists because of camera, in N frames, if the absolute value of difference between the any pixel of background and the current background is less than T_1, then the current pixel value will be assigned to the background model, or do not update the current background pixel value.

IV. Experimental Results and Analysis

The experiment equipment is Intel Core 2 T8100 / 2.1GHz, Memory 2GB. Select the test samples of video data from a section of the highway video surveillance, the video frame width 160, height 90, background samples were 60, dealing with frame rate greater than 15 fps.

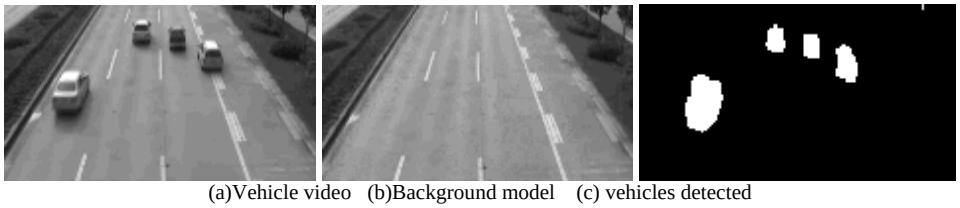
Figure 5 are screenshots of the 3.1 kernel density estimation algorithm experiments. First obtain a video frame 5 (a), and then create the background model, screenshot in Figure 5 (b), Using background subtraction method, subtracting 5 (a) and 5 (b), then transform it into binary image, and obtained foreground image, Screenshots as 5 (c). In this set of data, deadlock happened and the background did not updated, resulting in detection errors.



(a)Vehicle video (b)Deadlock of the background

(c)error detection

Figure 5. Vehicle detection in video when deadlock



(a)Vehicle video (b)Background model (c) vehicles detected

Figure 6. Vehicle detection in video

Figure 6 displays the experiment shots from adaptive threshold algorithm and deadlock avoidance test, get clear background, by which accurately detects the video of the vehicle.

V. CONCLUSION

This paper introduced a method of background subtraction model for detecting vehicles in video. The model was based on the framework of the kernel density estimation. While the median value is 0, keep the background pixel value of the point stable, and speeded up the processing. Estimate a suitable threshold by the probability density maps to determine whether the pixel locals in the background. Update the background pixel points whose values vary in a certain range within a certain number of frames to solve the problem of deadlock. The experiments show that this method can segment the moving vehicles rapidly and accurately from the video. And the shortage is neglecting the effect of overlapping and the shadow, Compute each pixel by a model leads to heavy computational burdens, it needs constant improvement.

ACKNOWLEDGMENT

The National High Technology Research and Development Program ("863" Program) of China (Funding No.: 2009AA01Z332)

REFERENCES

[1] M. Tomizuka, “Automated Highway Systems. An Intelligent Transportation System for the Next Century,” IEEE International Symposium on Industrial Electronics, pp.1-4, 1997.

[2] Hongjiang Zhang, Yihong Gong, et al, “Moving Object Detection, Tracking and Recognition,” The 3rd International Conference on Automation, Robotics and Computer Vision, 1994.

[3] YANG C H and CHUNG P C, “Knowledge-based automatic change detection and positioning system for complex heterogeneous environments,” Intelligence and Robotic Systems, pp. 85-98, 2002.

[4] R.M. Haralick and J.S. Lee, “The Facet Approach to Optical Flow, Image Understating Workshop,” Arlington, VA, pp.74-83, 1984.

[5] A. Verri and T. Poggio, “Motion field and optical flow: Qualitative properties,” IEEE Transactioins on Pattern Analysis and Machine Intelligence. vol. PAMI-11, pp. 490-498, May1989.

[6] R. JAIN, “Difference and accumulative difference pictures in dynamic scene analysis,” Image and Vision Computing, pp. 99-108, 1984.

[7] C.R. Wren, A. Azarhayejani, T. Darrell, “Real-time tracking of the human body,” IEEE Transactions on Pattern Analysis and Machine Intelligence, pp.780-785, 1997.

[8] A.M. Elgammal, D. Hanvood, L.S. Davis, “Non-parametric model for background subtraction”. In: Proceedings of the 6th European Conference on Computer Vision. Dublin, Ireland: Springer-Verlag, pp.751-767, 2000.

[9] C.Stauffer and W.E.L. Grimson, “Learning Patterns of Activity Using Real-Time Tracking,” IEEE Transactions on Pattern Analysis and Machine Intelligence, pp.747-757, 2000.

[10] C. Stauffer and W.E.L. Grimson, “Adaptive background mixture models for real-time Tracking,” CVPR’99 pp.246-252, 1999.