Multi-Lane Detection and Road Traffic Congestion Classification for Intelligent Transportation System

Li-Wu Tsai, Yee-Choy Chean, Chien-Peng Ho, Hui-Zhen Gu and Suh-Yin Lee
Department of Computer Science
National Chiao-Tung University
Hsinchu, Taiwan
{trick1412.cs98g, ned.cs98g}@nctu.edu.tw, {cpho, hcku, sylee}@cs.nctu.edu.tw

Abstract—Intelligent Traffic Systems have been widely used for traffic monitoring on roadway, and it is one of the most practicable tools to provide the instant road traffic information for everyone needs it, especially mobile users that demand instant information of road traffic. When traffic congestion arises, if the vehicles get the traffic information earlier, they can choose recommend alternate routes to avoid the traffic jam. Therefore, we propose a traffic congestion classification framework to allow classification of congestion in traffic video sequences from real-time surveillance. The framework consists of three procedures: the first one is the roadway mask with bidirectional roadway analysis, and then the virtual detectors are set up for each lane without lane marking detection. The second step is to estimate the three traffic parameters: flow, speed and density by virtual detectors. In the last procedure, three traffic parameters are utilized to classify the traffic congestion into four levels accurately. The experiment results show that the framework can perform well on more complicated roadway types and simplified procedure of vehicle tracking to reduce the computational cost caused by complex algorithm.

Keywords—lane detection; virtual detector; traffic parameter; traffic congestion; intelligent transportation system.

I. INTRODUCTION

In recent years, surveillance systems have been widely used for monitoring the roadway. A complete roadway traffic video surveillance system can record all events of a specific region to tape or to disc, and it is useful to utilize the abundant traffic video data which can be processed for calculation purposes of algorithms applied in a real-time intelligent transportation application. Therefore, a great deal of significant research [1] has been focused on automatic traffic events analysis, such as traffic congestion, accidents and violation. In this paper, we investigate the event that is related to road traffic congestion caused by traffic flow, speed and density.

Vehicle detectors and other similar sensors are used to gather the traffic information in traditional traffic control system. However, more information on the surveillance cameras can be provided to support establishment of complete traffic information. With the rapid development of mobile devices, if the vehicles get the traffic information by mobile devices earlier from a robust intelligent surveillance system, they can choose recommend alternate routes based on the detailed traffic information to avoid the traffic jam.

Aiming at the benefits arising from integration of video analysis technique and intelligent mobile devices, we propose a traffic congestion classification framework that utilities the surveillance video to identify the congestion levels of traffic in real-time automatically. The traffic congestion levels are associated with some important traffic parameters, such as density, speed and flow. According to these three traffic parameters, the level of traffic congestion is classified into four levels: heavy, medium, mild and light.

As shown in Fig. 1, the proposed framework contains three major procedures. The first one is to find out the roadway region with bidirectional roadway analysis. For the purposes of real-time response and accuracy promotion, monitoring the regions that we are interested in is helpful. In addition, the virtual detectors are set up for each lane of roadway without lane marking detection. The second is to extract the moving vehicles. Only the moving vehicles that trigger the virtual detector are tracked on the roadway for estimating traffic parameters. In the third step, traffic parameters are utilized to make an accurate classification of traffic congestion levels for both directions of bidirectional roadway at the same time. We present the experiment result on the freeway surveillance videos, which demonstrate the accuracy of traffic congestion classification and real-time response of our framework.

II. RELATED WORKS

In traffic surveillance video, the region that we are interest in is roadway. Finding out the region in advance can reduce the computation of video processing and decrease the

![Figure 1. Block diagram of proposed framework.](image-url)
errors caused by noises. Li Bo et al. [2] proposed an algorithm based on Multi-resolution Hough Transform to detect the roadway by lane marking. For the reason that lane marking is not always visible, the authors [3] put forward a method to detect bidirectional roadway without lane marking detection. However, the previous methods are affected by unbalanced traffic flow and constrained by three roadway types only.

Moving vehicles extraction is an important technique in video processing. Segmenting the video frames into foreground and background components is useful to detect the moving vehicles. In [4], B. Chen et al. proposed a real-time background model initiation and maintenance algorithm. In [5], an algorithm was proposed to eliminate cast shadow on the basis of four general observations. Because most events of vehicles happened on the roadway are associated with their trajectories, individual vehicles are tracked over time. However, tracking all the moving vehicles on the roadway is time-consuming and complicated.

Along with vehicles detection and tracking, numerous works deal with the vehicles activity analysis. For some events detection, the features of vehicles, such as size, speed and moving direction, can help to understand what happened in the surveillance video [3]. Other more complex events are mostly detected with some machine learning algorithms [6, 7]. For the purpose of traffic congestion analysis, the usage of virtual line detectors was proposed in [8, 9, 10]. This idea is more suitable to analyze traffic congestion. Nevertheless, a limitation of their algorithm is that it is hard to eliminate the shadow and filter noise.

In the present paper, we propose a roadway detection and bidirectional roadway analysis method to conquer the problem of unbalanced traffic flow and the limitation of three roadway types. In addition, the virtual detectors are set up for each lane to simplify the procedure of vehicle tracking and shadow elimination for traffic parameters estimation. This approach significantly reduces the cost caused by other complex algorithm. Finally, the traffic parameters are utilized for traffic congestion classification in real-time surveillance video.

III. INITIALIZATION PROCEDURE

A. Roadway Detection

For the scenes in traffic surveillance videos, the region of our interest is only the roadway. The region is an area where most movements occur. We accumulate the difference between two consecutive frames to obtain the position and shape of roadway. The difference image $D$ is defined as follows:

$$D(x, y) = D^{i-1}(x, y) + \begin{cases} 1, & \text{if } |F^i(x, y) - F^{i-1}(x, y)| > T_d \\ 0, & \text{otherwise} \end{cases}$$

where $F^i(x, y)$ denotes the current frame $i$ and $T_d$ is the predefined value for identifying the pixels with movements. After accumulation of difference for a sequence of frames, the pixels with difference are regarded as the candidate $R_c$ of roadway.

$$R_c(x, y) = \begin{cases} 1, & \text{if } D(x, y) > 0 \\ 0, & \text{otherwise} \end{cases}$$

And then, remove the noise by density filtering in (3).

$$\frac{1}{(2k+1)^2} \sum_{i=j-k}^{j+k} \sum_{j=i-k}^{i+k} R_c(i, j) < T_k$$

where $k$ control the size of filter and $T_k$ is a threshold determined by the characteristic of roadway. Currently, 0.6 is used. In $R_c$, order the isolated components in descending according to their size from $S_1$ to $S_n$ and choose the first $K$ component as the roadway, where

$$K = \arg \min \sum_{i=1}^{n} \frac{S_i}{S} > T$$

B. Bidirectional roadway analysis

Most surveillance camera can capture images at specified angle and scale range that contains multiple lanes of traffic in both directions. The bidirectional roadway analysis can be applied to monitor the traffic situation for both directions separately. During the period of difference accumulation in roadway detection, the occurrence of differences means some motions appear at the same time. Accumulation of those motions can approximately reveal the direction of
roadway. Hence, we estimate and accumulate the motion vectors for those pixels with difference in roadway detection. Motion image \( M \) is defined as follows:

\[
M'(x,y) = M^{i-1}(x,y) + \begin{cases} 
1, & m_i'(x,y) > y \\
-1, & m_i'(x,y) < y 
\end{cases}.
\]  

(5)

where \( m \) denotes the motion vector in \( y \)-axis between frame \( i \) and \( i-1 \). After the accumulation of motion, the \( M(x,y) \) is simplified into two kinds of direction: 0 is for DOWN and 1 is for UP.

\[
M(x,y) = \begin{cases} 
0, & M(x,y) > 0 \\
1, & M(x,y) < 0 
\end{cases}.
\]  

(6)

In order to separate the roadway into two directions, the positions of both directions are decided by the average \( x \)-position of two kinds of motion.

\[
X_{\text{DOWN}} = \frac{1}{N_{\text{DOWN}}} \sum_{(x,y) \in R} x, \quad \text{if } M(x,y) = 0.
\]  

(7)

\[
X_{\text{UP}} = \frac{1}{N_{\text{UP}}} \sum_{(x,y) \in R} x, \quad \text{if } M(x,y) = 1.
\]  

(8)

where \( N_{\text{DOWN}} \) and \( N_{\text{UP}} \) are the number of pixels of motion DOWN and UP in roadway mask respectively. If \( X_{\text{DOWN}} \) is smaller than \( X_{\text{UP}} \), the moving direction of the left side of roadway is DOWN and the right side is UP, otherwise the left side is UP and the right side is DOWN, as illustrated in Fig. 3(a).

And then, a center line \( Y \) for separating bidirectional roadway is evaluated. Each pixel at the top boundary of roadway mask is considered as a core, and rotates the line \( Y \) from left boundary to right boundary of roadway mask for each core. For each rotation, we evaluate the error rate of motion classification that base on the center line \( Y \), and the rate of recall of classification is also considered. The error rate \( E \) is small, the separation is better. The \( Y \) with minimum \( E \) is considered as the actual center line.

\[
E = \frac{E_{\text{left}} + E_{\text{right}}}{R_{\text{left}} + R_{\text{right}}},
\]  

(9)

where \( E_{\text{left}} \) is error rate of classification in left side of \( Y \), \( E_{\text{right}} \) is error rate classification in right side of \( Y \), \( R_{\text{left}} \) is recall of classification in left side of \( Y \), \( R_{\text{right}} \) is recall of classification in left side of \( Y \). Using this line, the roadway mask are redefined as bidirectional roadway mask. An example of bidirectional roadway mask and the result of center line evaluation are shown in Fig. 3(b).

C. Virtual Detectors Setting

The virtual detectors are set up on each lane for traffic information collection. The appropriate positions for detector are the center between two lane markings. However, if the type of lane marking is a dashed-like line or the lane marking is not visible, the lane markings always fails to be detected. Based on the observation, we know the center point of lane usually is the center point of moving vehicles. So the center point of moving vehicles is retrieve to detect the center of lane.

To obtain the center point of moving vehicle, vehicle extraction based on background subtraction technique, which is described in section IV, is used. Every vehicle will be identified by bounding box after extraction. The center points of vehicles and the average width of vehicles are gathered from the bottom line of bounding box. Figure 4(a) shows an example of center points of moving vehicles. After collecting all center points of vehicles, the Modified Basic Sequential Algorithm Scheme (MBSAS) [11] clustering algorithm is used to cluster these center points at every row of video frame. The average width of vehicles is the threshold for MBSAS. The mean values of clusters are also the center points of each lane. The positions of virtual detectors are determined by those center points.

A detection row \( d \) is regarded as an expected row of the virtual detector. Let \( N_i \) is the number of cluster of row \( i \). From \( N_{\text{left}} \) to \( N_{\text{right}} \), choose the value \( N \) with max count as the number of virtual detectors, and select the row with \( N \) lanes and closest to \( d \) as the actual detection row. Finally, we set the virtual detectors on mean values of clusters at detection row and identify their monitoring direction of traffic flow by center line \( Y \) which is calculated in bidirectional roadway analysis, as illustrated in Fig. 4(b).

Figure 3. The example of bidirectional roadway analysis: (a) motion image \( M \). The pixels with motion DOWN is colored white; the pixels with motion UP is colored gray. The pixels without motion are colored black; (b) Bidirectional roadway mask and the result of center line evaluation.

Figure 4. The example of virtual detectors setting: (a) The illustration of center points of moving vehicles; (b) Virtual detectors are set up for each lane. The detectors with different shapes monitor different directions. (The red triangles monitor the traffic flow DOWN; the green rectangles monitor the traffic flow UP)
IV. VEHICLE EXTRACTION

For the characteristic of the surface roadway is stable, the background subtraction method is an appropriate way to segment the foreground image for the advantages of integrity of information and low computation. So, in this paper, mixture of Gaussians background model [12] is applied to extract the vehicles on the roadway.

Shadow elimination base on gradient feature [13] and color reflectance [14] are used to avoid that vehicles connect each other by shadow. Finally, cluster the adjacent pixels in foreground as a single vehicle. The bounding boxes mark every extracted moving vehicle.

V. TRAFFIC PARAMETERS ESTIMATION

In the proposed framework, three traffic parameters: flow, speed and density are needed simultaneously to analyze the traffic congestion. Based on the bidirectional roadway analysis, the traffic parameters can be estimated for both directions of roadway separately.

A. Flow

The virtual detectors on each lane of the roadway are for counting the number of vehicles that trigger the virtual detectors per unit of time. The vehicle trigger the virtual detector when it pass through the detector and the foreground pixels of the vehicles occupy at least the quarter of the triggered virtual detector. This condition is for reducing the erroneous judgment caused by noises. The ratio can be changed with the quality of surveillance videos.

For the usage of virtual detectors, the vehicles just are tracked when they trigger the virtual detector, so the vehicle tracking on whole the roadway is unnecessary. A key is that only one virtual detector on each lane instead of just one virtual detector for all the lanes. This approach will simplify the vehicles matching in vehicle tracking procedure. Because only one vehicle can occupy the virtual detector at the same time in normal situation, we match the vehicles occupying the same virtual detector in two consecutive video frames for determining whether the two vehicles are the same. In our framework, the color histograms of vehicles are used to match the vehicles. Flow \( n \) is defined as the number of moving vehicles that trigger the virtual detectors.

B. Speed

The speed of a vehicle is estimated when the vehicle triggers the virtual detector. In principle, a slower moving vehicle will trigger the virtual detector for more consecutive frames. Hence, the speed approximation can be done by counting the number of frames that a single vehicle triggers the virtual detector. Speed \( S \) is defined as follow:

\[
S = \frac{fps}{n} \sum_{i=1}^{f_i}
\]

where \( f_i \) is the number of frame of a vehicle \( i \) trigger the virtual detector, \( n \) is flow and \( fps \) stands for frame per second of the surveillance video.

C. Density

After segmenting out all the pixels of foreground image in current frame, the traffic density are calculated based on the ratio between the number of all pixels of foreground and the number of pixels of roadway. To reduce the influence of vision depth, the ratios are calculated row by row. The average of whole ratios is the density \( D \) as follow:

\[
D = \frac{1}{n} \sum_{i=1}^{r_i} \frac{P_{ji}}{P_{ri}}
\]

where \( r \) is the height of roadway, \( P_{ji} \) is the number of foreground pixels of row \( i \), \( P_{ri} \) is the number of row of roadway \( i \).

VI. TRAFFIC CONGESTION LEVEL CLASSIFICATION

For a sequence of video frames in a time interval, traffic parameters: speed and density are used to make traffic congestion levels classification. The traffic congestion levels are classified into four conventional levels: heavy, medium, mild and light. The congestion evaluation equation in (12) calculates the value \( C \) that stands for the degree of traffic congestion on the roadway. The higher the value is, the more congestive the traffic is. The threshold of the value for congestion level classification is determined by training data in advance. The threshold between two levels is the average value of means of two levels. The design of evaluation of congestion is based on the relation between traffic parameters.

\[
C = D \frac{S}{S}
\]

VII. EXPERIMENTS

The proposed framework for traffic congestion level classification has been examined on freeway surveillance videos that are captured from different surveillance camera [12]. As to the position of virtual detecting line, it could be chosen manually along with the different captured scene. In the experiments, we verify the accuracy of classification for these conditions: heavy, medium, mild and light.

These are 147 roadways with single direction from traffic surveillance video in our experiment. The length of one clip of video has 60 seconds and the size of 352 × 240 frame. The traffic parameters are calculated for every frame, and average those values in 60 seconds to evaluate the

<table>
<thead>
<tr>
<th>Ground Truth</th>
<th>Light(56)</th>
<th>Mild(35)</th>
<th>Mid.(26)</th>
<th>Heavy(30)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Light(56)</td>
<td>52</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Mild(35)</td>
<td>4</td>
<td>30</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Mid.(26)</td>
<td>0</td>
<td>1</td>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>Heavy(30)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>30</td>
</tr>
</tbody>
</table>

Accuracy: 0.93 0.86 0.96 1.00
congestion degree. We get the 93.2% average accuracy for the classification. The confusion matrix of congestion classification is shown in Table I.

Examples of roadway detection with bidirectional analysis are shown in Fig. 5. The scenes which contain the interchange also can be analyzed by our algorithm. Furthermore, the examples of virtual detectors setting are shown in Fig. 6, which displays the color bar in each lane is the virtual detector. The positions of virtual detectors are not actually in the center of lane. There is some inaccuracy caused by behaviors of driving.

For the experiment, we use the computer with AMD 2.8 GHz dual-core CPU. The processing of one video frame needs around 0.1 second. On the whole, the proposed framework works well with the traffic surveillance video that provides at least 5 frames per second, so using the key frames of video are enough to accomplish the classification.

VIII. CONCLUSION

A traffic congestion classification framework with three procedures is described and examined in traffic surveillance video with different traffic congestion levels. Automatic roadway detection, bidirectional roadway analysis and virtual detector setting method are proposed to overcome the unbalanced traffic flow and roadway-type limitations. In addition, simplified procedure of vehicle tracking for traffic parameters estimation significantly not only reduces the cost caused by other complex algorithm, but also solves the difficulty of shadow elimination of previous works. Based on the methods described above, we can enhance the ability of the traffic parameters estimation. Subsequently, we describe a method of optimizing these parameters to maximize the accurate classification of traffic congestion in real-time traffic surveillance video for intelligent transportation system. In the future, the sequence of congestion classification results can be utilized for instant analysis with advanced data mining technique.

REFERENCES