Archives of
Transport System
Telematics

Volume 7

Issue 2

May 2014

Detailed evaluation and analysis of vision-based online traffic parameters estimation approach using low-resolution web cameras

M. ALI[a], L. C. TOUKO TCHEUMADJEU[b], S. ZUEV[a], S. RUPPE[b]
[a] GERMAN AEROSPACE CENTER (DLR), Institute of Optical Sensor Systems, Information Processing for Optical Systems, Rutherfordstraße 2, 12489 Berlin, Germany
[b] GERMAN AEROSPACE CENTER (DLR), Institute of Transportation Systems, Rutherfordstraße 2, 12489 Berlin, Germany
EMAIL: musharaf.ali@dlr.de

ABSTRACT

In this paper, we give an overview and a detail analysis of our approach for vision-based real-time traffic parameters estimation using low-resolution web cameras. Traffic parameters estimation approach mainly includes three major steps, (1) stable background estimation, (2) vehicle detection, mean speed and traffic flow estimation, and (3) traffic scene classification into three states (normal and congested). The background image is estimated and updated in real-time by novel background estimation algorithm based on the median of First-in-First-Out (FIFO) buffer of rectified traffic images. Vehicles are detected by background subtraction followed by post-processing steps. By exploiting the domain knowledge of real-world traffic flow patterns, mean speed and traffic flow can be estimated reliably and accurately. Naive Bayes classifier with statistical features is used for traffic scene classification. The traffic parameter estimation approach is tested and evaluated at the German Aerospace Center's (DLR) urban road research laboratory in Berlin for 24 hours of live streaming data from web-cameras with frames per second 1, 5 and 10. Image resolution is 348 x 259 and JPEG compression is 50%. Processed traffic data is cross-checked with synchronized induction loop data. Detailed evaluation and analysis shows high accuracy and robustness of traffic parameters estimation approach using low-resolution web-cameras under challenging traffic conditions.

KEYWORDS: background estimation, vehicle detection, time mean speed, traffic flow, traffic state classification

1. Introduction

The goal of reduced traffic congestion and increased traffic mobility can be achieved by automatic real-time traffic monitoring using sensor technology. The quality of traffic-monitoring measurements depends on the type of traffic sensors used. Most commonly used sensors are the point sensors like video cameras and loop detectors. Typical point sensors are good at measuring time domain characteristics such as traffic flow and mean speed in time. Traffic cameras and induction loops are the two main point sensors. Traffic cameras have a great potential as they provide good spatial and temporal resolution. Moreover, they are cost effective and easy to maintain [1].

The objective of the current research is to investigate the feasibility of automatic real-time traffic parameters estimation and traffic states estimation using low-cost and low-resolution un-calibrated web-cameras.

Environmental factors and occlusions play a significant role in the quality of traffic parameters estimation. Moreover the camera technology used in traffic- monitoring has a significant impact on the performance of traffic monitoring systems. The main challenges for video-based real-time traffic-monitoring system from un-calibrated traffic cameras are as follows. No information on camera interior and exterior orientation parameters readily available. Web cameras used in the research have low frame and spatial resolution. This implies that lots of valuable information in the scene regarding vehicle detection and vehicle movement has been lost. Camera jittering due to the windy conditions can create false foreground detection. Ambient illumination changes such as passing clouds, other moving objects like walking people, moving bicycles, swinging trees can create false moving foregrounds. In dark weather conditions, dark colored cars have low contrast against the background roads, so they have a high probability of identification as background. The shadow of vehicles creates a false foreground on sunny days. Under heavy traffic flow conditions, blobs of different vehicles may merge into one.

To simplify the problem, we made some assumptions which are as follows. The traffic moves largely toward or away from the camera and the road section should be straight up to 40 meters minimum [2]. The video cameras used here are un-calibrated, in the sense that camera constant, principal point location, and image affinity parameters are irrelevant. Regarding exterior orientation: the ground in front of the camera is planar and the cameras are fixed to a static structure like a pole or a building. The speed of the vehicles should be finite and within legal limits. There should be no sudden lane changes by the vehicles in the time between images in the image sequence. Motion is constraint to the road plane.

With these assumptions, the vehicles are treated as though they travel in one dimension along a straight line, either toward or away from the camera in the image sequence.

The rest of the paper is organized as follows: In Section II, we describe the basic framework for initialization step. In Section III, we explain the proposed traffic scene classification. The background estimation approach is elaborated in Section IV. Vehicle detection step is explained in Section V. Velocity estimation and vehicle counting procedures are described in Section VI and Section VII, respectively. We demonstrate the offline and online experimental results in Section VIII. Finally, conclusions are made in Section IX.

2. Initialization step

The first step is the manual selection of the detection zone from the traffic scene. The length of the detection zone must be at least 40 meters long along the direction of traffic flow. The main objective of this step is to remove the unwanted information from the scene so that the only moving objects in the scene are the vehicles as shown in Fig. 1. Selection of the detection zone improves the quality of the vehicle detection algorithm and decreases the computational cost.

After the selection of the detection zone, the second step is the perspective correction in which the original image of the detection zone is transformed from projective view to an orthographic-view image as shown in Fig. 1. Through perspective correction, all the pixels in the scene become on the same scale for further processing and measurements. Perspective correction is performed according to the method described in [3].

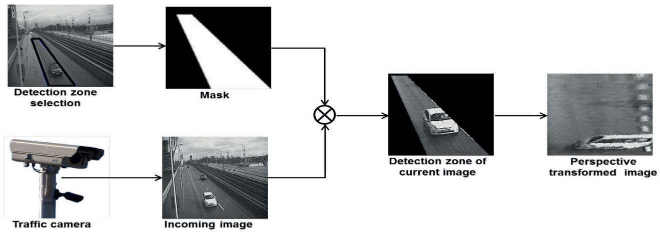


Fig. 1. Detection zone selection, region of interest extraction and perspective transformation of incoming image [own study]

3. Traffic scene classification

Traffic scene classification module is the integral part of our background estimation approach. It makes background estimation approach traffic invariant. Traffic scene is classified into two classes, (i) Free, (ii) Congested. We use a trained two-class naive Bayesian classifier to classify traffic scene. The first-class w_1 corresponds to traffic congestion state and the second-class w_2 corresponds to traffic free state. Feature vector **x** is a N-D vector $(e, s, e_d, m^{3rd})^T$, where features e, s, ed ,m3rd denote the entropy, standard deviation, edge density and 3rd moment of the current image , respectively. i.e,

$$P(w_i \mid x) = \frac{P(w_i)P(x \mid w_i)}{P(x)} \tag{1}$$

The class conditional probability density $P(x \mid w_i)$ is modeled by a Gaussian function. The parameters of the Gaussian function for each class is obtained by maximum-likelihood estimation [4] . We assume equal prior probabilities $P(w_i)$ for congestion and free state classes. The evidence $P(x)$ is a constant and scales both posteriors equally. It therefore does not affect classification and can be ignored.

4. Background estimation

Stable background estimation is the key step for vehicle detection. Most of the state-of-the-art background estimation algorithms [5, 6, 7] did not discuss the problem of stable background image estimation under changing traffic conditions e.g. congestion. In long traffic congestion scenarios, stable background image estimation is a quite challenging.

We proposed a novel background estimation algorithm based on the median of FIFO buffer of incoming rectified orthographic images. Incoming rectified orthographic images have to pass through the congestion detection test as described in Section III. If congestion is detected in the current image, then it is not added into FIFO buffer. If the traffic state in current image is free-flow then current image is added into the FIFO buffer and the FIFO buffer is updated according to the First-in-First-out method. Median of current FIFO buffer generates a stable background image which is invariant of changing weather, illumination and traffic conditions. The buffer size is 21, which is determined empirically according to the real-time requirements of the system. The background estimation approach is described in Fig. 2.

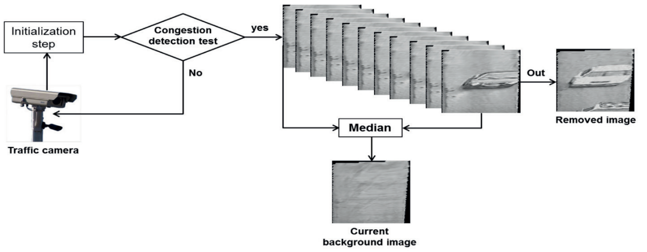


Fig. 2. FIFO-buffer based background estimation approach [own study]

5. Vehicle detection

The overall quality of the real-time traffic-monitoring system depends upon the robust vehicle detection. In the static camera scenario, moving vehicles can be detected through a simple background subtraction operation. Assume that *I(x, y)* is the current image and *CB(x, y)* is the current updated background image. The difference image *D(x, y)* is used to detect moving vehicles defined as follows.

$$D(x,y) = \begin{cases} 0 \;\; if \; | I(x,y) - CB(x,y) | < T \\ 1 \;\; otherwise \end{cases} \tag{2}$$

Where T is the predefined threshold. T is calculated as the average of the difference image $D(x,y)$ [8]. To eliminate the small nosiy blobs and to fill small holes in the blobs, detected foreground objects are post processed with morphological operations. Fig. 3 shows the example of the vehicle detection step.

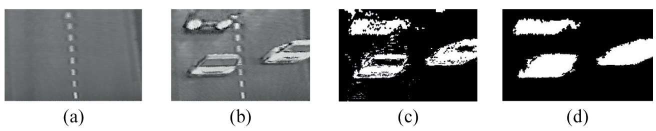


Fig. 3. Example of vehicle detection. (a) Current background image. (b) Current image of the traffic scene. (c) Difference image. (d) Difference image after morphological steps [own study]

6. Velocity estimation

For incoming traffic scene, the detected vehicle closest to the camera field of view in the current image is only investigated for finding its correspondence with the detected vehicles in the previous image. By using the centroids information of the detected vehicles, the velocity of the vehicles can be measured by Euclidean distance reliably.

For simplicity, we model vehicle objects by their centroid information, let $V_t^i = (x_t^i, y_t^i)$ vehicle *i* at time *t*, where x_t^i is the horizontal centroid position, and y_t^i is the vertical centroid position of the vehicle *i* in the image at time *t*.

For incoming traffic flow scenario, correspondence between the vehicle V_t^1 in the current *image (t)* and the vehicles V_{t-1}^j in the previous *image(t-1)* can be found by two constraints.

Position constraint: For incoming traffic scenario under normal traffic flow conditions, the horizontal centroid position x_t^1 of the vehicle V_t^1 in the current image(t) must be greater than the horizontal centroid position x_{t-1}^j of the corresponding vehicle V_{t-1}^j in the previous image (t-1).

$$x_t^1 \;>\; x_{t-1}^j \rightarrow \; true \tag{3}$$

This condition is assumed on the basis of the principle that a vehicle cannot travel in reverse direction under normal incoming traffic flow situations.

Minimum Euclidean distance: if the position constraint holds true, then two vehicles in the adjacent images with minimum Euclidean distance have correspondence.

$$\text{Distance} = \arg\min \left(\sqrt{\left(x_t^1 - x_{t-1}^j \right)^2 + \left(y_t^1 - y_{t-1}^j \right)^2} \right) \tag{4}$$

we can calculate the scale factor in the image by using geometric relationships or markers that are inherently available in the traffic scene such as lane markers in the middle of the road section. By calculating the length of the lane markers in pixels in image coordinates and by knowing their length in meters in real-world coordinates, any point on the image can be back-projected onto the road. Therefore, the calculated scale factor (meters/pixel) allowed us to calculate the distance between any two points on the road [9].

7. Vehicle count

For incoming traffic scene, the detected vehicle closest to the camera in the current *image(t)* and the detected vehicle closest to the camera in the previous *image(t-1)* are only under investigation for vehicle counting. The vehicle-counting algorithm comprises of two conditions as follows.

Condition 1: If the vehicle detected in current *image(t)* and no vehicle was detected in the previous *image(t-1)*. Then increment the counter as a new vehicle appeared in the current image such that.

$$\begin{aligned} &if \; v_t^i \;\Rightarrow\; v_{t-1}^j \;\rightarrow\; false \\ &\quad count = count + 1 \end{aligned} \tag{5}$$

Condition 2: If the vehicle V_t^1 detected in current *image(t)* closest to the camera field of view with the centroid position $\left(x_t^1, y_t^1\right)$ and vehicle V_{t-1}^1 detected in previous *image(t-1)* closest to the camera field of view with centroid position $\left(x_{t-1}^1, y_{t-1}^1\right)$, if x_{t-1}^1 coordinate of vehicle V_{t-1}^1 is greater than x_t^1 coordinate of vehicle V_t^1 then increment the counter.The condition 2 is given

$$\begin{aligned} &if \; \left(x_{t-1}^1 + a \;\geq\; x_t^1\right) \rightarrow true \\ &\quad count = count + 1 \end{aligned} \tag{6}$$

where α is the adjustment factor and it has a small integer value from 0 to 5 depends upon the frame rate of the video sequence. Detailed describtion of velocity estimation and vehicle count algorithm is described in [10].

8. Results and discussion

We carried out a comprehensive test campaign to validate our real-time traffic parameters estimation approach. We tested the proposed system on traffic surveillance scenes obtained from the German Aerospace Center's (DLR) urban road research laboratory in Berlin for 24 hours of live streaming data and offline data from traffic-cameras. Both day time and night time surveillance scenes under different traffic conditions are tested. The frame rates for online and offline image sequences are 1, 5 and 10 FPS. Image resolution is 348 x 259 pixels and JPEG compression is 50%. Processed traffic data is cross-checked with synchronized induction loop data. Fig. 4 shows some snapshots of the traffic scenes. The region inside the polygon is the detection zone as shown in Fig. 4. For several lanes in one direction, each lane is selected as a separate detection zone and processed independently.

(a) (b) (c)

Fig. 4. Experimental traffic sequences [own study]

8.1 Offline test

Offline validation is carried out for a 20 minutes interval for both daytime and night videos from above mentioned traffic scenes. The statistics in Table. 1 are collected from the lane 1 in sequence (a) , lane 3 in sequence (b) and lane 1 in sequence (c). Ground truth is the manual count of vehicles from 20 minutes of offline sequence. As induction loops are fixed and can not accomodate change of lanes by the vehicles. So loop count is prone to errors with respect to ground truth. Absolute error of traffic flow and traffic speed between induction loop and traffic camera is measured by the aggregation of one minute of data.

Table. 1 Offline comparison of measured (loop) and processed (traffic camera) velocity and vehicle count [own study]

Seq	Day/Night	FPS	Ground truth count	Loop count	Traffic camera count	Absolute Error Diff (Loop,Traffcam) in [veh/min]	Error (GroundTruth, Traffic camera) in %	Absolute Error Velocity (loop-traffic camera) In [km/h]
a	Day	10	106	88	110	1,6	3.77	6
a	Day	5	114	105	117	1,15	2.6	6,96
a	Day	1	112	96	113	1,13	0.89	9,56
b	Night	10	27	19	29	0,12	7.40	3,13
b	Night	5	27	22	27	0,18	0.00	4,3
b	Night	1	28	22	29	0,3	3.57	10,66
c	Day	10	156	127	147	1,25	5.76	7,07
c	Day	5	188	158	192	2,45	2.12	6,57
c	Day	1	177	158	166	1,3	6.22	16,7

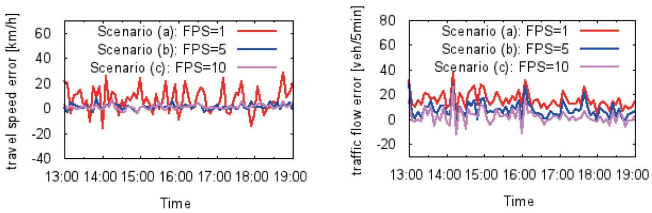
8.2 Online test

Second phase of test campaign is the evaluation of system's real-time capability to estimate traffic parameters. System is tested for 6 hours of live streaming data from 13:00 to 17:00 hours at 04.04.2014. Three different configurations with frame rates 1 FPS, 5 FPS and 10 FPS has been tested. Processed data from traffic camera and measured data from corresponding induction loop is aggregated for over the period of 5 minutes.

Let $(V_{loop}(t), V_{trafficcam}(t))$ be the measured and processed travel speed from the induction loop and traffic camera, respectively at time(t). Similarly $(Q_{loop}(t), Q_{trafficcam}(t))$ is the corresponding measured and processed traffic flow. The statistical evaluation requires the travel speed error $\Delta V(t)$ and the traffic flow error $\Delta Q(t)$ as described in equation (7).

$$\begin{aligned} \Delta V(t) &= V_{loop}(t) - V_{trafficcam}(t) \\ \Delta Q(t) &= Q_{loop}(t) - Q_{trafficcam}(t) \end{aligned} \qquad (7)$$

Fig. 5 shows the errors curve generated for travel speed and traffic flow from equation (7) for the three camera configurations (1 FPS, 5 FPS and 10 FPS). Error curves in Fig. 5 shows high accuracy in estimated travel speed and traffic flow for 5 FPS and 10 FPS with respect to induction loop data. While estimated travel speed and traffic flow from 1 FPS shows slightly high variance with respect to induction loop data.



(a) (b)

Fig. 5. Online (a) travel speed errors and (b) traffic flow errors [own study]

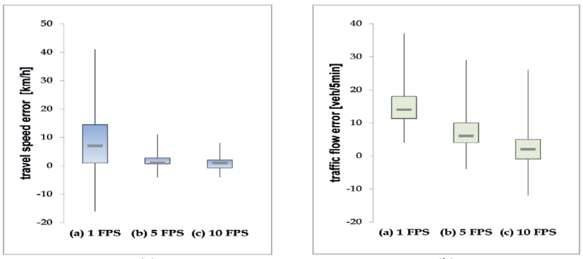
Table 2. lists other statistical measures calculated from $\Delta V(t)$ and $\Delta Q(t)$ in equation (7) to validate the accuracy and precision of estimated travel speed and traffic flow.

Table. 2. Online travel time and trafic flow errors [own study]

	Travel Speed Error [km/h]			Traffic Flow Error [veh/5min]		
	1 FPS (a)	5 FPS (b)	10 FPS (c)	1 FPS (a)	5 FPS (b)	10 FPS (c)
Minimum	-16	-4	-4	4	-4	-12
25th percentile	1	0,75	-0,75	11,25	4	-1
Median	7	1	1	14	6	2
75th percentile	14,50	2,75	2	18	10	5
Maximum	41	11	8	37	29	26
Mean	7,71	1,16	1,05	15,25	7,36	2,01
Standard deviation	10,52	2,53	2,15	6,30	6,19	6,08
Data Count	80	80	80	80	80	80

All statistical errors data for travel speed and traffic flow listed in Table. 2 for the three scenarios (a) 1FPS, (b) 5 FPS and (c) 10 FPS is plotted and shown in the Fig. 6. The box-and-whisker plot shows that the processed data from the camera achieved higher accuracy for frame rate 5 FPS and 10 FPS.as compared to 1FPS data sequence.

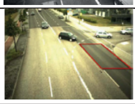
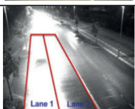
Our approach is simple and general enough to work for incoming and outgoing traffic scenarios under the assumptions described in section I. However, there are several factors that can affect the accuracy of the algorithms such as lane change by the vehicles in camera field of view, irregular road section, stop- and-go vehicle movement, large vehicle movement in the scene such as trucks and buses that cover most of the camera field of view and also covers the adjacent lanes thereby generating false foreground objects, and heavy traffic flow condition in which distance between vehicles is small (less than 4 pixels) so that multiple vehicles are detected as a single object. Our approach might not work for the night time scenes with incoming traffic scenario because vehicles are not identifiable due to the bright glare from head lights direct into the camera line of sight. Table. 3 shows some main reasons of error during automatic traffic parameters estimation from camera.



(a) (b)

Fig. 6. Boxplot errors of (a) travel speed and (b) traffic flow from the loop and traffic cam data of all scenarios. The end of the whiskers represent minimum and maximum errors [own study]

Table. 3. General traffic scene problems for traffic parameters estimation [own study]

Problem	Description	Example	Impact-Factor on estimated traffic parameters
1	Large trucks cover the whole detection zone		Medium
2	Bad camera view : large vehicles in adjacent lanes covers the detection zone		High
3	Small detection zone		High
4	Over exposure of vehicle head lights at night		Very High
5	Change of lane by vehicles		Medium

9. Conclusion

For an intelligent traffic surveillance system, it is important to estimate traffic parameters in real-time under different traffic conditions. In this paper we designed, implemented and evaluated the real-time traffic monitoring system using low resolution un-calibrated web-cameras. One of the main contributions of this work includes the novel FIFO buffer based background estimation algorithm which is adaptive to changing weather and traffic conditions. Statistical features based naive Bayesian classifier is used to classify traffic scene. Traffic scene classification module is an integral part of background estimation algorithm. It ensures generation of stable background image even under long congested traffic conditions. When the traffic state is free flow then traffic parameters are estimated. If traffic is congested or stop-and-go then system generates the flag of congested traffic condition. After detailed tests and evaluation, we concluded that camera configuration with frame resolution 5 FPS , image resolution 348 x 259 and JPEG compression 50% is optimal for real-time traffic monitoring. Higher frame resolutions and image resolutions are computationally expensive for real-time systems with no significant improvement in results.

Bibliography

[1] CHO Y., RICE J.: Estimating velocity fields on a freeway from low-resolution videos, IEEE Transactions on Intelligent Transportation Systems, vol. 7, no. 4 (2006)

[2] SCHOEPFLIN T. DAILY D.: A correlation technique for estimating traffic speed from cameras, TRB-03-3414 (2003)

[3] MALINOVSKIY Y., WU Y., WANG Y.: Video-Based Vehicle Detection and Tracking Using Spatiotemporal Maps, Transportation Research Record: Journal of the Transportation Research Board, vol. 2121 (2009)

[4] DUDA R.O., HART P.E., STORK D.G.: Patern Classification, 2nd ed. Hoboken, NJ: Wiley (2001)

[5] BRUTZER S., HOFERLIN B., HEIDEMANN G.: Evaluation of Background Subtraction Techniques for Video Surveillance, IEEE conference on Computer Vision and Pattern Recognition (2011)

[6] SEN-CHING C., CHANDRIKA K.: Robust techniques for background subtraction in urban traffic video, Visual communications and image processing, Proceedings of the SPIE, vol 5308 (2004)

[7] GOYETTE N., et. al.: Changedetection.net: A new change detection benchmark dataset, in Proc. IEEE Workshop on Change Detection (CDW'12) at CVPR'12, Providence, RI, 16-21 Jun 2012 (2012)

[8] XIOKUN L., PORIKLI F.M.: A hidden Markov model framework for traffic event detection using video features, Image Processing ICIP, Vol 5 (2004)

[9] DAILY D., CATHEY F.: The Automated use of un-calibrated CCTV cameras as quantitative speed sensors phase 3", Final report T269 (2006)

[10] ALI M., et. al.: Real-time traffic monitoring using low-resolution web-cameras, 19th ITS World Congress (2012)