

Reliable Infrastructural Urban Traffic Monitoring Via Lidar and Camera Fusion

Yi Tian, Hangxin Liu, and Tomonari Furukawa
 Virginia Tech

ABSTRACT

This paper presents a novel infrastructural traffic monitoring approach that estimates traffic information by combining two sensing techniques. The traffic information can be obtained from the presented approach includes passing vehicle counts, corresponding speed estimation and vehicle classification based on size. This approach uses measurement from an array of Lidars and video frames from a camera and derives traffic information using two techniques. The first technique detects passing vehicles by using Lidars to constantly measure the distance from laser transmitter to the target road surface. When a vehicle or other objects pass by, the measurement of the distance to road surface reduces in each targeting spot, and triggers detection event. The second technique utilizes video frames from camera and performs background subtraction algorithm in each selected Region of Interest (ROI), which also triggers detection when vehicle travels through each ROI. Based on detection events, vehicle location is estimated respectively. The final location estimation is derived by fusing the two estimation in the framework of Recursive Bayesian Estimation (RBE). Vehicle counts, speed estimation and classification are then performed using the vehicle location estimation in each time step. The approach achieves high reliability by combining the strength of both sensors. A sensor prototype has been built and multiple field experiments have been completed. High reliability is demonstrated in experiment by achieving more than 95% accuracy both in vehicle counting and classification.

CITATION: Tian, Y., Liu, H., and Furukawa, T., "Reliable Infrastructural Urban Traffic Monitoring Via Lidar and Camera Fusion," *SAE Int. J. Passeng. Cars – Electron. Electr. Syst.* 10(1):2017, doi:10.4271/2017-01-0083.

INTRODUCTION

Driven by the increasing needs for more efficient and safe mobility in last 20 years, the development of Intelligent Transportation System (ITS) has attracted increasing academic and industrial efforts as a potential solution to modern traffic challenges.[1][2] ITS brings multiple benefits including transportation system safety and mobility improvement, energy cost reduction by providing real-time optimized travel route, smart infrastructural controlling and better policy making. [3] In the structure of ITS, traffic monitoring system plays a fundamental role by constantly providing traffic data to traffic administration, individual drivers and traffic system and city planners. The performance of the traffic monitoring component in the system has direct influence on the functioning of ITS. As ITS is developed to be intensively data dependent, it requires modern traffic monitoring systems able to provide more reliable and informative traffic data to enhance the reliability and expand the functionality of ITS.

To date, a number of traffic monitoring solutions using different approaches already exists in both industry and academia. These solutions mostly fall into two categories, infrastructural sensors and mobile sensors. Mobile sensors, which is also known as probe-vehicle sensors, are sensors installed in individual vehicles that provide real-time traveling data. Most probe-vehicle sensors are based on techniques including Signpost Based Automatic Vehicle Location (AVL), Automatic Vehicle Identification (AVI), Ground-based Radio Navigation, Cellular Geolocation and Global

Positioning System (GPS). [4,5,6,7,8] On the other hand, infrastructural sensors, also referred as stationary sensors, include intrusive sensors and non-intrusive sensors. Intrusive sensors such as pneumatic tube detectors, inductive detectors loop (IDL), embedded magnetometers and weigh-in-motion detector, are sensors installed in or on the road surface.[9,10] And Non-intrusive sensors like Video Image Detection (VID) sensors, infrared sensors, Microwave-Doppler and radar and ultrasonic sensors normally are installed over the road or standing alongside.[11,12,13] Both categories of approaches have been further developed and extensively evaluated in regards to performance, cost and potentiality.

Mobile sensors normally have following advantages: Low function cost per data, continuous data provision, automated data collection and no disruption of traffic after installation. [4] However, several issues have hindered the real-world application of the approach. [14] pointed out that probe sensors are required to be securely installed on more than 5% of vehicles to provide comprehensive traffic information. And these sensors need to be connected to the same network. Since the application of probe sensor network also requires infrastructural supporting facilities, the initial cost in purchasing and installment is typically higher than stationary sensors. Concerns about motorist privacy leaking also has been raised and discussed. [15] Furthermore, probe vehicle approaches can only provide macroscopic traffic information. And for specific curves of a road, it may not be able to provide precise speed information. On the other hand,

stationary traffic monitoring systems have been the mainstream solution in practice. Intrusive sensors, compared with non-intrusive ones, are easier to achieve high reliability since they typically have stronger noise immunity. However, these sensors mostly require high maintenance and installation cost especially when the sensors are embedded in the road. Non-intrusive sensors normally have lower installation cost. But under different environmental condition, the performance of these sensors is more likely to be unstable. For instance, Video Image Detection (VID) sensors based on computer vision techniques face challenges from changing illumination factor. Even optimized VID approach has non-satisfying performance at night or bad weather days. [16]

This paper presents a novel infrastructural traffic monitoring approach which estimates traffic information by combining two sensing techniques. The approach uses input from a vision sensor such as camera, and an array of laser range finder sensors such as Lidar and output traffic condition information including passed vehicle counting, estimated vehicle speed and vehicle classification based on its size (length). The first technique uses local background subtraction algorithm to process video frames from camera and perform traffic monitoring. The second technique, at the same time, achieving the same goal by constantly measuring the distance from laser transmitter to the multiple target ground region. Both techniques monitor traffic independently and the system then fuses the outputs. Information fusion process based on grid based Recursive Bayesian Estimation is mathematically derived to demonstrate the higher reliability compared with the result from individual techniques. Following the formulation, field experiments are conducted and the results are analyzed.

The paper is organized as follows. [Section 2](#) in detail describes the principles of each individual techniques used for traffic monitoring. In [section 3](#), the presented traffic monitoring approach overview and estimation framework are addressed. The observation likelihood fusion process is presented. In [section 4](#), the system implementation is presented with the field experiment result to demonstrate the efficacy of the proposed approach, followed by the discussion on different aspects of the proposed approach and future research directions.

LIDAR MEASUREMENT AND VISION-BASED METHOD FOR TRAFFIC MONITORING

The technique presented in this paper is based on the fusion of Lidar measurement and vision-based information [17]. As shown in [Fig. 1](#), while Lidar measurement relies on a Laser transmitter and receiver and actively conducts the measurement, vision-based method uses a camera passively collecting information. The different physics of the measurement leads to different performance of the two methods in different environment and the combination improves the overall reliability by cross-validating the output from both methods.

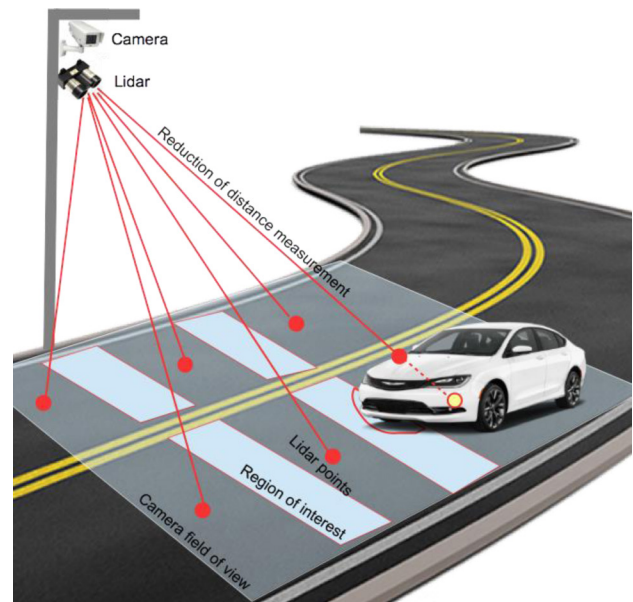


Figure 1. Lidar measurement and vision-based method for traffic monitoring.

Lidar Measurement for Traffic Monitoring

In traffic monitoring, each Lidar measures the distance between the transmitter and projecting point on the road or other obstacles and provides point information about the road plane. As shown in [Fig. 1](#), a reduction of the distance measurement indicates a possible vehicle or pedestrian occupying the detecting point on the road. Based on this, an array of Lidars that provides multi-point distance measurement can be used to sample a 2D region of interest on the road and acquire information of vehicle counts. By calibration of the sampling point, vehicle speed estimation and size classification can also be estimated.

Compared with vision based method, the strongest advantage of the Lidar measurement lies in its independence from sunlight or any other external light source. As a result, Lidar provides consistent measurement regardless of time of day, shadow or other illumination problems. The problem of Lidar measurement is the limited detecting zone. As a point-measurement, assumption is made that all objects pass through the detecting line between transmitter and projecting point on road, are a passing vehicle; and every passing vehicle passes through this line. Depending on the road condition and sensor installation, more Lidars are possibly required to make the assumption hold.

Vision-Based Method for Traffic Monitoring

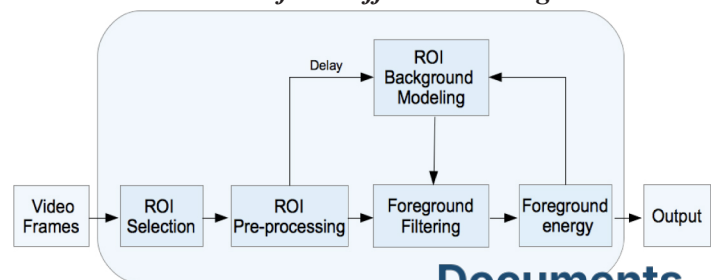


Figure 2. Background subtraction algorithm working flow.

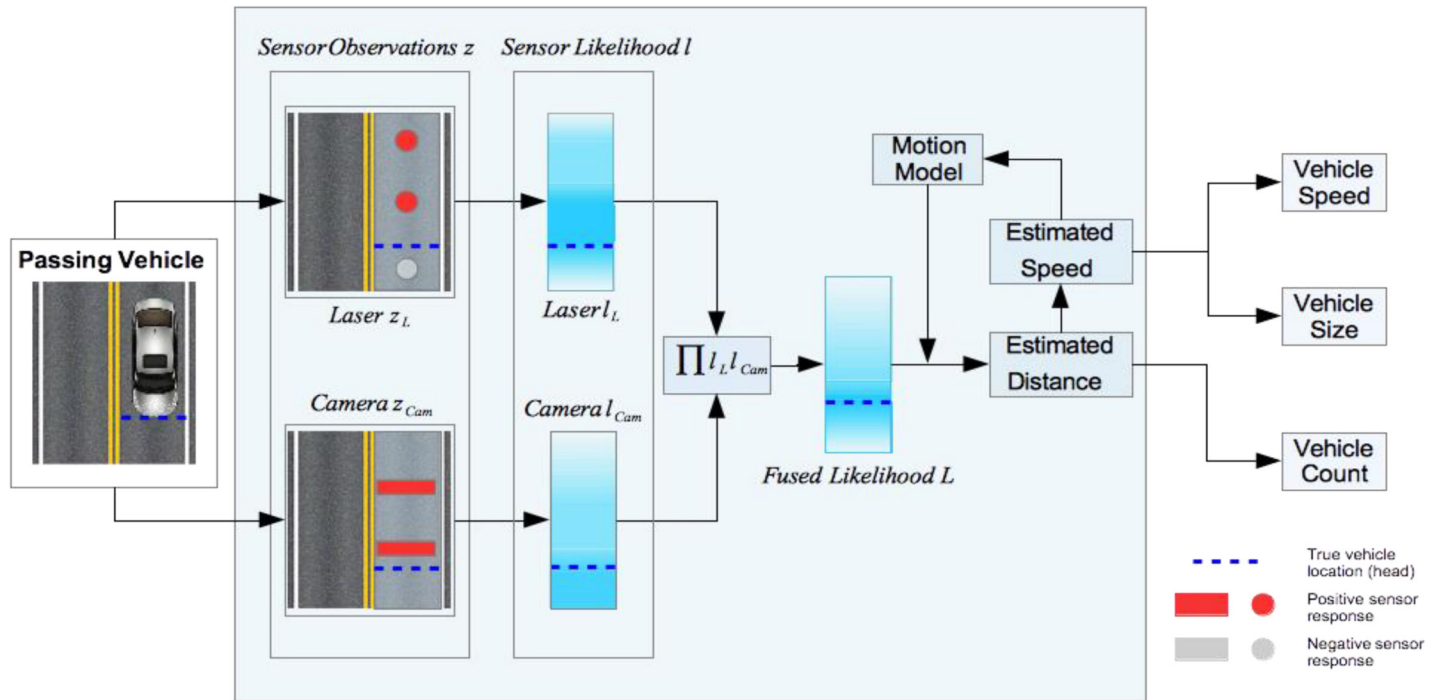


Figure 3. Proposed approach overview

Vision-based traffic monitoring is based on video frame information and can be achieved through different approaches. To minimize the computation complexity while remaining reliability, background subtraction is implemented in the presented technique. In the implementation, similar to Time of Flight (ToF) measurement, multiple regions of interest (ROI) are selected. As shown in Fig. 2, background subtraction has following major steps:

- ROI Selection.** Since the algorithm is based on local information, global modeling is not necessary. ROIs are selected based on real location of the image region.
- Pre-processing.** This step is used to remove noise and resize the ROI images. Each incoming frame will be preprocessed using Gaussian filter or other filters based on image quality.
- Background modeling.** Model of the ROI can be initialized by obtaining different statistical measures such as mean value of a pixel of the first n incoming frames, where the number n is selected based on the quality of the camera, ranging from 10 to 20. The model needs to be updated due to the changing background in real world. While the incoming frame triggers a non-vehicle-detection event, the corresponding ROI will be considered as new background and have higher influence in the background model update. If a vehicle-detection event is triggered, although the model is still updating, the new frame's weight will be much smaller.
- Foreground filtering and foreground energy.** With the updated ROI background model, the foreground energy is obtained by comparison between new frame and the model. A high foreground energy value indicates a possible vehicle or pedestrian occupation in the ROI, which is similar to Time of Flight (ToF) measurement.

Multiple ROIs improve the reliability of vehicle counting and enable the system to conduct vehicle speed estimation and size classification.

Compared with Lidar measurement, the advantage of vision-based method lies in its provision of large amount of information. Normally, vision sensors such as cameras have a large field-of-view, which makes it possible to distinguish small object from large ones that are considered as passing vehicles. A large field-of-view also enables the selection of more ROIs, which can improve the overall reliability of the output, including vehicle counting, speed estimation and classification. On the other hand, vision sensors can provide RGB information, which makes the detection more easy and reliable. The major disadvantage of the vision-based method is its dependence on external illumination, which is hard to be controlled or known ahead. As a result, illumination variation and shadow are always potential factors decreasing the performance of the system. At night, vision sensors have to rely on illumination from light poles or sometimes only from headlights of vehicles. And that significantly reduced the detecting accuracy.

Whilst both methods have common issues such as working in extreme weather condition [4], they have complementary performance in most scenes and a combined approach can potentially lead to more reliable and accurate output. The next section will describe the proposed approach that fuses sensor output from Lidars and cameras.

TRAFFIC MONITORING WITH LIDAR AND CAMERA FUSION USING RECURSIVE BAYESIAN ESTIMATION

Figure 3 shows the overview of the proposed approach for traffic monitoring that conducts vehicle counts, speed estimation and classification based on vehicle size. Observations used include Lidar