

## **A Real-time Traffic Congestion Estimation Approach from Video Imagery**

**Li Li<sup>1</sup>, Jian Huang, Xiaofei Huang, Long Chen**

*Joint Lab for Intelligent Computing and Intelligent Systems, School of Electronic Information  
Wuhan University, 129 Luoyun Road, Wuhan, Hubei 430079, P.R. China*

---

**Abstract:** This paper presents a novel approach to detect traffic congestion on roads in a natural open world scene observed from TV cameras placed on poles or buildings. In this system, three algorithms are proposed to estimate the road status. They are based on background subtraction, frame differencing and time-spatial image respectively. The experimental results on real road traffic congestion estimation show that the time-spatial method is the most robust in complex lighting and traffic environment. The detailed algorithms and the comparison results are given in the paper.

**Keywords:** Traffic congestion estimation; Background subtraction; Frame difference; Time-spatial imagery

---

### **1. Introduction**

With the development of traffic infrastructure, many cities like Wuhan, China, installed hundreds surveillance CCD in urban area. Currently, these video always transfer to a monitoring center and analyze manually. The purpose of the work presented here was to provide video imagery based urban traffic monitoring methods to realize real-time traffic congestion estimation.

Traffic congestion in urban road and freeway networks leads to a strong degradation of the network infrastructure and accordingly reduced throughput, which can be countered via suitable control measures and strategies [1]. Most traffic information in ITS is based on ground sensors like induction loops, bridge sensors and stationary cameras. Sensor based system only used in few main roads due to high cost. Video based system used more wildly now, however, it's inefficiency in traffic congestion estimation. Thus because few video based methods were used in the system.

We made a survey of research work elsewhere, many works have been done on vehicle tracking

applications that use image sequences, which mainly focused on vehicle classification and tracking [2-5]. In our system, the surveillance CCD camera is installed in a relatively far distance from a highway and the vehicles are visualized as small objects with minimum detail on their geometrical model. The focus of our system is to estimate the road congestion situation rather than the detailed vehicle types and counts.

For video based traffic congestion estimation proposal few publications were reported. Hence, we proposed three algorithms to analyze and compare the performance and robustness in different environments.

Normally the traffic situation can be roughly categorized into two states, open flow and congestion. However, these two categories are not enough to describe the traffic situation. Thus, in our system we proposed three traffic patterns; congestion, Mild congestion (MC) and Open flow (OF) are defined. Congestion: there are a large number of vehicles and almost all of the vehicles run very slowly or completely stopped. Mild congestion: there are a large number of vehicles and most vehicles run slowly. Open flow: vehicles run at normal speed. According to this definition, two parameters can be used to determine the road situation: the occupation rate of the vehicles on the

---

<sup>1</sup> Corresponding author.  
Email address: ll@eis.whu.edu.cn

road and the status of the movement of the vehicles. In the Open flow situation, the occupation rate of the vehicles on the road is low and the vehicles are moving at normal speed; in the mild congestion situation, the occupation rate of the vehicles is relatively high and the movement is slow; in congestion situation, there's hardly any movement and the occupation rate is very high. So our task can be categorized into two tasks:

- Detect the movement status of the vehicle.
- Compute the road occupation rate of the vehicle.

Several ways are tried to detect the movement status of the vehicle and road occupation rate.

Detection of changes between sequences of frames is a major task in our system. In terms of detection of the movement status, *background modeling* and subtraction is a very popular approach because it's easy to implement, but in some cases like heavy congestion and abrupt change in light, it's hard to get a satisfying background model.

Another method is *frame differencing* followed by contour detection, which has relatively steady performance against environmental changes. However when the camera places very high and the vehicles seem very small in the video, the results is unsatisfying.

Traffic congestion situation in a certain site reflects the relationship of moving objects – vehicles between temporal and spatial. Based on this principle we proposal *time-spatial image* based method for congestion estimation. The most similar thing that we found was the work by Antonio Albiol *et al.* [6] for people counting. They built space-time images and use morphological tools for detection. The concept is the same, but the objective and environment are difference. They focus on objet segmentation and isolation, in our case object situation is the most important information. However they do not seem to have the problem of objects stagnation.

Recently, A. Liu *et al.* [7] and J. Wu *et al.* [8] proposed methods based on virtual line for vehicle detection. Their approach has some similarities with ours – they use video base image and detect vehicle. There are however many differences between their work and ours; they use background information and virtual line for vehicle detection optimize, while we not use background but get pattern characters from the generated time-spatial image. The environment shown in their examples is in freeway and daylight condition, while our case is more complex in urban area and intercourse.

To be an effective traffic surveillance tool, a video image processing system should meet several stringent requirements [9]:

- Function under a wide range of traffic conditions: light traffic, congestion, varying speeds in different lanes.
- Function under a wide variety of lighting conditions: sunny, overcast, twilight, night, rainy, etc.
- Operate in real-time.

In our system, we first raised time-spatial image methods for traffic congestion estimation and used different methods in complex environment. All of the algorithms meet the requirements and demonstrate good results. The time-spatial method is the most robust in complex lighting and traffic environment.

The rest part of the paper is organized as follows. In section 2 describes the proposed methods and algorithms. Section 3 presents comparison experimental results and Section 4 concludes this paper.

## 2. Traffic congestion estimating methods

### 2.1 Background modeling based method

Background modeling is a very popular approach for motion detection since it's easy to implement. In this algorithm, first we construct a background image, update it dynamically, then taking absolute differences between each incoming frame and a background model of the scene.

The sequence of major procedure is described as follows:

**Step1:** Predefine region of interest (ROI) mask of the frame. A ROI corresponds to a traffic lane. Since traffic camera is assumed to be stationary, these regions are entered once at the beginning.

**Step2:** Construct a background image. Although the simplest background model is the temporally averaged image, a background approximation that is similar to the current static scene is median image. So we compute the original background image by taking the median for each pixel over the previous N frames. It is reported that median value is far more robust than mean value [10].

In order to avoid the problem of changes in ambient light, this background reference image is updated periodically (Fig.3). In the update procedure, only those pixels that do not have a large difference with the corresponding pixel of the last updated version of the background image (i.e. they still belong to the background), are considered in the update procedure. The updating equation is as follows:

$$B_{k+1}(x) = B_k(x) + (\alpha_1(1 - M_k(x)) + \alpha_2 M_k(x))(I_k(x) - B_k(x)) \quad (1)$$

Where  $B_k(x)$  is the pixel value of current

background image,  $Bk + I(x)$  is the image to be updated,  $I_k(x)$  is the pixel value in the newest frame, and  $\alpha 1, \alpha 2$  is the update coefficient. And the equation to compute  $Mk(x)$  is listed below,

$$M_k(x) \begin{cases} 1, & |I_x - B_x| > T_b \\ 0, & \text{otherwise} \end{cases} \quad (2)$$

**Step3:** Thresholding after the background subtraction. After evaluating many threshold methods for change detection, we decide to use the method proposed in 2006[11]. This non-parametric algorithm divided the image into some blocks, then computes a threshold of each block of an image adaptively based on the scatter of regions of change (ROC) and averages all thresholds for image blocks to obtain the global threshold. Compared with some gray-level distribution based methods, this method has significantly better performance (Fig.4).

**Step4:** Use closing and Opening morphological operators to fill little apertures and remove small particles depending on the size of the structuring element selected for these operations (Fig.5).

**Step5:** Calculate vehicles road occupation ratio  $r_v$  and change rate  $r_c$ .  $r_v$  which is obtained by the sum of the vehicle area  $a_v$  divided by the area of the ROI  $a_r$ , and the changing area rate  $r_c$ , is obtained by

$$a_v(n-1) - a_v(n) / a_r \quad (3)$$

**Step6:** Congestion estimation by following criteria:

Road status is set to congestion if:

$$CG: \{r_v > T_v \text{ And } r_c < T_c\} \quad (4)$$

Likewise, the road status is set to mild congestion if:

$$MC: \{r_v > T_v \text{ And } r_c > T_c\} \quad (5)$$

And road status is set to open flow if:

$$OF: \{r_v < T_v\}. \quad (6)$$

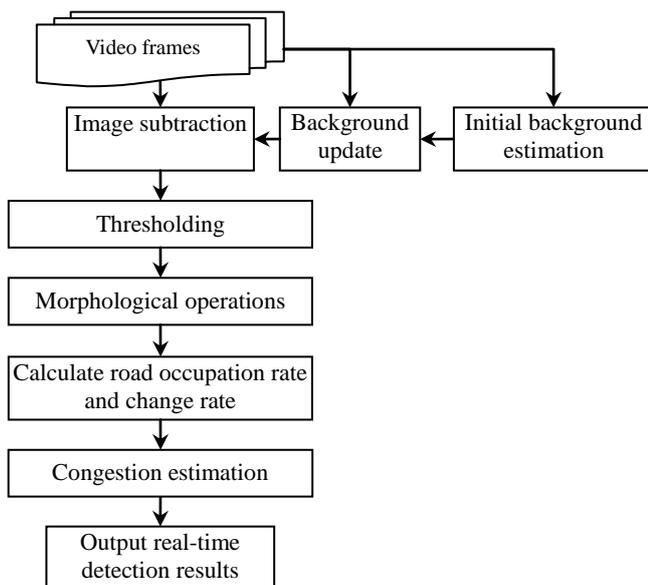


Fig.1. Flow chart for background modeling based method.

Background modeling has some limitation in some cases. When the road becomes congested for a long time, like 2 minutes or above, it's hard to set up an initial background. When the ambient light changes suddenly, the background updating procedure may also fail to track the change.

Fig.1. shows the flowchart of the background modeling based method.



Fig.2. Original image.



Fig.3. Generated background image.

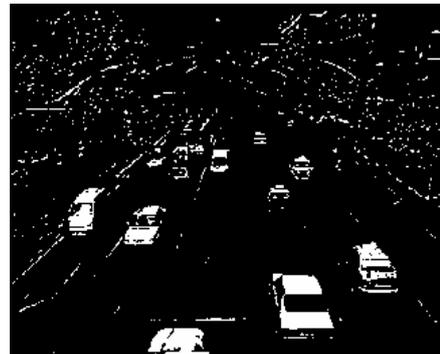


Fig.4. Difference image after thresholding.



Fig.5. Image after close and open operations.

## 2.2. Frame difference based detection method

In this method, frame differencing method is applied to determine where a motion occurred. Compared with background subtracting method, frame differencing is much more stable against light changes in certain environment. The sequence of major procedure is described as follows:

**Step1:** Select frames by timer. The absolute difference of two frames is taken at an interval of 3 seconds (Fig.7-9).

**Step2:** Threshold the difference image (Fig.10).

**Step3:** Use downsampling, dilate morphological operation, and upsampling step to avoid the noise and some discrete holes in the difference image.

**Step4:** Use contour detection method to determine the number of contours in the binary image, thus estimating the number of the moving vehicles (car\_num) (Fig.11). Four variations of algorithms described in [11] to retrieve borders.

**Step5:** Use Canny Edge Detector to detect edges [13]. It has four stages:

Stage 1. Image Smoothing

Stage 2. Differentiation

Stage 3. Non-Maximum Suppression

Stage 4. Edge Thresholding

**Step6:** Calculate the road occupation rate.

It takes grayscale image on input and returns bi-level image where non-zero pixels mark detected edges (Fig.12). Then we compute  $re$  the percentage of non-zero pixel has taken of the ROI.

**Step7:** Congestion estimation by following criteria:

Road status is set to congestion if:

$$CG: \{re > Te \text{ And } Car\_num < 1\} \quad (7)$$

Road status is set to mild congestion if:

$$MC: \{re > Te \text{ And } Car\_num \geq 1\} \quad (8)$$

Road status is set to open flow if:

$$OF: \{re < Te\} \quad (9)$$

Fig.6. shows the flowchart of the background modeling based method.

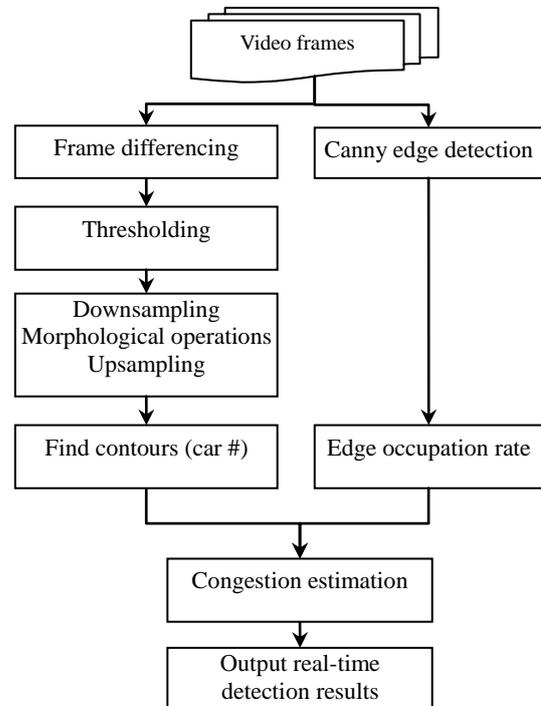


Fig.6. Flow chart for frame difference and contour detection method.



Fig.7. Original image 1 (20:00:16pm).



Fig.8. Original image 2 (20:00:19pm).

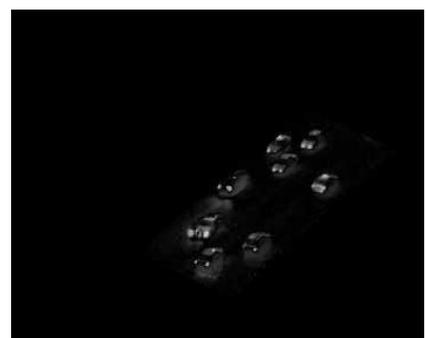


Fig.9. Image differencing after mask.



Fig.10. Image after thresholding.



Fig.11. Contour Image

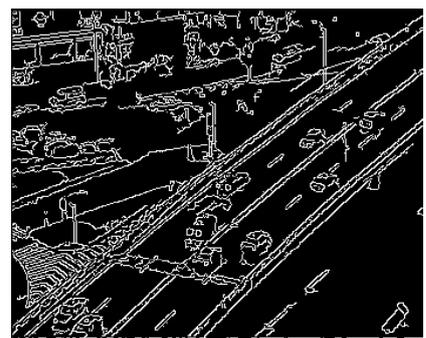


Fig.12. Canny Edge

### 2.3. Time-Spatial image based method

#### A. Generation of Time-spatial image

In two dimension video image, the relationship between moving objects and static background isn't easy to find. Since a vehicle usually takes more than one frame to cross a *virtual detection line* on the road, some time memory must be included in the algorithm [6].

We acquire a new frame to store certain lines of the frame onto what we have called *Time-spatial image*. These stacks of lines contain all the required information to detect vehicle status, and can be considered itself as an image.

Time-spatial image merges the characteristic both of temporal series and spatial sequence. It's good for reflecting differences of moving vehicles and static objects. In time-spatial image, image spreads in space and time coordinate as Fig.13.(a). Such images can be perceived as detector staring-map by line detector. We could say that stacking act like an information condenser, passing from a sequence from Time-spatial image slice. Different rows are reflected in the corresponding detection line in time. In the paper a white line selected as line detector, see white lines in each frame of Fig.13.(a)., and in chronological order from left to right these detection lines generated time-spatial images Fig.13.(b).

Under this generation method, time-spatial images reflect the image of a fixed region (line) changes with time. If no objects moving in detect region, there is no gray changed. Otherwise, gray will be changed by moving objects.

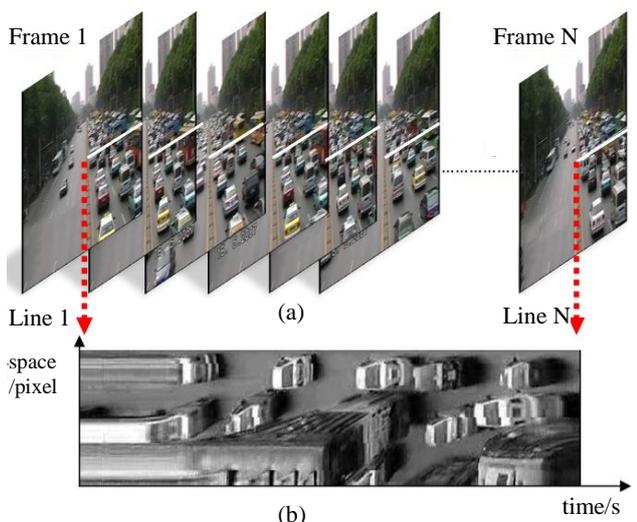


Fig.13. The Generation Procedure of Time-Spatial Image.

In certain sample frequency, the faster the movement objects, the shorter it stayed in the region;

the contrary is longer. In the image, the length of gray changes  $t$  is inversely proportional to object velocity  $v$ . After normalized, it's according to following equation:

$$v=1/t \quad (10)$$

When objects is static, the length  $t$  (x axis) is infinite; when speed is not zero, the length  $t$  is limited, with Eq.(10)

In time-spatial image, horizontal axis is the time line shows vehicle time information, while the vertical one is space-axis shows vehicle space information. The image height equal to the detection line (white line) length; and the image width ( $t$ ) can be calculated as:

$$t=N/a \quad (11)$$

$N$  is the number of frames, and  $a$  is the frame rate per second. In the paper  $N=100$ ,  $a=15$ , thus  $t=6.67s$ .

#### B. Collection rule of Time-spatial image

To ensure traffic congestion estimate accuracy, a set of rules are proposed in detection line setting and time-spatial image generation. These rules are listed as follows:

**Rule1:** set the line perpendicular to the direction of cars travelling and near intersection of roads;

**Rule2:** span of the line includes vehicles' projection area on the road;

**Rule3:** width fix to 1pixex; length longer than 2 lanes width in the same direction, no less 50 pixels;

**Rule4:** frame number no less than 100 to get enough information and no more than 150 to ensure real-timing.

Fig.14 gives the time-spatial images of all kinds traffic situation in scene 1.

#### C. Relationship of Time-spatial image and CG

From time-spatial images, it can directly reflect the state of road congestion and trends. When vehicle speed faster, image block length is shorter and area is smaller; when speed slower, the length is longer and area also will be increases; when road congestion, the images show many straight lines or horizontal stripes. The obvious characteristics can be detected with Hough transform, Fig.14-15 shows typical situation of OF and CG.

We use Canny Edge Detector to detect edges as 2.2.described. Then use Hough transform to detect edge stripes or lines. Then calculate the length  $L$ , direction  $D$  and line number. Filter the lines with these parameters, and get valid horizontal line number  $N$ . The edge occupation rate  $r$  can be calculated as 2.2 described, compare  $r$  with threshold  $T_r$  and  $N$  with  $T_N$ . If both of them greater than given threshold, it belongs to congestion status, otherwise road is over flow. This can be defined by:

$$CG: \{ N > T_N \text{ And } r > T_r \text{ And } L > T_L \} \quad (12)$$

$$OF: \{ \text{otherwise} \} \quad (14)$$

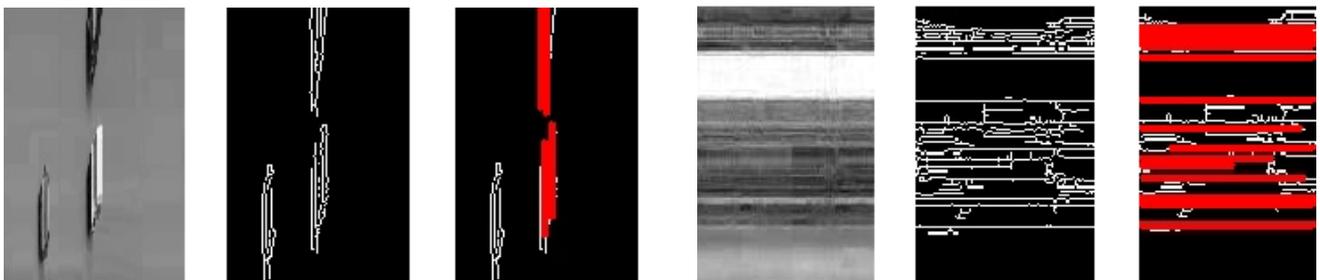
$$MC: \{ N > T_N \text{ And } r > T_r \text{ And } L < T_L \} \quad (13)$$



(a) OF, no vehicle in the period (b) OF, with vehicle in the period (c) OF to CG  
Fig.14. Time-spatial images examples under different traffic conditions



Fig.15. 2minutes time-spatial images analyze in scene 2 (original size 50\*1800 )



(a) Over flow (with vertical line ) (b) Congestion (with horizontal line)  
Fig.16. Time-spatial image congestion analysis

#### D. Congestion estimation by Time-spatial Image

After learning congestion estimation strategy by image like Fig.16, the time-spatial image can be used to detect congestion by Hough transform and edge detection as following steps:

**Step1:** Image preprocessing, change color frame to grey;

**Step2:** Setting detection line as rule1-4. In our work line can be set in any direction, thus can meet different layout of intercourse and lanes;

**Step3:** Generate Time-spatial image;

**Step4:** Use Canny Edge Detector to detect edges as 2.2.described.

**Step5:** Use Hough-transform to detect line pattern. Then calculate the length  $L$ , direction and line number. Filter the lines with these parameters, and get valid horizontal line number  $N$ .

**Step6:** Calculate the edge occupation rate  $r$ ;

**Step7:** Congestion estimation according to eq.(12-14)

Fig.17 shows the flowchart of the time-spatial image detection algorithm.

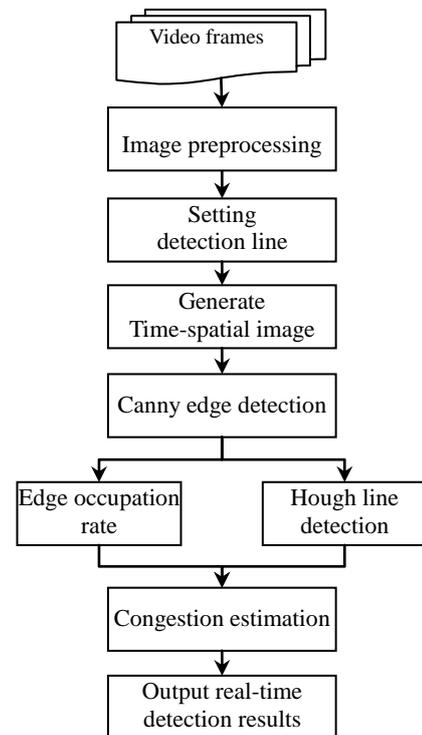


Fig.17. Flow chat for time-spatial image base method.

### 3. Experimental results

The proposed method is evaluated on real video clips which covered more than 20 roads and intersection in Wuhan City of China. These video are collected in various traffic and light conditions: night, evenfall and daytime. In some of the video, the camera is placed on the viaduct and jitters seriously. Each site includes OF, CG and MC clips. Fig.18 shows different clips for congestion estimation.

Original video frames are color image, with size of 352\*288 (CIF/PAL), and frame rate of 15FPS. Each frame is converted to eight bits gray scale in image preprocessing procedure.

From Table 1, 20 sites materials divided into three groups according to corresponding situation. Each column gives statistic results under different methods. It can be seen the detection rate of proposed methods are from 87.8% to 96.6%. The detection rates vary with the complexity of road traffic and environment situation. Generally, the average detection rates in daylight are greater than that of night and jitter situation. M3 shows very robust in different situation and have higher detection rates than M2 and M1.

The false estimation results are mainly due to low contrast image, small vehicle block and irregular lane direction, etc. In normal condition our experimental results showed high accuracy compared with manual results.

### 4. Conclusions

This paper presents some traffic congestion estimating approaches that is designed to operate under challenging conditions like lighting transitions between night/day and day/night, camera vibration due to wind etc.

Background subtraction method introduced in part 2.1 is easier to implement but more vulnerable to ambient light change. In Frame difference and Contour detection method, frame difference method is used to detect the move status of the car, and canny edge detector to check the occupation rate. But its results are very sensitive of detecting area. Base on the research work of these two methods, we propose a novel approach for video congestion estimation.

The concept of Time-spatial image was introduced in our approach and was found to have good congestion estimation performance.

Future work would mainly setup congestion estimation model in disadvantageous situation.

### Acknowledgements

This work is supported by National Nature Science Foundation of China under Grant No. 40571102. Some of the video used in the work were provided by Wuhan Sanli Communication LTD. of China.

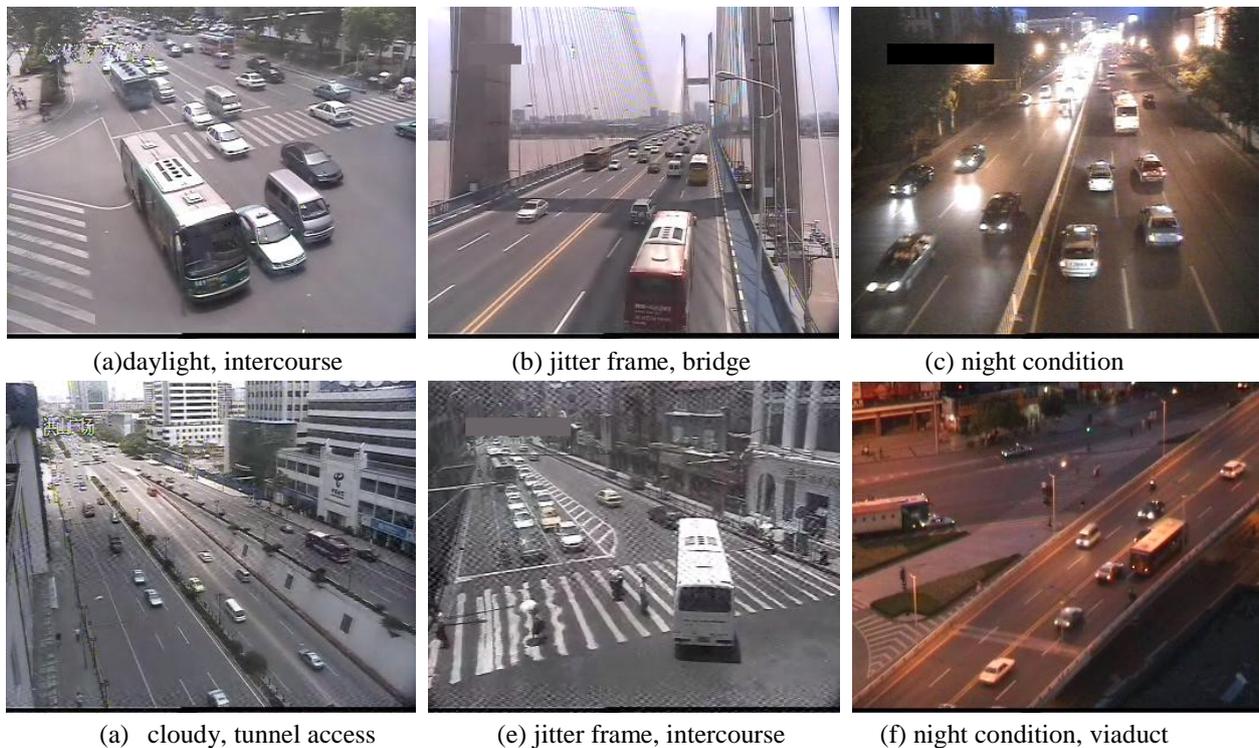


Fig.18. Different Testing Environment

Table 1. Detection rate comparison results of three methods

Site	Situation	Manually		M1		M2		M3	
		#OF	#CG	#OF	#CG	#OF	#CG	#OF	#CG
1	daylight	13	7	12	6	13	6	12	7
2	daylight	15	11	14	9	14	10	14	10
3	daylight	12	9	12	8	12	8	12	8
4	daylight	13	7	13	6	13	6	12	6
5	daylight	14	6	13	6	13	6	13	6
6	daylight	20	5	19	5	19	5	19	5
7	daylight	17	6	16	5	16	5	16	5
8	daylight	11	9	11	7	11	8	11	9
9	daylight	15	8	15	8	15	8	15	8
10	daylight	19	6	17	6	18	6	18	6
11	jitter	13	10	12	9	12	9	12	9
12	jitter	16	9	15	8	15	8	15	8
13	jitter	13	7	12	6	12	7	12	7
14	jitter	17	7	15	5	15	6	15	7
15	jitter	15	5	13	5	13	5	14	5
16	jitter	14	11	14	10	13	10	14	10
17	night	20	15	18	13	18	14	19	14
18	night	22	13	19	11	20	11	20	12
19	night	27	8	24	7	25	7	24	7
20	night	30	5	26	5	28	5	29	5
TOTAL		336	164	310	145	315	150	316	154
Detection rate (daylight)				95.3%	89.2%	96.6%	91.9%	95.3%	94.6%
Detection rate (jitter)				92.0%	87.8%	90.9%	91.8%	93.2%	93.9%
Detection rate (night)				87.9%	87.8%	91.9%	90.2%	92.9%	92.7%
Detection rate (Average)				92.3%	88.4%	93.8%	91.5%	94.0%	93.9%

Note: M1: background based method  
M2: frame difference based detection method  
M3: time-spatial image based detection method

## References

- [1] Markos Papageorgiou, Christina Diakaki, Vaya Dinopoulou, Apostolos Kotsialos, and Yibing Wang, "Review of Road Traffic Control Strategies", In: *Proc. of the IEEE*, VOL. 91, NO. 12, pp. 2043-2067, 2003.
- [2] B. Li, Q. Chen, "Freeway Auto-surveillance From Traffic Video", In: *Proc. of 6<sup>th</sup> International Conference on ITS Telecommunications*, pp.167-170, 2006.
- [3] X. Ma W. Eric L. Grimson, "Edge-based rich representation for vehicle classification", In: *Proc. of the Tenth IEEE International Conference on Computer Vision*, pp1-8, 2005.
- [4] Hiroshi Inoue, Mingzhe Liu, and Shunsuke Kamijo, "Vehicle Segmentation by Edge Classification Method and the S-T MRF Model", In: *Proc. of the IEEE ITSC*, pp.1543-1549, 2006.
- [5] Z. Yang, H. Meng, Y. Wei, H. Zhang, X. Wang, "Tracking Ground Vehicles in Heavy-traffic Video by Grouping Tracks of Vehicle Corners", In: *Proc. of the IEEE ITSC*, pp.396-399, 2007.
- [6] Antonio Albiol, Inmaculada Mora, and Valery Naranjo, "Real-time High Density People Counter Using Morphological Tools", *IEEE Trans. Intelligent Transportation System*, vol. 2, No.4, pp.204-218, Dec. 2001.

- [7] A. Liu, Z. Yang, J. Li, "Video Vehicle Detection Algorithm based on Virtual-Line Group", In: *Proc. of IEEE APCCAS*, pp.1148-1151, 2006
- [8] J. Wu, Z. Yang, J. Wu, A. Liu, "Virtual line group based video vehicle detection algorithm utilizing both luminance and chrominance", In: *Proc. of IEEE IEAC*, pp. 2854-2858, 2007
- [9] Benjamin Coifman, David Beymer, Philip McLauchlan, Jitendra Malik, "A real-time computer vision system for vehicle tracking and traffic surveillance", *Transportation Research Part C 6* (1998) 271-288
- [10] L. Wang, W. Hu, T. Tan, "Recent developments in human motion analysis", *Pattern Recognition 36*, pp.585 –601, 2003.
- [11] Su, C., Amer, A., "A Real-Time Adaptive Thresholding for Video Change Detection, Image Processing", In: *Proc. of 2006 IEEE International Conference on 8-11 Oct.* pp.157 -160, 2006.
- [12] S. Suzuki, K. Abe. "Topological Structural Analysis of Digital Binary Images by Border Following", *CVGIP*, v.30, n.1. pp. 32-46, 1985.
- [13] J. Canny. "A Computational Approach to Edge Detection", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(6), pp.679-698, 1986.