

Research Article

A Calculation Method for Vehicle Movement Reconstruction from Videos

Hao Feng,¹ Weiguo Shi ,² Feng Chen,² Young-Ji Byon,³ Weiwei Heng,¹ and Shaoyou Pan¹

¹Academy of Forensic Science, Shanghai 200063, China

²Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, 4800 Cao'an Road, Shanghai 201804, China

³Department of Civil Infrastructure and Environmental Engineering, Khalifa University of Science and Technology, P.O. Box 127788, Abu Dhabi, UAE

Correspondence should be addressed to Weiguo Shi; shiweiguo@tongji.edu.cn

Received 21 May 2020; Accepted 5 June 2020; Published 23 June 2020

Academic Editor: Kun Wang

Copyright © 2020 Hao Feng et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This paper proposes a new enhanced method based on one-dimensional direct linear transformation for estimating vehicle movement states in video sequences. The proposed method utilizes a contoured structure of target vehicles, and the data collection procedure is found to be relatively stable and effective, providing a better applicability. The movements of vehicles in the video are captured by active calibration regions while the spatial consistency between the vehicle's driving track and the calibration information are in sync. The vehicle movement states in the verification phase are estimated using the proposed method first, and then the estimated states are compared with the actual movement states recorded in the experimental test. The results show that, in the case of camera perspective of 90 degrees, in all driving states of low speed, high speed, or deceleration, the error between estimated speed and recorded speed is less than 1.5%, the error of accelerations is less than 7%, and the error of distances is less than 2%; similarly, in the case of camera perspective of 30 degrees, the errors of speeds, distances, and accelerations are less than 4%, 5%, and 10%, respectively. It is found that the proposed method is superior to other existing methods.

1. Introduction

As the video image records the vehicles' motions within the monitoring range and can provide objective raw data for further quantitative analysis, video surveillance equipment has been applied to provide important clues for detection and forensics of vehicle-related applications. For instance, Byon et al. [1–3] used the vehicle state, which was identified in terms of speed and acceleration on the associated road section from video recording, as input information to transportation mode detection algorithms. Tarkowski et al. [4] used video recording and reconstructed a motorcycle-motor car collision with both vehicles moving in the same direction and performing simultaneous maneuvers: a left turn and overtaking. Wu et al. [5] calculated the wheel track on an operational highway by video analysis when analyzing the design of visual intervention pavement markings.

Vehicles' speed has a significant impact on severity of road traffic accidents [6–9], and a reliable method to extract valid data and reconstruct vehicle movement is the critical success factor on accident analysis by videos. There have been several traditional techniques such as background subtraction and vehicle segmentation using hand-crafted features that are sensitive to noise. For instance, the optical flow method [10], interframe difference method [11], background difference method [12] resulting from the change of pixels for detecting moving objects and their associated speeds, and color method [13] usually used in the field of video velocimetry. Gao et al. [14] found the method of vehicle speed identification based on video images which helps to detect the vehicle moving distances in accordance with the fixed object and the structure of the target vehicle in the road scene environment. According to the definition of interval running speed and

point speed, the speed of the target vehicle can be estimated with a video frame rate. At present, the method has become the main method of using video image for a vehicle's speed identification in the practice of forensic identification of traffic accidents [15–19], but due to the location of the calibration and the video frame rate, the method cannot fully reflect all motion states of the target vehicle in the monitoring range. Kumar et al. [20] presented a semiautomatic 2D solution for vehicle speed estimation from monocular videos, and they employed the state-of-the-art object detector Mask-RCNN to generate reliable vehicle bounding boxes and propose a two-stage algorithm to approximately transform the measurements in the image domain to the real world. However, none of these methods are able to estimate the vehicle trajectory (especially the trajectory of the curve) under the view except the case of videos recorded from overlooking perspectives. Furthermore, low speed and multiobjective situations will lead to larger detection errors.

The combination of photogrammetry and computer technology provides a better method for road geometry measurement and vehicle speed calculation [21]. The transformation relationship between the image point coordinate system and the object space coordinate system can be established by using a direct linear transformation principle in the close-range photogrammetry technique [22], which is suitable for photographic measurements of images taken by currently dominant nonmeasurement cameras [23, 24]. Yang et al. [25] collected information of pavement and vehicle body in traffic accident scenes based on the direct linear transformation principle of close-range photogrammetry, which can be used as the input condition of software for accident simulation. Through the road surveillance video, Miao [26] and Han [27] selected four fixed points on the road markings as control points. By matching the control points of the image plane coordinates and the object plane coordinates combined with the direct linear transformation principle, the track of a target vehicle on the ground and the associated speed of the vehicle can be estimated. However, the method takes four fixed points on the road as control points, and the selection of road control points is often affected by the number and distribution of road marks and road grade [28]. And the relative distances among the four feature points need to be measured where accidents would occur, resulting in significant limitations with respect to the application conditions and the required precisions in the set-up. Bullinger et al. [29] presented a method to reconstruct three-dimensional object motion trajectories in stereo video sequences, and they embed the vehicle trajectories into the environment reconstruction by combining the object point cloud of each image pair with corresponding camera poses contained in the background SfM (Structure from Motion) reconstruction. This method proved more reliable than monocular trajectory reconstruction approaches, but the invalid stereo matches caused by strong surface reflections are not easy to avoid. Hasith et al. [30] used a detector to detect the object

bounding boxes in an image frame and associated these detections to tracks using the Hungarian algorithm based on pairwise dissimilarity scores calculated between detections in the current frame against the tracks in the memory and proposed a method to address MOT (Objective of multiple object tracking) by defining a dissimilarity measure based on object motion, appearance, structure, and size. It is an efficient and advanced framework tracking accuracy, but the tracking accuracy needs to be improved.

Overall, the traditional and existing methods of estimating the vehicle motion cannot fully reflect the driving states of the vehicle, or the required conditions are greatly restricted or do not have enough accuracy. This paper develops a one-dimensional direct linear transformation method to solve the entire trajectory of a vehicle in the image sequence of recorded video, by capturing linear structure features of vehicle body and using the trajectory of the vehicle as calibration information. The calibration information of the method is derived from the profile of a target vehicle. The data collection process is found to be stable and effective, which eliminates the dependence, as in other existing methods, on other information such as fixed distances in the image, and the requirement for the quantity and distribution of the road markings. The calibrated area covers the whole motion of the vehicle in the video image, which keeps the consistency between traffic trajectory and calibration information space as much as possible. Based on video frame rates, the minimum step length of the entire vehicle trajectory is used to calculate the distance, which reduces the phenomenon of external interpolation during the solving process and improves the accuracy of results.

2. Basic Assumptions and Calculating Methods

In order to discretize the vehicle's continuous movement from a recorded video with certain image sequences, the video frame rate (i.e., recording frequency) and the vehicle's position in each image frame are used to compute the change in displacement during each time interval between each consecutive pair of image frames.

2.1. Basic Assumptions. Because of the continuity of a vehicle's movement and the high video frame rates commonly used in the emerging camera devices, it can be assumed that the moving trajectory within the two adjacent image frames are approximately a linear motion. Therefore, the corresponding displacement between adjacent frames of the video can be assumed to be equal to its actual distance, as shown in Figure 1.

2.2. Solution Ideas. A direct linear transformation does not need interior and exterior orientation elements of the camera, and it is suitable for photographic measurements of images taken by nonmeasurement cameras. The basic formula for solving the direct linear transformation is shown in the following:

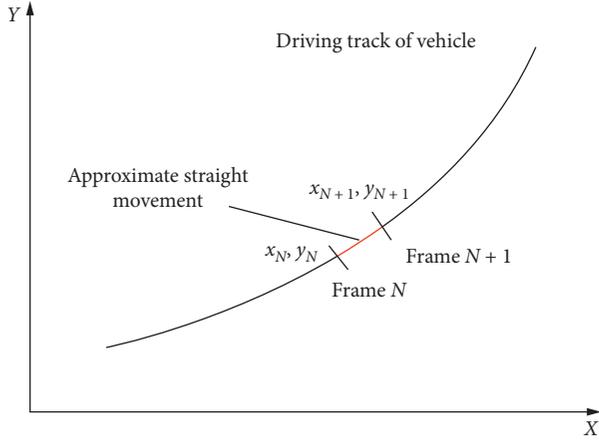


FIGURE 1: Approximate linear motion of the vehicle between adjacent image sequences.

$$\left. \begin{aligned} x + \frac{l_1 X + l_2 Y + l_3 Z + l_4}{l_9 X + l_{10} Y + l_{11} Z + 1} &= 0 \\ y + \frac{l_5 X + l_6 Y + l_7 Z + l_8}{l_9 X + l_{10} Y + l_{11} Z + 1} &= 0 \end{aligned} \right\}, \quad (1)$$

where x and y are the image plane coordinates, and X , Y , and Z are object space coordinates. And $l_{1\sim 11}$ are the direct linear transformation factors.

It is assumed that two adjacent frames can be approximated to form a linear motion, so the linear feature of the vehicle body profile can be considered to be on the same line in two adjacent frames. To make the variable Z and Y as constants, equation (1) can be rewritten as

$$x + \frac{l_1 X + l_2}{l_3 X + 1} = 0. \quad (2)$$

Equation (2) is a one-dimensional direct linear transformation formula, which can solve the one-dimensional linear transformation problem of the image space into the object space. According to Figure 2, four intersections where the line at the outer end of the wheel axis meets the rim edge in the $No. N$ frame images are set as reference points, named a_N , b_N , c_N , and d_N . The four marked points make up three segments adjacent to a straight line. Then, any three of them are selected as control points. And based on the distance of the pixels and the actual distance between them, one-dimensional direct linear transformation coefficients $L_N = [l_1, l_2, l_3]$ of $No. N$ frame image are solved. On the next frame image ($No. N+1$ Frame), combined with the image point coordinates of a_{N+1} and a_N and the direct linear transformation coefficients L_N , the object space distance of a_{N+1} and a_N are calculated.

Then, we can find the $No. N+1$ frame image's corresponding one-dimensional direct linear transformation coefficient l_{N+1} in accordance with the image distance and the actual distance of any three points of a_{N+1} , b_{N+1} , c_{N+1} , and d_{N+1} . Similarly, depending on the image point coordinates a_{N+2} and a_{N+1} of the next frame image $No. N+2$, combined with the direct linear transformation

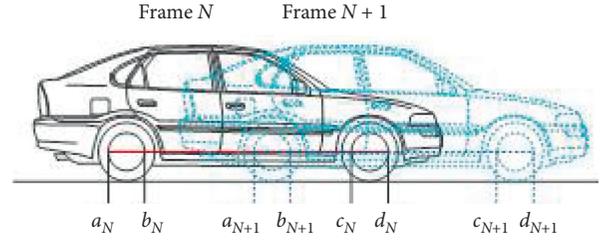


FIGURE 2: Distance calibration between adjacent image sequences.

coefficient l_{N+1} , the actual distance of a_{N+2} and a_{N+1} can be calculated. Calibration and calculation can be continued frame by frame in the direction of the vehicle's movement. Finally, the driving state of the vehicle can be obtained.

3. Calculating Procedure

The solution procedure described in the "Solution Ideas" section can be summarized in the following steps:

- (1) Extract the image and time-stamp of the target vehicle in the desired duration of video frame by frame.
- (2) In each frame, four intersection points are formed between the straight line at the outer end of the wheel axis on one side of the target vehicle and the rim edge as the reference points, respectively set to a , b , c , and d .
- (3) With the $No. N$ frame image as an example, take three points a , b , and d (any three-point combinations) as control points and calculate distances l_{ab}^N and l_{bd}^N for pixel plane coordinates, followed by combining with its corresponding actual distance. Then, solve the one-dimensional direct linear transformation coefficient l_k^N of the image line and the actual object line.
- (4) Based on the image point coordinate a_{N+1} of $No. N+1$ frame image, combining with the direct linear transformation coefficients l_k^N for $No. N$ frame image, calculate the actual motion distance of point a from $No. N$ to $No. N+1$ frame image.
- (5) On $No. N+1$ frame image, take a , b , and d as control points, based on three-point image plane coordinates, compute distances l_{ab}^N and l_{bd}^N , combining with its corresponding actual distance. Then, solve the one-dimensional direct linear transformation coefficient l_k^{N+1} of the image line and the actual object line.
- (6) Based on the image point coordinate of a_{N+2} in $No. N+2$ frame image, combining with the direct linear transformation coefficient l_{k+1}^N , calculate the actual motion distance between $No. N+1$ and $No. N+2$ frame images.
- (7) Calculate the direct linear transformation coefficient l frame by frame by the method above, and set point a as the reference point of the actual driving distance of the target vehicle.

- (8) According to the driving distance of the target vehicle in the image sequence, combining with the frame rate of the video, calculate the vehicle's speed and acceleration corresponding to the specified images, frame by frame.
- (9) Output file information, vehicle feature point spacing, extracted video image range, image sequence and time-stamp information, feature point image coordinates on each frame, one-dimensional direct linear transformation coefficient l_k^N , the distance of feature point a from one frame to the next frame in the video, and speed and acceleration in every frame.

4. Design of Validation Test Scheme

This section describes the validation test set-up as the following:

- (i) Test material: a manual transmission car, Kistler S-350 photoelectric five-wheel instrument, LED monitors, two cameras, SIRIUS digital collector, a notebook computer (for channel setup and data collection), three-dimensional laser scanners, checkerboard, tape, and BMW label.
- (ii) Test site: a 150-m-long straight lane.
- (iii) Test scenario: Figure 3(a) shows test vehicles and equipment, and 3(b) shows a test scenario. There are six BMW labels on both sides of the test lane, and two cameras on the side of the driveway. The camera directions are set in the direction of the vehicle at 90° and 30° . The test vehicle is equipped with laser five-wheel instrument, digital collector, LED monitor, and car camera. 90° direction camera is 10 m far away from the side of the lane, and the test car runs three times at different driving speeds. DWDEsoft saves the collected video and data information to the notebook computer. The frame rate of the video collected in the test is 30 fps, and the resolution ratio is $640 * 480$ pixels. Collected data include travel speed and distance (collection frequency is 100 Hz).

5. Solution and Analysis

5.1. Low-Speed Running

5.1.1. In the Camera Perspective of 90 Degrees

(1) *Image Preprocessing and Inter-Frame Displacement Solution.* In PC-Rect 4.2, the image of the checkerboard taken by the test camera is processed and the lens distortion correction file is obtained. In MATLAB, the collected video is decomposed into multiple frames, the video frame rate is 30 fps and the total number of frames is 73, and the file format is in RGB24 bit. Then, 28 continuous frames are extracted when the test car passes through the video region. The image is corrected by applying lens distortion processing. As shown in Figure 4 (the original image) and Figure 5 (the corrected image), the corrected lane edge line is

approximately parallel to the reference line. The corrected images are renumbered in the order of 0 to 27.

Analytical process of the No. 0 frame image is summarized as in the following:

- (i) First of all, make a gray-scale change of the image of three-dimensional matrix information and translate it into a two-dimensional matrix. Calibrate the vehicle contour of three feature points, as shown in Figure 6. From left to right, there are three points of a , b , and d (a point corresponds to rear hub back-end point A of the vehicle, b point corresponds to rear hub front end point B , and d point to front hub front end point D).
- (ii) Second, the image plane coordinate system is established in the upper left corner of the image as the coordinate origin. Extract pixel coordinate values for three feature points $a = (18, 331)$, $b = (39, 332)$, and $c = (166, 332)$ in MATLAB. The ordinate values are slightly different, but in the vertical view, equation (2) does not involve ordinate operations during the calculation and hence can be ignored. The measured test axle distance is 280.3 cm, hub diameter is 44 cm, with a point (the test vehicle rear hub rear end point) for the coordinate origin set up for the test vehicle's right body plane coordinate system (simply referred to as the object square coordinate system), then object coordinate values of A , B , and D , three points are $A = (0, 0)$, $B = (44, 0)$, and $D = (324.3, 0)$
- (iii) At the same time, through the one-dimensional direct linear transformation equation (2):

$$J \cdot L_N + K = 0, \quad (3)$$

$$\text{where } J = \begin{bmatrix} X_1 & 1 & x_1 X_1 \\ X_2 & 1 & x_2 X_2 \\ X_3 & 1 & x_3 X_3 \end{bmatrix}, L_N = \begin{bmatrix} l_1 \\ l_2 \\ l_3 \end{bmatrix}, \text{ and } K = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Put above three points' pixel coordinate values and the object-square coordinate values in the formula and obtain the matrix of coefficient $L_0 = [-0.48 \quad -18.00 \quad 0]^{-1}$.

- (i) For No. 1 frame image, the same method is obtained on the frame image by a , b , and c . Three-point pixel coordinates are $a = (33, 331)$, $b = (55, 332)$, and $c = (187, 331)$.
- (ii) Relative to No. 0 frame image, three points in frame image moves in the x direction by approximately one pixel, which is approximated by a straight-line motion and consistent with the assumption. Put a point x -axis coordinate value $x_a = 33$ in equation (2), then get point A in No. 1 frame of object horizontal axis value $X_a = 31.25$ cm. Therefore, it can be considered that the point A from the No. 0 frame to the No. 1 frame moved horizontally 31.25 cm in the object space. The test vehicle has run $S(A_1 - A_0) = 31.25$ cm. Also, according to equation (3), you can get the No. 1 frame transformation coefficient matrix $L_1 = [-0.51 \quad -33 \quad 0]^{-1}$.



FIGURE 3: Test vehicle, equipment, and scene.



FIGURE 4: Original images taken by the video camera.



FIGURE 5: Images corrected in distortion.



FIGURE 6: Three feature points calibrated on vehicle exterior contour.

- (iii) With the same method to all frame image processing, calculate the displacement between two adjacent frames of the test vehicle for dS (called

frames displacement in the following) and the results are shown in Table 1.

(2) *Driving Speed Solution.* From Table 1, the frame displacement dS value is in the range of 33–38 cm, with a mean of 34.84 cm. The video frame rate is 30 fps, and according to $v = dS/dt$, we solve the interframe velocity v , as Figure 7 shows. Meanwhile, by the SIRIUS data collection, we record the test speed v_0 value at a collection frequency of 0.01 s, v_0 and v comparison analysis as shown in Figure 8.

From Figures 7 and 8, the time t fourth-order fitting curve of calculated speed v is basically the same as that of the test-gathering speed curve, and the overall trend is decreasing. Standard deviation of $v(\mu + \sigma)$ ranges between 36.3 km/h and 39.3 km/h, with a mean of 37.7 km/h. Recorded v_0 ranges between 37.2 km/h and 38.9 km/h, included in the range of 36.3 km/h~39.1 km/h. The average speed is 38.1 km/h, which is different from the calculated average v_0 approximately by 1.0%. This shows that the proposed method is valid for speed calculation.

(3) *Acceleration Solution.* In order to calculate acceleration a , we first fit to fourth-order fitting polynomial on v ,

$$v = 25t^4 - 48t^3 + 25t^2 - 4t + 38. \quad (4)$$

Then, we take the derivative equation (4) resulting in acceleration a :

$$a = 100t^3 - 143t^2 + 50t - 4. \quad (5)$$

At the same time, since the data collection instrument does not record the acceleration value over time. This paper takes the derivative of v_0 with respect to t and gets a scatter acceleration a_0 . Then, we fit a third-order polynomial from the scatter value a_0 , as shown in Figure 9.

There is a large deviation between the a_0 fitted curve and calculated a . It is also noted that the distribution of a_0 is scattered. Because the change of test speed is small ($<2 \text{ m/s}^2$), close to a uniform motion, while sampling period is as short as, the speed error reflected on the acceleration is amplified 100 times by the sampling frequency.

We notice the time T is only 0.9 s and the change in vehicle acceleration is usually approximately linearly variable under nonemergency situations of 0.9 s. Therefore, it is reasonable to make a linear comparison of the acceleration values, as shown in Figure 10. The differences between two slopes of the two fitted lines and mean values are 7.6% and 5.1%, respectively.

TABLE 1: Calculation results of the displacement between 28 image sequences (unit: cm).

No. n frame image	1	2	3	4	5	6	7
$dS(A_n - A_{n-1})$	35.38	34.55	35.71	34.88	32.80	37.60	33.59
No. n frame image	8	9	10	11	12	13	14
$dS(A_n - A_{n-1})$	35.96	33.97	37.19	36.12	34.95	36.41	34.33
No. n frame image	15	16	17	18	19	20	21
$dS(A_n - A_{n-1})$	33.35	34.25	34.40	36.01	34.96	35.58	34.56
No. n frame image	22	23	24	25	26	27	—
$dS(A_n - A_{n-1})$	34.65	35.25	31.60	34.84	34.09	33.62	—

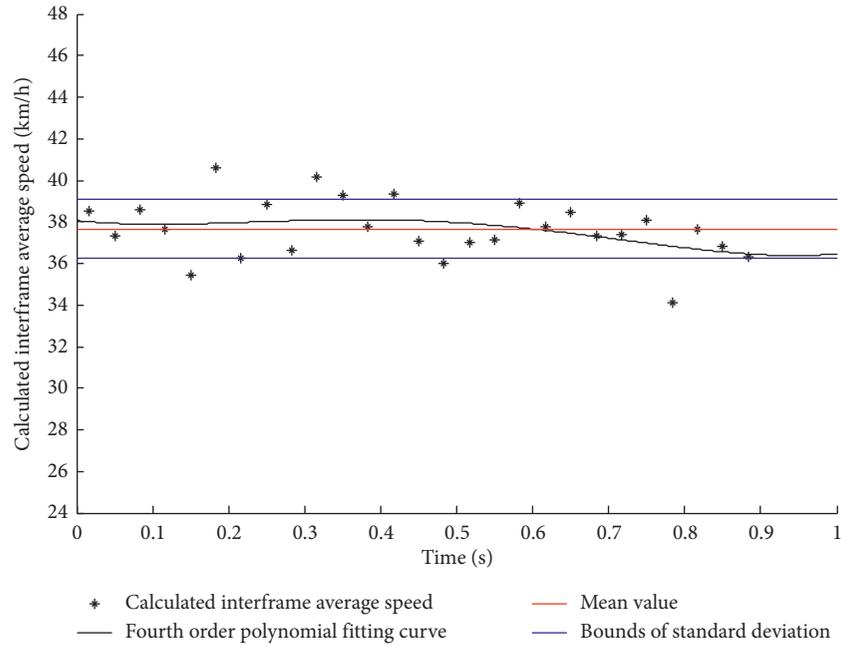


FIGURE 7: Calculated average speed and its fitting curve.

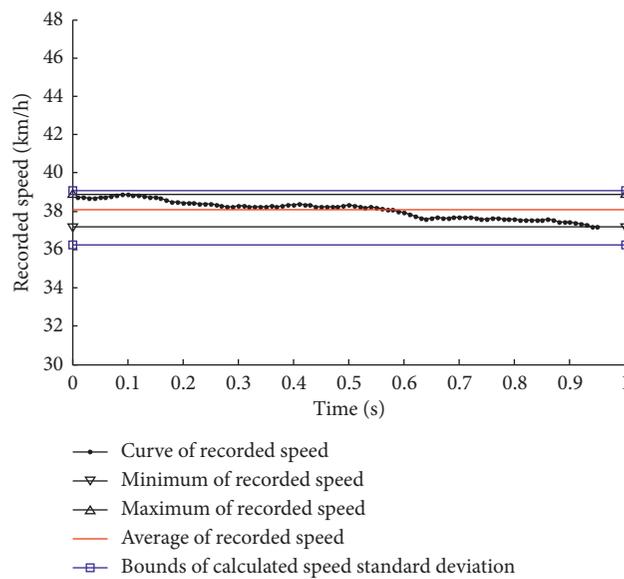
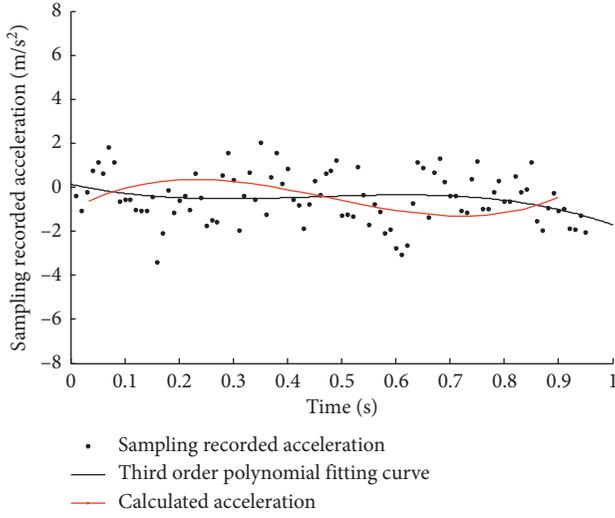
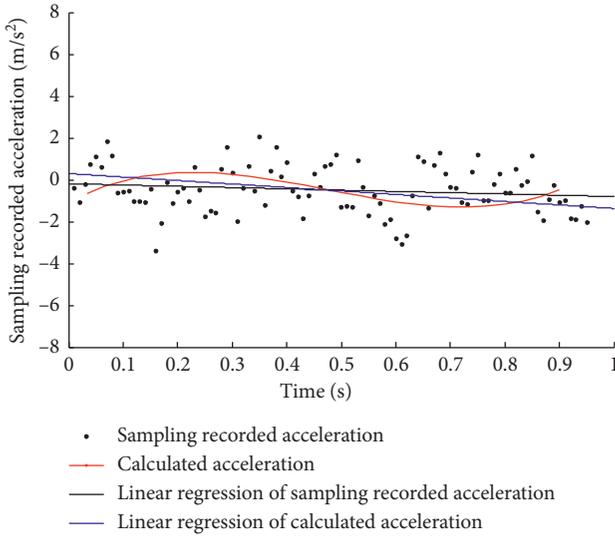


FIGURE 8: Compared analysis between v_0 and v .

FIGURE 9: Compared analysis of a and a_0 .FIGURE 10: Linear regression of a and a_0 .

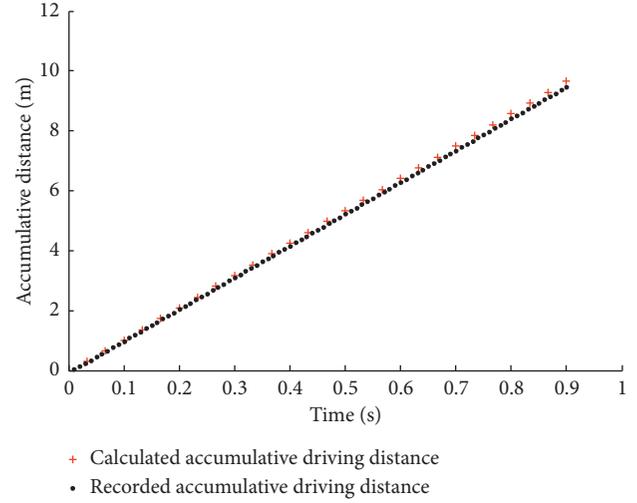
From this result, we can see that this method is also applicable to the solution of the mean acceleration value.

(4) *Travel Distance Calculation.* The travel distance S is found by taking an integral of the frame displacement dS about T . Figure 11 shows the calculated driving distance S by this method compared with test recording S_0 .

S and S_0 reflect that the car is approximately under uniform motion and the linear regression equation is

$$\begin{aligned} S &= 10.8t - 0.07, \\ S_0 &= 10.6t - 0.07. \end{aligned} \quad (6)$$

The two slopes and intercepts are similar. The cumulative driving distance in 0.9 s results in $S = 9.65$ m, and the driving distance of the test record is $S_0 = 9.47$ m, between which the difference is 1.9%. It is found that the calculation of the driving distance is valid.

FIGURE 11: Compared analysis of S and S_0 .

5.1.2. In the Camera Perspective of 30 Degrees

(1) *Value Correction of Feature Point Coordinate.* After the video frame solution and distortion correction processing are applied, we extract 42 continuous frame images of the test vehicle. The images are renumbered in the order of 0 to 41.

Due to the 30° of deflection angle between the lens orientation and the driving direction of vehicle, it is difficult to ensure that the feature points of the body of the vehicle are located on the same line. As shown in Figure 12, there is no guarantee that the pixel distance between adjacent feature points in the image coordinate system can be correctly estimated.

First, 3 feature points for each frame are selected (126 feature points in all). Second, a linear regression of 126 feature points is conducted, as shown in Figure 13. The ordinate value of each feature point is revised to obtain the coordinate value, and it is found that the modified ordinate values change more uniformly and regularly as shown in Table 2.

(2) *Driving Speed Solution.* According to the corrected coordinates and equation (2), the same method solves dS frame displacement between two adjacent frames. The derivation of dS results in the average speed v between frames, as shown in Figure 14. Comparisons of calculated v_0 and recorded v are shown in Figure 15.

As Figure 14 shows, standard deviation of v ($\mu \pm \sigma$) ranges between 33.6 km/h and 38.7 km/h with the mean of 36.1 km/h. Recorded v_0 ranges from 36.5 km/h to 38.7 km/h, resembling the ranges of 36.7 km/h~38.6 km/h as in the estimated range. The average speed is 37.5 km/h and the deviation from the calculated average v_0 is 3.7%, meeting the error standards.

(3) *Acceleration Solution.* The fourth-order fitting equation of v is found as

$$v = 7t^4 - 25t^3 + 28t^2 - 12t + 38. \quad (7)$$



FIGURE 12: Difficult to ensure the selected feature points are on the same line.

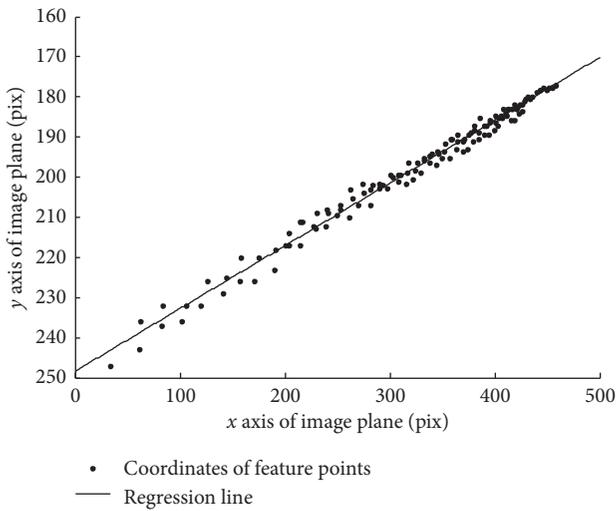


FIGURE 13: Coordinates and linear regression of 126 feature points.

TABLE 2: Partial ordinates of feature points corrected.

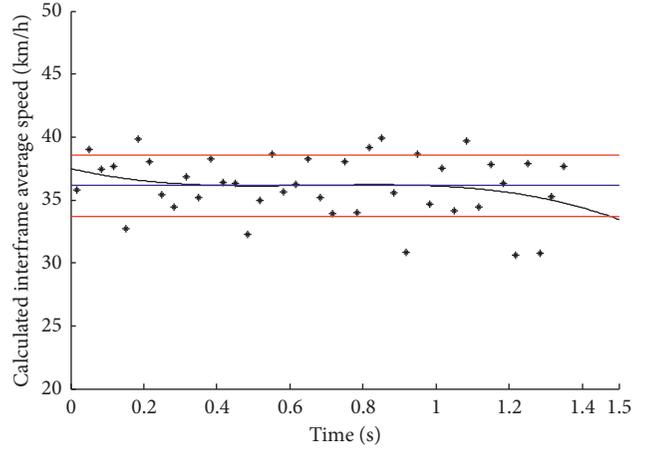
	<i>x</i>	<i>y</i>	<i>y_modify</i>	<i>x</i>	<i>y</i>	<i>y_modify</i>	<i>x</i>	<i>y</i>	<i>y_modify</i>
<i>a</i>	33	249	244	102	236	232	157	226	224
<i>b</i>	62	236	238	126	226	228	175	220	221
<i>c</i>	204	214	216	241	209	210	275	204	205
<i>a</i>	61	243	238	120	232	229	171	226	221
<i>b</i>	84	232	235	144	225	226	191	218	218
<i>c</i>	216	212	213	253	207	209	284	202	204
<i>a</i>	82	237	235	141	226	225	190	223	219
<i>b</i>	106	232	231	158	220	223	204	217	216
<i>c</i>	230	209	212	262	203	207	293	204	204

Then, the derivation of the equation results in the following for *a*.

$$a = \frac{dv}{dt} = 29t^3 - 74t^2 + 56t - 12. \quad (8)$$

Linear regressions of curve *a* and scatter *a*₀ are shown in Figure 16.

The difference between the mean values of two accelerations is 5.6%. Although the two fitting line slopes are opposite, the lines are below 0, indicating the acceleration is always negative.



◆ Calculated interframe average speed
 — Third order polynomial fitting curve
 — Mean value
 — Bounds of standard deviation

FIGURE 14: The curve of *v* and *t*.

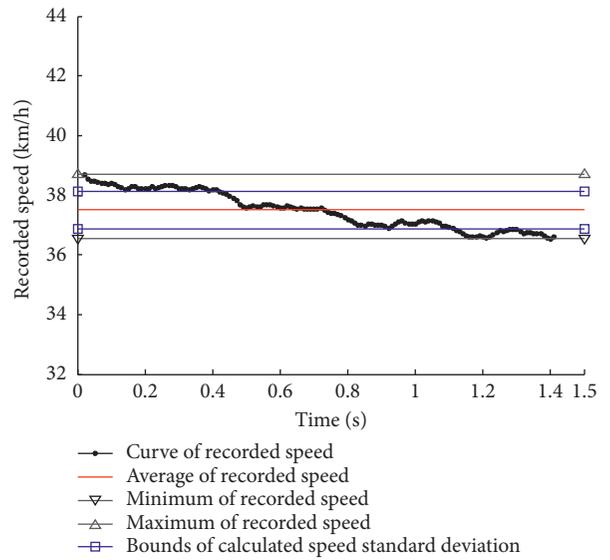


FIGURE 15: Compared analysis between *v*₀ and *v*.

(4) *Travel Distance Calculation.* The integral of frame displacement *dS* about *T* results in the travel distance *S*. Figure 17 shows the estimated driving distance *S* by this method compared with the test record of *S*₀. Two curves approximately coincide, in 1.4 s with *S* = 13.7 m and *S*₀ = 14.4 m, resulting in the difference of 4.9%, which satisfies the error requirements.

5.2. Middle- and High-Speed Running

5.2.1. In the Camera Perspective of 90 Degrees

(1) *Driving Speed Solution.* When 10 consecutive frames are extracted from the video and the same methods are applied

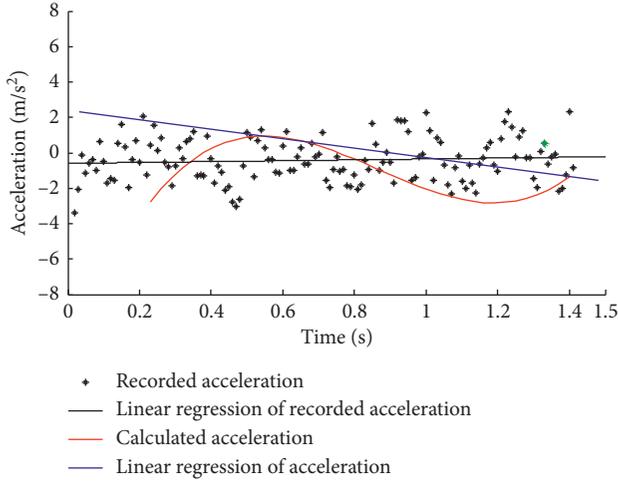


FIGURE 16: Linear regression of a and a_0 .

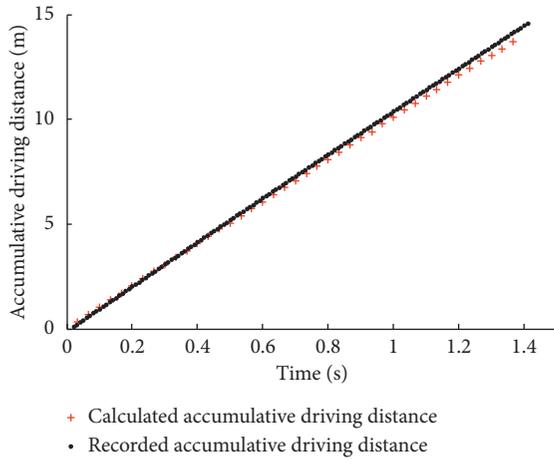


FIGURE 17: Compared analysis of S and S_0 .

to calculate the mean speed v , which is compared with record speed v_0 , the result is shown in Figure 18.

The average speed of v and v_0 are 76.99 km/h and 76.77 km, respectively, which is different from each other by approximately 0.3%.

(2) *Travel Distance Calculation.* The linear regression of travel distance S gives

$$S = 21.37t + 0.004. \quad (9)$$

The coefficient of determination $R^2=1$ indicates the regression is significant, and the vehicle is approximately under a uniform motion, while the speed is 21.37 m/s (76.93 km/h). The difference between the estimated $S = 6.42$ m and the observed $S_0 = 6.40$ m is 0.3%, meeting the error requirements. Figure 19 shows the comparison between the estimated and observed accumulative driving distances.

The comparisons of its acceleration values are not discussed here, as the experiment is very close to uniform motion and similar to the situation described in the “Low-

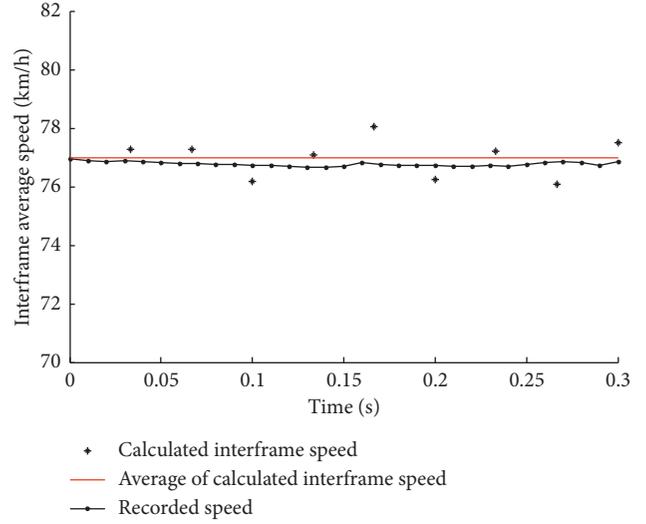


FIGURE 18: Compared analysis between v_0 and v .

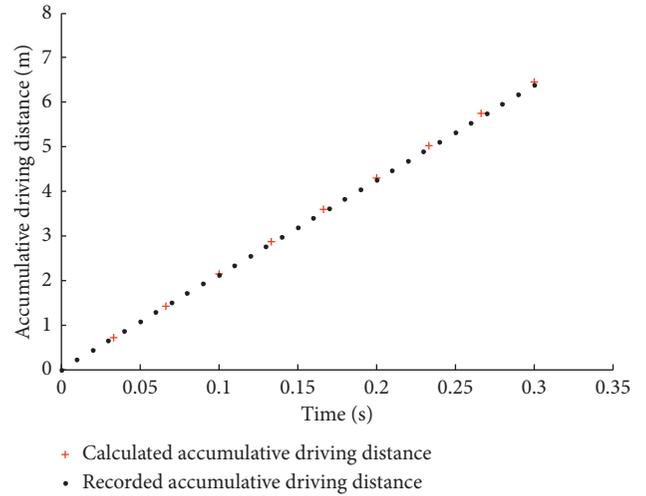


FIGURE 19: Compared analysis of S and S_0 .

Speed Running” section. The next section “*Deceleration Running*” will discuss especially for the deceleration values.

5.2.2. In the Camera Perspective of 30 Degrees

(1) *Travel Distance Calculation.* 25 continuous frame images of the test vehicle are extracted first, and 75 feature points coordinates are fixed, as shown in Figure 20.

Based on the modified feature point coordinate, we obtain v that is shown in Figure 21 with v_0 .

The average speed of v is 75.31 km/h and the standard deviation ($\mu \pm \sigma$) ranges from 70.58 km/h to 80.07 km/h. The average of v_0 is 76.75 km/h, which is different from the calculated average v_0 for approximately 1.9%, meeting the requirements.

(2) *Travel Distance Calculation.* Cumulative driving distance S and test record S_0 are shown in Figure 22.

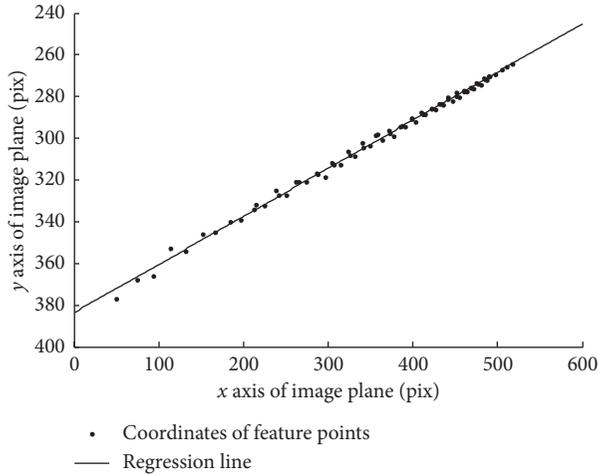


FIGURE 20: Coordinates and linear regression of 75 feature points.

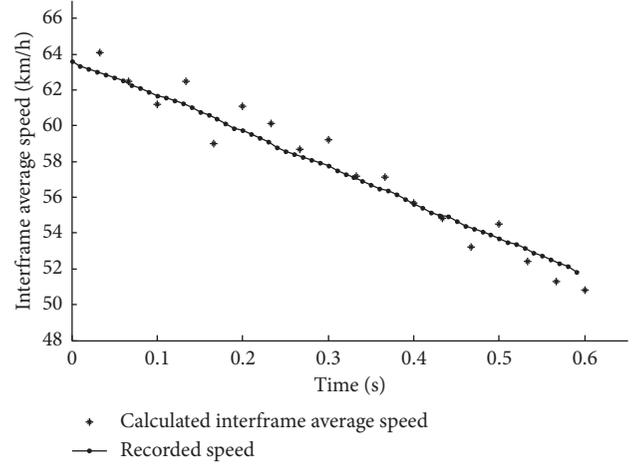


FIGURE 23: Compared analysis between v and v_0 .

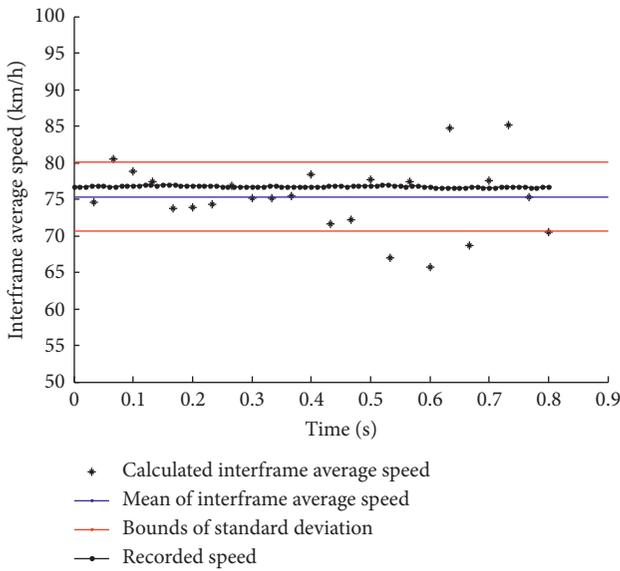


FIGURE 21: Compared analysis between v and v_0 .

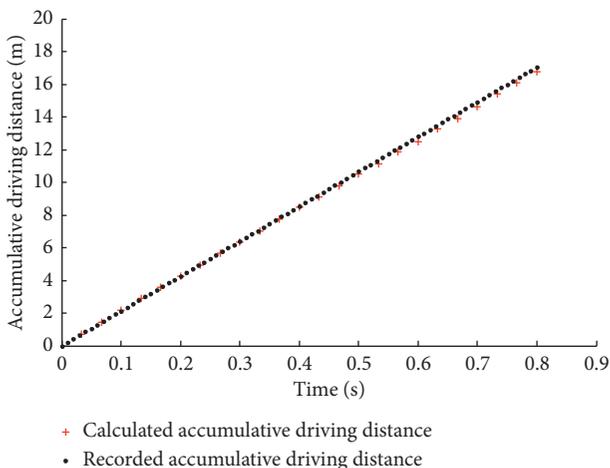


FIGURE 22: Compared analysis of S and S_0 .

After the regression, with coefficient of determination $R^2 = 0.98$, the vehicle is approximately under a uniform motion. In 0.8 s, the difference between calculated $S = 16.73$ m and record $S_0 = 17.05$ m is 1.9%.

5.3. Deceleration Running

5.3.1. In the Camera Perspective of 90 Degrees

(1) *Driving Speed Solution.* 10 frames of continuous images are extracted through the video region and the same methods are applied to calculate the mean speed v , which is compared with record speed v_0 shown in Figure 23.

We took F Inspection Method for a validation between v and v_0 for a significance test, then the resulting H value is 0 and P value is 0.38, which is greater than the significant level of 0.05. It indicates that the equation is homogeneous. It is found that there is no significant difference, and the mean speed difference is 1.3%.

(2) *Travel Distance Calculation.* In the above calculation of acceleration, it shows that the acceleration scatter value obtained directly through ds/dt is dispersed and has a larger error. Therefore, for the solution of acceleration in the deceleration state, we can choose to solve the cumulative driving distance S by curve fitting under different conditions and double derivations to obtain the acceleration curve.

The calculated result of S is shown in Figure 24 with an error of about 1.1% and nonlinear change.

(3) *Acceleration Solution.* Because the speed reduction of the test vehicle is basically unchanged through the video region, secondary fitting S

$$S = -3.1t^2 + 18t, \quad (10)$$

residuals to 0.043, then secondary fitting S_0

$$S_0 = -2.9x^2 + 18x, \quad (11)$$

residuals to 0.095, and then twice derivative of both

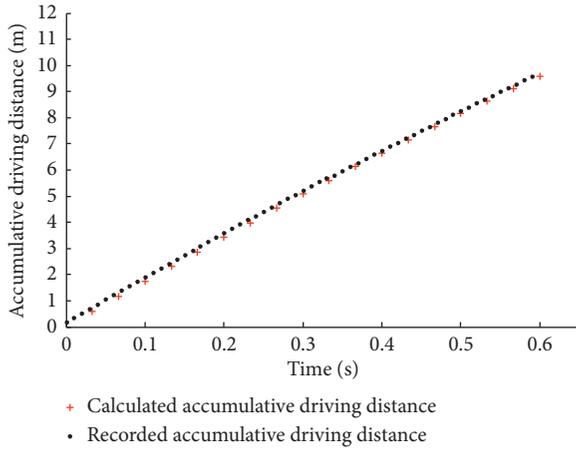


FIGURE 24: Compared analysis of S and S_0 .

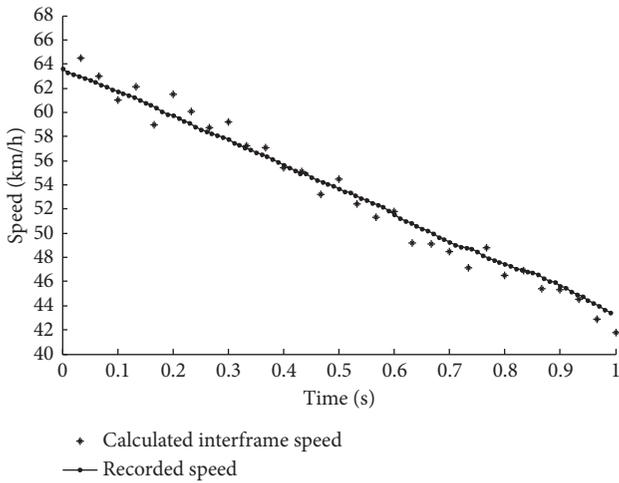


FIGURE 25: Compared analysis between v_0 and v .

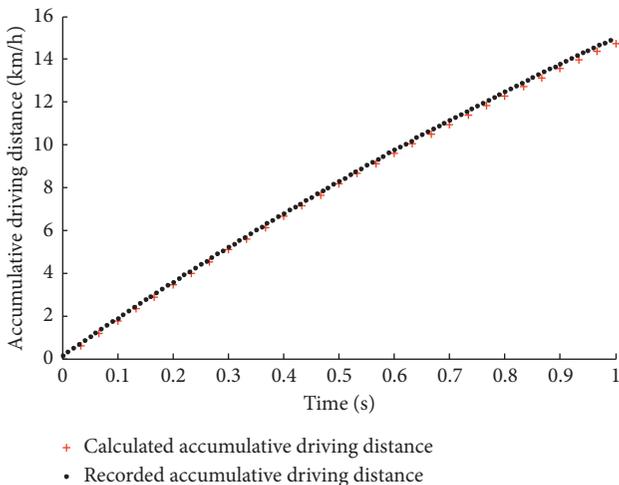


FIGURE 26: Compared analysis of S and S_0 .

$$\begin{cases} a = -6.2, \\ a_0 = -5.8. \end{cases} \quad (12)$$

The resulting acceleration difference is 6.9%.

5.3.2. *In the Camera Perspective of 30 Degrees.* By the same method, driving speed and distance can be obtained as follows in Figures 25 and 26. There is no significant difference in driving speeds, the mean error is 1.2%, the distance error is also 1.1%, and the acceleration error is 9%.

6. Conclusion

In this paper, a method based on a direct linear transformation is proposed to calculate the moving state of a vehicle in a video image, which can fully reflect the running state of the vehicle based on the collected data and is found to be stable and easy to obtain. The application condition is thoroughly described throughout the paper. Examples are given to show the estimation processes of the speed of the vehicle, acceleration, and driving distance under 2 angles for 3 different motion states.

In order to verify the validity of the method, vehicle motion states are estimated in experimental test cases; the results show that in the 90° angle of camera perspective, low-speed, under-medium-speed, and high-speed and deceleration running three states, the calculation of the speed and record value error is found to be less than 1.5%, the acceleration value error is less than 7%, and the driving distance value error is less than 2%. In 30° angle of camera perspective of view, low-speed, under-medium-speed, and high-speed and deceleration running three states, the calculation of the speed and record value error is found as less than 4%, the driving distance value error is less than 5%, and the acceleration value error is less than 10%. The performance of estimations made by the proposed method in terms of errors for vehicle speed and distance value is found to be higher than other existing methods, and the error of acceleration value complies with the error requirement, which shows that the method presented in this paper is effective in solving the driving state of the vehicles in the video.

This paper does not currently attempt to estimate the vehicle movement states of running trajectory under the view of overlooking camera. We will continue our research by adding vehicles' cross-lane curve states in the near future, and the selection criteria of the feature points of the body profile under overlooking camera will be newly defined.

Data Availability

All of the data related to this paper are available for peer researchers to validate.

Disclosure

The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This research was funded by the National Key Research Development Program of China: Research on Digital Simulation and Reappearance Technology of Road Traffic Accidents, Grant no. 2016YFC0800702-1, Shanghai Committee of Science and Technology Program for Science and Technology Development: Research on the Key Technology of Monitoring Video Forensic Identification, Grant no. 17DZ1205500, National Natural Science Foundation of China General Program: Study on traffic injury mechanism based on digital scene reconstruction, Grant no. 81571851, and ADEK Award for Research Excellence (AARE-2017), Grant no. 8-434000104, Central-Level Research Institutes Public Welfare Project, Grant no. GY2020G-7, and Central-level Research Institutes Public Welfare Project, Grant no. GY2018Z-3.

References

- [1] Y.-J. Byon, B. Abdulhai, and A. Shalaby, "Real-time transportation mode detection via tracking global positioning system mobile devices," *Journal of Intelligent Transportation Systems*, vol. 13, no. 4, pp. 161–170, 2009.
- [2] Y. J. Byon and S. Liang, "Real-time transportation mode detection using smartphones and artificial neural networks: performance comparisons between smartphones and conventional global positioning system sensors," *Journal of Intelligent Transportation Systems*, vol. 18, no. 1–4, pp. 264–272, 2014.
- [3] Y. J. Byon, J. Ha, C.-S. Cho, T.-Y. Kim, and C. Yeun, "Real-time transportation mode identification using artificial neural networks enhanced with mode availability layers: a case study in dubai," *Applied Sciences*, vol. 7, no. 9, 2017.
- [4] S. Tarkowski, A. Nieoczym, J. Caban, L. Gardyński, and J. Vrabel, "Reconstruction of road accident using video recording," *MATEC Web of Conferences*, vol. 252, Article ID 05023, 2019.
- [5] G. Wu, F. Chen, X. Pan, M. Xu, and X. Zhu, "Using the visual intervention influence of pavement markings for rutting mitigation-part I: preliminary experiments and field tests," *International Journal of Pavement Engineering*, vol. 20, no. 6, pp. 734–746, 2019.
- [6] J. I. Castillo-Manzano, M. Castro-Nuño, L. López-Valpuesta, and L. V. Vassallo, "The complex relationship between increases to speed limits and traffic fatalities: evidence from a meta-analysis," *Safety Science*, vol. 11, pp. 287–297, 2019.
- [7] F. Chen, M. T. Song, and X. X. Ma, "Investigation on the injury severity of drivers in rear-end collisions between cars using a random parameters bivariate ordered probit model," *International Journal of Environmental Research and Public Health*, vol. 16, no. 14, p. 2632, 2019.
- [8] F. Chen and S. Chen, "Injury severities of truck drivers in single- and multi-vehicle accidents on rural highways," *Accident Analysis & Prevention*, vol. 43, no. 5, pp. 1677–1688, 2011.
- [9] Z. Feng, M. Yang, W. Zhang, Y. Du, and H. Bai, "Effect of longitudinal slope of urban underpass tunnels on drivers' heart rate and speed: a study based on a real vehicle experiment," *Tunnelling and Underground Space Technology*, vol. 81, pp. 525–533, 2018.
- [10] M. Georges, "Computation of optical flow using dynamic programming," in *Proceedings of the IAPR Workshop on Machine Vision Applications*, pp. 249–252, Tokyo, Japan, November 1996.
- [11] R. Mech and M. Wollborn, "A noise robust method for segmentation of moving objects in video sequences," *IEEE in Proceedings of the IEEE international Conference on acoustics, speech, and signal processing*, vol. 4, pp. 2657–2660, Munich, Germany, April 1997.
- [12] D. Toth, T. Aach, and V. Metzler, "Illumination-invariant change detection. Image analysis and interpretation," in *Proceedings of the IEEE Southwest Symposium*, pp. 3–7, IEEE, Austin, TX, USA, April 2000.
- [13] J. Arunehru and M. K. Geetha, "A quantitative real-time analysis of object tracking algorithm for surveillance applications," *International Journal of Emerging Technology and Advanced Engineering*, vol. 3, no. 1, pp. 234–240, 2014.
- [14] F. Gao, S.-Y. Pan, and J.-G. Chen, "Speed determination using video recording," *Chinese Journal of Forensic Sciences*, vol. 5, pp. 46–48, 2009.
- [15] Beijing Traffic Management Bureau, *GA/T1133-2014 Vehicle Speed Determination Based on Video Images*, Standards Press of China, Beijing, China, 2014.
- [16] Institute of Forensic Science, Ministry of Justice, *2012 Assessment of Forensic Expertise Verification Documents*, Science Press, Beijing, China, 2013.
- [17] Institute of Forensic Science, Ministry of Justice, *2013 Assessment of Forensic Expertise Verification Documents*, Science Press, Beijing, China, 2014.
- [18] Institute of Forensic Science, Ministry of Justice, *2014 Assessment of Forensic Expertise Verification Documents*, Science Press, Beijing, China, 2015.
- [19] Institute of Forensic Science, Ministry of Justice, *2015 Assessment of Forensic Expertise Verification Documents*, Science Press, Beijing, China, 2016.
- [20] A. Kumar, P. Khorramshahi, and W. A. Lin, "A semi-automatic 2D solution for vehicle speed estimation from monocular videos," in *Proceedings of the 2018 IEEE/CVF Conference on computer vision and pattern recognition workshops (CVPRW)*, June 2018.
- [21] B. Coifman, D. Beymer, P. McLauchlan, and J. Malik, "A real-time computer vision system for vehicle tracking and traffic surveillance," *Transportation Research Part C: Emerging Technologies*, vol. 6, no. 4, pp. 271–288, 1998.
- [22] A. Yoneyama, C. H. Yeh, and C. C. J. Kuo, "Robust traffic event extraction from surveillance video," in *Proceedings of the Visual Communications & Image Processing*, January 2004.
- [23] Z. Z. Wang, *Photogrammetry Theory*, Wuhan University Press, Wuhan, China, 1979.
- [24] W. J. Feng, *Close-range Photogrammetry*, Wuhan University Press, Wuhan, China, 2002.
- [25] B. Yang, X. Jin, X. Zhang, X. Chai, J. Xu, and X. Hou, "Information acquisition and process reconstruction of traffic accidents based on digital photogrammetry," *Automotive Engineering*, vol. 32, no. 6, pp. 530–534, 2010.
- [26] X. Q. Miao, X. L. Jin, and X. Y. Han, "Based on the surveillance video of the accident scene reconstruction method digitizing," *China Journal of Highway and Transport*, vol. 31, no. 4, pp. 241–249, 2011.
- [27] X. Y. Han, X. L. Jin, and X. Y. Zhang, "Vehicle movement information reconstruction based on video images and DLT

- theory,” *Automotive Engineering*, vol. 34, no. 12, pp. 1145–1149, 2012.
- [28] T. Gong, “Research on layout project of control points in close-range photogrammetry,” *Journal of Southwest Jiao Tong University*, vol. 32, pp. 330–335, 1997.
- [29] S. Bullinger, C. Bodensteiner, and M. Arens, “3D object trajectory reconstruction using stereo matching and instance flow based multiple object tracking,” in *Proceedings of the 16th International Conference on Machine Vision Applications (MVA)*, May 2019.
- [30] K. H. Hasith, H. D. Wang, and H. D. Zhang, “Multiple object tracking with attention to appearance, structure, motion and size,” *IEEE Access*, vol. 99, pp. 104423–104434, 2019.