Research on Highway Traffic Event Detection Method Based on Image Processing

Li Xia Bao¹, Qiulan Wang¹*, Wenliang Qu¹ and Xianglun Mo²

¹Intelligent Transport System (ITS) R&D Center, Shanghai Urban Construction Design and Research Institute (Group)Co., Ltd., Shanghai, 20082, China
²Department of transportation, China University of mining and technology, Xuzhou, Jiangsu, 221116, China

*Corresponding author’s e-mail: wangqiulan@sucdri.com

Abstract. Aiming at the problem of the image processing technology of the expressway event detection system, this paper uses the background subtraction method and Camshift algorithm to detect and track the moving vehicle. Through sorting out the traffic event detection and processing process of Lian-Xu Expressway, the traffic event detection system is constructed, and the data processing module of the system is mainly studied to realize the detection and alarm of traffic events. The results show that the highway traffic incident detection system studied in this paper can complete the detection of traffic events in the video, improve the detection rate, and provide technical support for the highway safety management.

1. Introduction

Highway transportation is an important part of China's transportation system. Compared with other modes of transportation, the accident rate of road transportation is higher. Although there are fewer traffic accidents on expressways in the highway system, the speed of vehicles on expressways is faster and the response time of drivers is shorter. Once the accident happens, the severity will be higher than that on ordinary highways. The occurrence of highway traffic accidents is often accompanied by huge casualties and economic losses. The automatic detection technology of expressway traffic accident is helpful to find expressway traffic accident quickly, and take corresponding measures to reduce the occurrence of expressway traffic accident. At the same time, the expressway traffic accident detection technology is helpful to find the expressway traffic accident quickly, and start the corresponding emergency plan, and reduce the loss of life and property caused by the traffic accident.

Along with the construction of expressway information, the expressway management department has established a relatively perfect monitoring video system, but at the present stage, the detection of expressway traffic incidents is still dominated by manual detection, which mainly includes three ways: manual random watching of monitoring video, manual inspection and manual alarm. Although the video surveillance network covering the entire expressway has been established, the monitoring efficiency of the expressway video surveillance network has not been further explored, resulting in the low monitoring efficiency of the existing video surveillance network on the expressway. In view of the above problems, this research develops a set of automatic detection system of expressway traffic incident based on expressway surveillance video processing.

In the 1960s, the New York State government established the earliest event detection system in Lincoln Tunnel, which was based on the occupant condition of the traffic detector to detect traffic...
incidents. Until the 1990s, the detection of traffic events was mainly based on the detection of the change of traffic flow and occupancy on the road through the detector, and the calculation was carried out by mathematical algorithms[1]. With the development of mathematical methods and computer technology, the expressway incident detection algorithm develops rapidly, such as ANN, Fuzzy and wavelet theory. The practicability and effectiveness of these algorithms are limited in the context of the rapid increase of traffic volume and complex traffic conditions. In 2001, Yik Jung and Gerard Medioni et al. proposed an algorithm to detect the trajectory of moving targets. Later, M Brand and V Kettaker proposed a video-based Markov modeling method to solve the entropy minimum condition of joint distribution. Based on the Hidden Markov entropy minimum, this method can detect traffic events and segment accidents[2]. Srinivasan et al. proposed an urban vehicle detection algorithm based on state space transformation and neural network[3]. Wang Qi proposed to use neural network theory to detect traffic events, and proposed a traffic flow state estimation method based on BP neural network[4]. Wang Changfeng proposed to build an expressway traffic incident detection system by combining background update technology with analyzer, data server and manager, and realize alarm and automatic recording functions at the same time[5]. After studying the trajectory of vehicles, Yang Yuan proposed to design a traffic event detection system based on video to track the trajectory of vehicles[6]. After studying the wavelet SVM, Chen Zhijian constructed the SVM model by using polynomial, radial basis and Sigmoid functions, and realized the detection of expressway events by analyzing the running state of expressway and determining parameters[7]. In recent years, with the emergence of high and new technologies such as artificial intelligence and pattern recognition, a large number of new IT-based algorithms have been proposed, which are more accurate and efficient[8-11].

The above traffic incident detection technology is mainly to extract pictures from the video, and then study various image processing technologies. Through the processing of the pictures, the vehicles in the pictures are extracted, and various state detection and analysis are carried out on the extracted vehicles. Algorithms with good vehicle detection effect generally have poor real-time performance, and the detection accuracy with good real-time performance is not high[12-13]. Moreover, they track vehicles solely through video detection, which only can detect and track vehicles without completing a complete traffic incident detection alarm mechanism. In view of the above problems, this paper mainly studies the data processing module of the traffic incident detection system. The main research contents are as follows: First of all, the data processing module is divided into moving target detection algorithm and the moving target tracking algorithm, and then respectively using the method of background difference and joined the Camshift algorithm of kalman filter for target detection and tracking, finally built on matlab data processing module and video processing of the test, can realize the function of target extraction and tracking.

2. Detection process

Traffic events are basically random, and the time, place, severity and time of affecting traffic are uncertain. This randomness is determined by the randomness of traffic demand. There are many parameters for the description of traffic events. Too slow or too fast driving speed on the expressway may lead to the occurrence of traffic accidents on the expressway. At the same time, the vehicle speed also drops to zero after the occurrence of traffic accidents on the expressway[14-15]. Video based traffic event detection technology, the use of image processing technology, the first image in the video in the form of frame to extract the frame picture, and then the picture of the vehicle detection, the next picture of the frame of the vehicle tracking and speed detection, to determine whether the traffic event. The current running state of the vehicle can be judged by analyzing the speed vector and other parameters of the vehicle in the video and setting the threshold. When some parameter values of the vehicle reach the threshold, the system will alarm and identify the occurrence of traffic events. The expressway traffic event detection process based on image processing is shown in Figure 1.
Figure 1. Expressway traffic incident detection process based on image processing

3. Key technologies

3.1. Vehicle Detection Based on Background Difference Method
The basic steps of the background difference method are as follows: firstly, the background model is modeled, the background model $B(x,y)$ is obtained after all pixels are counted by using the selected background model or the images in the image sequence frames in the later stage, and then the image frame $F(x,y)$ that needs to be recognized is carried out difference operation with the background model to obtain the initial difference result $D(x,y)$. The above steps are repeated until the regional information of the determined target is obtained. The process of the modeling algorithm of the mixed Gaussian background is shown in Figure 2. The specific calculation steps are as follows:
Matching the existing distribution

The new
distribution

Match distribution
parameter updates

Updating background

Detecting prospects

Ending detection

Figure 2. Background Difference Method algorithm flow

(1) Parameters initialization of background model

Initialization is to use the first N frames of the video data, initialization processing a background model. Given that each pixel has a different Gaussian distribution to describe its value, the mean value and variance are used to represent the central value of the model function and the distance between the current pixel and the central value of the model respectively. The weight is used to represent the matching degree of pixels and each model. The parameters of the initial model are 0 except that the mean value of the first Gaussian distribution is the pixel value of each pixel in the first frame. If the distribution variance is different, select a larger value. The parameter formula of mean value and weight is as follows, Represents the mean of pixel values, represents weights.

\[
\mu_i = 255 \times \frac{i}{k} \quad (1)
\]

\[
\omega_i = \frac{1}{k}, \quad i = 1, 2, 3, \ldots k \quad (2)
\]

(2) New pixel matching judgment

After initializing the model, each pixel of the following video frame needs to be compared with the pixel of the same position of the previous model, and the model should be updated or not according to the judgment criteria. Judgment criteria are:

\[
|x_t - \mu_{k,t-1}| < 2.5\sigma_{k,t-1} \quad (3)
\]

where: \(x_t\) — Current pixel;
\(\mu_{k,t-1}\) — Mean value of future matching model;
\(\sigma_{k,t-1}\) — variance of future matching model.

(3) Update the background model
If the above criteria are met, then this point is the background point, and all parameters of the matching model need to be updated. The updating formula is as follows:

\[ \mu_{k,t} = (1 - \omega) \mu_{k,t-1} + \alpha x_t \]  
(4)

\[ \sigma_{k,t}^2 = (1 - \beta) \sigma_{k,t-1}^2 + \beta (x_t - \mu_{k,t-1})^2 \]  
(5)

\[ \beta = \frac{\omega_k}{\omega_{k,t-1}} \]  
(6)

\[ \omega_{k,t} = (1 - \alpha) \omega_{k,t-1} + \alpha M_{k,t} \]  
(7)

where: \( \omega_{k,t} \) —— Weight values at different times. The higher the matching degree of the Gaussian component is, the larger the weight value is;

\( \beta \) —— Degree of updating speed of each parameter;

\( M_{k,t} \) —— Whether the model matches the regulation values, the value is 0, 1.

(4) Normalization of weight

After updating the parameters of the model, the weight value changes, which needs to be normalized. In this paper, the ratio of weight to variance \( \frac{\omega_{k,t}}{\sigma_{k,t}^2} \) is used to sort each model. If the weight value is too low, the background judgment result has no reference value. Generally, M points are selected as the matching model, and the appropriate T value is selected first. The larger the T value is, the better the detection effect will be, but the amount of calculation will increase accordingly, so it is good to select the appropriate T value.

\[ M = \text{argmin}_m (\sum_{t=1}^{m} \omega_{k,t} > T) \]  
(8)

(5) Determine the foreground/background

When none of the k models satisfies Equation 6, the current point is judged as the foreground point. Since the model with lower weight was abandoned above, a new Gaussian component needs to be determined by the current pixel point. Let the Gaussian component be the mean value of the current pixel value, and the variance is determined as in the first step. Weight update and normalization are carried out again:

\[ \omega_{k,t} = (1 - \alpha) \omega_{k,t-1} \]  
(9)

\[ \omega_{k,t} = \frac{\omega_{k,t}}{\sum_{k=1}^{k} \omega_{k,t}} \]  
(10)

After the background model is determined, the original image is differentiated from the background model, then the foreground image is:

\[ D_t(x, y) = |I_t(x, y) - M_t(x, y)| \]  
(11)

where: \( M_t(x, y) \) —— background model;

\( I_t(x, y) \) —— Origin image.

3.2. Improved CamShift moving vehicle tracking algorithm based on Kalman filter

The full name of CamShift algorithm is "Continuously Adaptive Mean-Shift", CamShift algorithm is essentially Mean-Shift algorithm. Mean-shift algorithm is mainly iterative optimization to find the extreme point of density distribution in the sample, and the method is self-adaptive step size, as shown in the figure, is the target point, x is the objective function, and is the offset vector generated by it. In the application of target tracking, the target template can construct the kernel density estimation function based on the color histogram of the target, and search the target as the search feature.

The color probability distribution of the target template is represented by M-level histogram:

\[ q_{it} = C_i \sum_{i=1}^{i=k} \frac{\sigma_i^2}{\sigma_{i,t}^2} \delta [b(x_i^t) - u] \]  
(12)

where: \( q_{it} \) —— color probability distribution, \( u = 1, 2, m; \)

\( C_i \) —— normalization constant of, when \( \sum_{i=1}^{i=k} q_{it} = 1; \)

\( x_i^t \) —— The normalized pixel coordinates in the target template; \( i = 1, 2, n; \)

\( b(x_i^t) \) —— The bin of the pixel at in the quantized feature space.

Kernel density estimation was established for the target candidate region, and the probability distribution of color features was \( p_{it} \):

\[ p_{it} = \frac{1}{C_h \sum_{i=1}^{i=n} \frac{\sigma_i^2}{\sigma_{i,t}^2} \delta [b(x_i^t) - u]} \]  
(13)
where: \( x_i \) — The y-centered normalized pixel position information of the candidate target in the current frame;

\( h \) — Bandwidth, which can be used to determine the scale of a candidate target.

The Bhattacharyya coefficient, or Babbitt distance, often measures separability between classes in the classification process. Mean-shift algorithm uses Babbart distance to measure the similarity between template target and candidate target. The optimal position of the target is the position with the maximum coefficient, \( \max \rho(y_j, q) \).

\[
\rho(y_j, q) = \sum_{u=1}^{m} \sqrt{p_u(y_j)q_u}
\]  

(14)

In the actual solution, in order to simplify the calculation, Mean-shift algorithm will move iteratively through the local density features of the target region, so as to converge to the local maximum value of the probability distribution of similarity, and finally reach the optimal value position. This leads to the problem that when the target is moving too fast or the background is too complex, the result of iteration may not be the optimal result, that is, the moving target may not be fixed at the iterated position in one frame.

The iterative steps of Mean-shift algorithm are as follows:

(1) Initializing the target

In the step of initializing the target, information such as the target model, the location of the target and the size of the target will be extracted. The model is generally the color histogram feature of the region, namely in formula 15:

(2) Calculate the position of the candidate target model in the previous frame to initialize the candidate target position in the current frame and calculate the candidate target model, \( \{p_u(y_0)\}_{u=1,2,m} \).

(3) Calculation of Similarity

The similarity between the target and the candidate target is measured by the Babbitt distance coefficient.

\[
\rho(p(y), q) = \sum_{u=1}^{m} \sqrt{p_u(y)q_u}
\]  

(15)

(4) Calculate the next position of the candidate target \( y_{1} \).

Mean-shift algorithm adopts the mathematical method of the fastest descent. The steepest descent method does not directly calculate the step size and direction of numerical descent. Instead, the model similarity matching function is first expanded by first-order Taylor and then approximated, so as to get the next position of the iteration. After Formula 18 is expanded by Taylor at, the linear approximation is obtained.

\[
\rho(p(y), q) \approx \frac{1}{2} \sum_{u=1}^{m} \sqrt{p_u(y)q_u} + \frac{1}{2} \sum_{u=1}^{m} p_u(y) \left( \frac{q_u}{p_u(y_0)} \right)
\]  

(16)

Substituting Eq. 15 and Eq. 16 into the above equation, it can be obtained:

\[
\rho(p(y), q) \approx \frac{1}{2} \sum_{u=1}^{m} \sqrt{p_u(y)q_u} + \frac{c_h}{2} \sum_{i=1}^{n_h} w_i k(\frac{y-x_i}{h})
\]  

(17)

where

\[
w_i = \sum_{u=1}^{m} \frac{q_u}{\sqrt{p_u(y)}} \delta(b(x_i)-u)
\]  

(18)

As can be seen from Eq.19, the first part is independent of \( y \) and maximizes the Babbitt distance coefficient, while the second term needs to be maximized. The second term represents the kernel density estimation at coordinate \( y \) under the function. Pixel weights are calculated by. Then the local maximum point of kernel density estimation is calculated by Meanshift iteration, and the new target position is:

\[
y_{m+1} = \frac{\sum_{i=1}^{n_h} x_i w_i g(\frac{y-x_i}{h})}{\sum_{i=1}^{n_h} w_i g(\frac{y-x_i}{h})}
\]  

(19)

where: \( g(x) = -k^2(x) \).

(5) Calculate \( y_1 \) candidate target model
Calculate \( \{p_u(y_1)\}_{u=1,2,m} \) at \( y_1 \), and calculate the degree of similarity between \( y_1 \) and the target by Equation 14.

(6) Update the candidate target location

When \( \rho(p(y_1), q) < \rho(p(y_0), q) \), the candidate site target was updated with the order of \( \frac{1}{2}(y_0 + y_1) \rightarrow y_1 \), and was compared with \( \rho(p(y_1), q) \).

(7) Satisfy iteration exit

When the candidate position target is updated, go back to the second step to find the best matching position until the end condition of iteration is met \( \|y_1 - y_0\| < \varepsilon \) or the set number of iterations is met to exit the cycle.

CamShift algorithm can well deal with the phenomenon that moving vehicles appear far, near and large in the video, and well adapt to the change of the size of the moving target. However, when there are too many moving vehicles or the speed is too fast, the moving vehicles cannot be guaranteed to be tracked at all times. In the process of tracking the target by the algorithm, once the problem of occlusion or moving too fast occurs, the target process will appear the phenomenon of tracking loss, which needs to be assisted by other algorithms. In this paper, Kalman filter algorithm is selected to assist. Kalman filter has many advantages, the following are listed as follows: the recursive filtering process makes the Kalman filter does not need to occupy a large amount of memory of the computer, increase the real-time and availability; The calculation accuracy is high. Kalman filter is one of the best filters in the existing generative model tracing. The algorithm converges quickly. Algorithms can adapt to different kinds of systems, so they are widely used.

Aiming at the defects such as occlusion in the process of target tracking and the loss of target due to too fast transport speed, the CamShift algorithm can be improved effectively and the robustness of the algorithm can be increased by introducing the Kalman filter algorithm with prediction mechanism. The specific steps are as follows:

1. Initialize and extract the moving target through the foreground detection algorithm to obtain the target's position, regional radius and color histogram.
2. Iterate, match and track the moving target region through CamShift algorithm. At the same time, the current position information is transmitted into the Kalman filter and used as the measured value.
3. Determine whether the current frame has occlusion or target loss. Target occlusion is the moving target region is less than the set threshold value.
4. If there is occlusion or abnormal loss of target, the Kalman prediction mechanism will be turned to. The Kalman tracking mechanism introduced in the last section will be used to update the target position of the previous frame. Therefore, the target position information of this frame can be predicted, and the prediction result will be transmitted back to the CamShift algorithm tracking mechanism. CamShift algorithm starts the iterator to obtain the information of the current frame target.
5. After the occlusion disappears or the target is tracked again in the next frame, the tracking mechanism of CamShift algorithm continues, and the Kalman filter model continues to serve as the optimization assistant.
4. Evaluation criteria for algorithm performance
The performance evaluation of traffic incident detection algorithms usually requires certain indexes and methods. In general, the algorithm is evaluated from different aspects such as the validity of event detection and detection efficiency. In this paper, the detection rate $R_T$, the omission rate $R_F$ and the false alarm rate $R_A$ are selected as the criteria to evaluate the merits of the algorithm.

Detection rate $R_T$ represents the ratio between the number of vehicles detected by the algorithm in a video and the total number of vehicles appearing in the video. The formula is as follows:

$$ R_T = \frac{TP}{TP+FN} $$

where:
- $TP$ —— The number of vehicles detected;
- $FN$ —— Number of vehicles not detected.

In the video, the number of undetected vehicles plus the correct number of detected vehicles is the total number of vehicles appearing in the video.

The omission rate $R_F$ is the ratio of the number of undetected vehicles in a video to the total number of vehicles appearing in the video.

$$ R_F = \frac{FN}{TP+FN} $$

which is

$$ R_F = 1 - R_T $$

False alarm rate refers to the number of targets mistakenly detected as vehicles and the number of all vehicles detected, including but not limited to all vehicles appearing in the video.

$$ R_A = \frac{FA}{TP+FN+FA} $$

where:
- $FA$ —— Number of vehicles incorrectly detected

5. Experimental Analysis
The data set in this paper comes from the vehicle in VOC2007 data set, and the test platform is built in the MATLAB program. The test video is the test video released by Google and the computer display screen of the high-speed monitoring hall. Enter the test video into Matlab, run the code, and the video after processing is shown as follows:

As can be seen from Figure 4, when there are fewer vehicles, the tracking effect of the algorithm on a single vehicle can meet the requirements of the system. The distance between three vehicles is appropriate, and each vehicle can achieve real-time tracking and monitoring. Figure 5 shows that when there are many vehicles, the tracking effect is poor. For example, the vehicles on the left are blocked into the fleet, and the vehicles cannot move. In the background subtraction method, the vehicles become the background and cannot be detected. In the right lane, the vehicles are running normally,
but due to the instability of the overall picture, it is impossible to detect and track the normally moving vehicles on the right side.

![Figure 4. Simple traffic status CAMSHIFT tracking algorithm implementation](image1)

![Figure 5. Traffic jam CamShift tracking algorithm implementation](image2)

### 6. Test results

In this test, the three videos shot at the Lian-Xu high-speed were imported into the data processing system, and the output videos were analyzed after the system processing. The analysis result is to evaluate the two algorithms according to the evaluation criteria in Section 4.1.

|                  | TP | FN | FA | \( R_T \) | \( R_F \) | \( R_A \) |
|------------------|----|----|----|-----------|-----------|-----------|
| Video 1          | 43 | 41 | 17 | 51.2%     | 48.8%     | 15.3%     |
| Video 2          | 80 | 27 | 24 | 74.8%     | 25.2%     | 18.3%     |
| Video 3          | 53 | 42 | 13 | 55.8%     | 44.2%     | 12.0%     |

![Figure 6. Comparison diagram of results](image3)
Figure 6 is a data bar chart result comparison of the table results. Among the detection results of the three videos, the detection rate of video 1 and video 3 is similar, and the detection rate of video 2 is the highest. In the video 2, the road is four lanes in both directions, with emergency lanes, and the traffic flow is large and the road passes smoothly. Video 1 and Video 3 are also two-way four-lane, with emergency lane and moderate traffic flow. However, a traffic incident occurred in one lane, causing traffic jam on one side of the road. Vehicles are all parked on the road and block each other, so the detection system cannot identify the vehicles that block most of the body, so the detection rate is lower than that of normal video. In video 2, although the speed is fast and the traffic flow is large, the traffic flow is balanced and there is no jam. The vehicles appear on the road and the whole picture is displayed. Therefore, it is relatively easy for the algorithm to identify the vehicles in the video.

7. Conclusion
In this paper, based on the existing software and hardware of the Lian-Xu Expressway, the construction of the traffic incident detection system is completed, and the system requirements and system architecture are analyzed and designed, with the emphasis on the study of the data processing module. The two main steps of the data processing module are target detection and target tracking, and the workflow of the data processing module is designed. When the background subtraction method was used to construct the data processing module, because the ordinary background subtraction method could not adapt to the changing highway monitoring scene, this paper added Gaussian background modeling to update the model to adapt to the changes of light and so on, and increased the accuracy of the target detection algorithm. In this paper, Kalman filter is combined with CamShift moving target tracking algorithm to solve the short-time occlusion problem and prevent the loss of tracking target. The experimental results show that the detection system based on the background difference method can realize the detection and tracking of vehicles, and the requirements of the detection system can be fulfilled when the traffic flow is small and the road running condition is stable. However, when the traffic flow is large or the video display effect is shaky, the detection efficiency of the detection system based on background subtraction method is low.

References
[1] Moshe Levin, Krause Gerianne-M, Budrick James-A. Incident-detection algorithms. part 2. online evaluation[M]. 1979.
[2] Angshuman Guin. An Incident Detection Algorithm Based On a Discrete State Propagation Model of Traffic Flow[J]. 2004.
[3] D Srinivasan, Jin X, Cheu R-L. Adaptive neural network models for automatic incident detection on freeways[J]. NEUROCOMPUTING, 2005, 64473-496.
[4] Wang Qi. Freeway traffic incident detection based on neural network and support vector machine [D]. Southwest Jiaotong University, 2006
[5] Li Weimin, Wang Changfeng. Traffic incident video detection system in expressway management [J]. Journal of Guangxi Academy of Sciences, 2005, (S1): 86-88
[6] Yang Yuan. Traffic condition and abnormal traffic event detection based on target trajectory [D]. Chang'an University, 2014
[7] Wenhan Luo, Xing Junliang, Milan Anton, et al. Multiple Object Tracking: A Literature Review[J]. 2014.
[8] He Biao. Application of Automatic Video Traffic Incident Detection System on Chongqing Wankai Expressway[J]. China Transportation Information Industry, 2007, (08): 20-21.
[9] Liu Weining, Zeng Heng, Sun Dihua, etc. Traffic congestion discrimination model based on video detection technology [J]. Application Research of Computers, 2010, 27(08): 3006-3008.
[10] Wang Peng, Huang Kaiqi. Video-based night highway vehicle incident detection[J]. Journal of Image and Graphics, 2010, 15(02): 301-306.
[11] Jiang Xinchun, Liu Xiaoming. Analysis of the safety state of highway vehicle traffic behavior based on video detection[J]. Highway Transportation Science and Technology, 2010, 27(01): 122-125.

[12] Chen Zhijian. Freeway traffic incident detection algorithm based on wavelet SVM [a / / transformation and innovation of urban traffic development mode: Proceedings of China urban traffic planning 2011 Annual Conference and the 25th Symposium [C]. Wuhan, Hubei, China: 2014: 1363-1371

[13] Shaoqing Ren, He Kaiming, Girshick Ross, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2015, 39(6): 1137-1149.

[14] Ross Girshick, Donahue Jeff, Darrell Trevor, et al. Rich feature hierarchies for accurate object detection and semantic segmentation. 2014: 580-587.

[15] Gao Bingbing, Gaoge Hu, Xinhe Zhu, et al. A Robust Cubature Kalman Filter with Abnormal Observations Identification Using the Mahalanobis Distance Criterion for Vehicular INS/GNSS Integration.[J]. Sensors (Basel, Switzerland), 2019, 19(23).