Real-Time Video-Based Traffic Measurement and Visualization System for Energy/Emissions

Brendan Tran Morris, Member, IEEE, Cuong Tran, Member, IEEE, George Scora, Mohan Manubhai Trivedi, Fellow, IEEE, and Matthew J. Barth, Senior Member, IEEE

Abstract—The ability to monitor the state of a given roadway in order to better manage traffic congestion has become increasingly important. Sophisticated traffic management systems able to process both the static and mobile sensor data and provide traffic information for the roadway network are under development. In addition to typical traffic data such as flow, density, and average traffic speed, there is now strong interest in environmental factors such as greenhouse gases, pollutant emissions, and fuel consumption. It is now possible to combine high-resolution real-time traffic data with instantaneous emission models to estimate these environmental measures in real time. In this paper, a system that estimates average traffic fuel economy, CO$_2$, CO, HC, and NO$_x$ emissions using a computer-vision-based methodology in combination with vehicle-specific power-based energy and emission models is presented. The CalSentry system provides not only typical traffic measures but also gives individual vehicle trajectories (instantaneous dynamics) and recognizes vehicle categories, which are used in the emission models to predict environmental parameters. This estimation process provides far more dynamic and accurate environmental information compared with static emission inventory estimation models.

Index Terms—Intelligent transportation systems (ITS), real-time energy/emissions estimation, traffic measurement and management, vehicle specific power energy and modal emissions, visual tracking and classification.

I. INTRODUCTION

W
ith the increase in roadway congestion, the ability to monitor the state of the roadway network, by a variety of means, has become critical. Over the last decade, there has been a tremendous amount of research conducted on intelligent transportation systems (ITSs) for advanced traffic monitoring and management. Traffic management centers (TMCs) located around the world are becoming increasingly sophisticated. They bring in data from large networks of sensors for traffic analysis purposes. Efficient operation of these centers requires both an up-to-date view of current conditions for speedy response, and historical data for modeling, planning, and prediction.

A prime example of such a system is California’s Performance Measurement System (PeMS) [1], which gathers measurements from 30,000 inductive loops embedded in the highway and distributed across the state in addition to police incident reports and lane closure information. PeMS has made the fundamental flow, occupancy, and speed data, along with basic traffic calculations, accessible to the research community, thus inspiring novel traffic management solutions.

While capacity and congestion have historically been the major motivating factors of transportation management, new performance metrics have recently garnered attention. In addition to standard traffic metrics, there is a strong interest in traffic-related emissions related to:

1) pollutants (e.g., carbon monoxide (CO), hydrocarbons (HC), nitrides of oxygen (NO$_x$), and particulate matter);
2) greenhouse gases [e.g., carbon dioxide (CO$_2$)];
3) energy (fuel consumption).

Estimating the emissions inventory for vehicles traveling on the roadway network is an active field due to emission requirements from government institutions such as the U.S. Environmental Protection Agency (EPA) and the California Air Resources Board (CARB). Both the EPA and CARB have sophisticated emission models [2], [3] that can be used to determine emissions for specific scenarios, and most roadway planning must utilize these models to determine the impacts of future activity. The transportation community is now beginning to see the value of combining both real-time transportation data and emissions modeling to predict instantaneous emissions or energy usage on a road network on a link-by-link basis. Unfortunately, there is no PeMS-like system for accurate roadway emissions measurements. Attempts have been made to utilize link-based traffic volumes and average speeds along with speed-emissions curves for estimation [4], but these approaches lack sensitivity. They do not account for vehicle profiles and the differences between various vehicle types and instantaneous activity, which drastically affect emissions.

To provide real-time link-based emissions (and fuel economy), this paper has developed the CalSentry system that combines the vehicle classifier and traffic flow analyzer (VECTOR) [5] system, which is a computer vision-based highway measurement system, with vehicle-specific power-based (VSP) [6] energy/emission profiles derived from the comprehensive modal...
emission model (CMEM) [7] and the motor vehicle emission simulator (MOVES) [2]. This system provides subtle vehicle dynamics (instantaneous speed and acceleration) through visual tracking along with categorization of the type of vehicles on the road to accurately estimate vehicle-specific emissions. The system could be deployed within a larger sensor network to provide a streaming data source for traffic management systems such as PeMS, thus providing useful real-time information for policy makers, planners, and health officials, and facilitating further transportation-related emissions research.

II. RELATED STUDIES

The world's rapidly growing and expanding population has led to traffic congestion even on well-planned road networks. This congestion results in a loss of time and productivity and contributes a high economical cost. To tackle the congestion problem without continual road construction projects, traffic management and control approaches have been adopted to better utilize the existing roadway infrastructure. However, the development of effective management and control strategies requires significant data collection. Historical data are needed to learn and develop models, whereas real-time measurements can provide up-to-date indicators of performance for prompt response. In addition, these data are needed over large coverage areas with varied conditions. Effective monitoring systems must be therefore scalable, provide distributed and cooperative sensing, be robust to a wide range of environmental conditions, and have efficient transmission and storage.

The dominant sensor for traffic management has been the inductive loop sensor, which is able to detect the presence of a vehicle based on an induced magnetic field. This simple spot detector provides the count of vehicles that have passed over it (flow) and the amount of activation (occupancy) in a time period. By collecting the loop readings from many sensors, traffic researchers have built complex models for analysis. PeMS [1] which was developed at the University of California, Berkeley and now under the control of the California Department of Transportation provides raw loop readings and fundamental traffic performance measures.

However, inductive loops are costly to install and maintain (only 62% of California's PeMS loop sensors are in working order), making the search for alternative sensing solutions [8] or augmentation schemes [9] appealing. Video cameras have emerged as a popular ITS device within TMCs for human operator monitoring and verification. Cameras provide complementary analysis that is difficult to obtain using traditional sensors. They provide wide spatial coverage or field-of-view (FOV) which captures higher order dynamics in vehicles and traffic, and rich information content for more complex analysis (e.g., classification and trajectory analysis [10]–[12]).

A. Visual Traffic Monitoring

Highway monitoring is one of the oldest applications for vision researchers. The inherent structure of roads coupled with a vehicle's rigid body constrains the vision processes. Early vision-based traffic monitoring researchers looked to mimic the popular inductive loop sensor counts by manually defining virtual loops in the camera FOV [13]. While both easy to manage and effective, the virtual loops did not take advantage of the spatial coverage afforded by the camera, reducing the wide FOV into several small point sensors. Subsequently, most researchers began to focus on moving object detection and tracking [14]. In this paradigm, a count is generated for every tracked vehicle [5], [15].

However, video processing is not without challenges. Camera placement and view are critical for successful deployment. Imaging provides roadway coverage over long distances but also causes perspective distortion, which greatly affects the apparent size of vehicles and leads to occlusion. Significant effort has gone into developing detection methods that can resolve occlusion such as feature grouping in the image plane [15] or using multiple homography transformations in 3-D space [16].

Cameras also have difficulties dealing with changing environmental conditions such as illumination changes. Researchers have attempted to distinguish cast shadows on the roadway from vehicles using shadow detection and suppression techniques [17]. In addition to shadows caused by lighting, it is quite difficult to operate vision systems at night without the use of costly low-light sensitive or infrared cameras, reducing the effective operating time.

B. Vehicle Classification

The switch to video-based traffic monitoring is particularly useful for vehicle classification because of the appearance information contained in an image. Loop-like sensors only generate a 1-D signature, which makes it difficult to resolve differences between vehicles (large and fast moving versus small and slow). Generally, these systems count the number of axles to distinguish between large and small vehicles.

Buch et al. [18] devotes two large sections to top-down (object-based) and bottom-up (part-based) visual classification techniques for urban traffic. It is noted that, classification often degenerates into a detection problem because the techniques are designed for matching. Even [16], which is a recent vehicle classification paper, only makes a distinction between two different classes of vehicles based on stable features.

An earlier work by Gupte et al. [19] classified vehicles by their length. However, length measurement precision was found to be low and not flexible to camera views. More detailed classification has been tackled using shape and appearance techniques [5], using a linearity feature [20] or explicit vehicle models. Edge matching techniques have been designed for use with 3-D wire-frame vehicle models [21]. Although the four wire models were quite simple and had low resolution, they only operated at 5 Hz. More generic and adaptive 3-D models have been explored to provide a deformable vehicle model with higher resolution [22]. Results were shown on a difficult 5-class problem where a distinction was made between two- and four-door sedans.
Despite these efforts, it is still difficult to leverage high-resolution imaging while maintaining the computational efficiency required for real-time monitoring.

C. Emissions/Energy

Environmentalists and health officials have long been concerned with the effects of air pollution on air quality, but only recently has there been a major shift in focus to the transportation sector. Traditionally, monitors that measure pollutants in the air in parts per million or per billion are used to evaluate air quality. Unfortunately, the number of monitoring sites is limited and, therefore, cannot accurately depict the spatial distribution of pollutants. In addition, the sparse data does not necessarily express traffic emissions specifically since measurements are not close to the source. In fact, mobile measurement systems have shown pollutants found near highways significantly exceed the maximum values reported by fixed measurement monitors [23].

To isolate the effects of transportation emissions, a wide range of modeling techniques has been adopted [4]. The simplest models are easy to calculate and rely on average speed but do not account for real-world driving characteristics [24]. More complex modal models operate at a higher time resolution (in seconds) and account for more detailed vehicle and traffic characteristics, such as specific engine operation and vehicle movements. The modal models have recently gained traction because supporting data can more easily be obtained via Global Positioning System (GPS) [25], and they can be integrated into microscopic traffic simulation models [26]. While promising, these techniques are difficult to scale to large areas. Communication networks must be established, and there must be a high penetration rate for GPS-based emission calculations. Microscopic traffic and emissions simulation is computationally expensive for large networks because trajectories and emissions must be calculated for each vehicle at each simulation time step over the entire network and are affected by errors in the traffic models themselves.

Currently, there is no tool available to calculate transportation-related emissions in real time for a large number of vehicles. In addition, there is no way to present this information to stakeholders to manage or plan future decisions.

III. CalSentry Highway Emission Management System

This paper presents CalSentry, the first real-time integrated highway transportation measurement and management system for emission/energy estimation. This vision-based system combines four major components, as shown in Fig. 1:
1) visual traffic measurement;
2) dynamics-based emission estimation;
3) real-time visualization;
4) database for record keeping.

The elements shown in purple in Fig. 1, comprise the parts of the VECTOR highway monitoring module [5], which is a visual tracking and analysis system suitable for distributed traffic understanding [27]. The emission modeling and estimation block, shown in green in Fig. 1, utilizes VECTOR analysis to estimate the amounts of pollutants produced by vehicles on the roadway using real-time emission modeling [2], [7]. Adhering to a framework designed for thematic contextualization [28], measurements and model estimations are stored and utilized for appropriate visualization of system output (the orange block in Fig. 1). As a testament to its robustness, the CalSentry system has been in continuous daily operation at a single site since early 2011, collecting highway data during the daylight hours when vehicles are visible.

At the host site, a traffic operator can view live highway video feeds, i.e., both in raw and in processed form. Fig. 2(a) shows the output of the VECTOR system. The highway video is processed in real time to display object detection and tracking results. Tracked vehicles have a color-coded bounding box to indicate the current emission score (based on dynamics and vehicle type) with red indicating a higher score. In addition, on the left side of Fig. 2(b), in red and white, are moving
time-series plots of highway flow. The plots are updated every video frame to give the instantaneous count of vehicles traveling either north or south while providing a short 30-s history. Using the VSP approach [6] (described in Section V-B), the roadway emissions are estimated based on tracking information and vehicle type (determined by visual processing). Similar to the moving flow plots, the instantaneous VSP value with 30-s history is displayed in yellow and blue. The time-series plots are intended to show the evolution of conditions on the road. On the right of Fig. 2(b) are two bins representing the total accumulated emissions in the north and south directions in a sliding 30-s window. The bins are color coded as yellow, orange, and red, to indicate low, medium, and high amounts respectively, of greenhouse emissions.

The diagnostics plots of Fig. 2(b) provide immediate up-to-date measurements but are quite variable due to traffic congestion conditions. The emission score, which includes the four greenhouse gases and pollutants \( (\text{CO}_2, \text{CO}, \text{HC}, \text{NO}_x) \), is accumulated and aggregated over 30-s increments for more stable timescales. By adopting the standard loop detector aggregation scheme, emission statistics can be directly correlated and used in the same way as the traditional highway measures of flow, occupancy, and speed. In fact, they could be combined not only with VECTOR highway measurements but also any loop data, such as those warehoused by PeMS.

The final output component in the CalSentry system is a remote user interface. A public website, utilizing the Google Maps API, was constructed to provide interested parties access to the emission measurements. Fig. 2(c) shows the map with a color-coded view of a highway link. Using the same color scheme as the bins above, i.e., \( \{\text{yellow}, \text{orange}, \text{red}\} = \{\text{low}, \text{medium}, \text{high}\} \), the highway is colored to indicate the 30-s aggregate emission value for the link. In this snapshot, the northbound direction is colored yellow, indicating a low emission level, whereas the southbound direction is orange, indicating a medium emission level. The map is similar to the more familiar navigation maps that have been color coded for speed. Similar emission coverage could be provided with more CalSentry nodes to give a better sense of the current emission conditions in a city or a region.

Although not visualized, an important part of the CalSentry system is the database for historical record keeping. With increased coverage and data, the database will be valuable for displaying trends (e.g., the evolution of emission “hot spots” in a city over the course of a day) and as input to support larger scale emission modeling. This data will help transportation engineers and policy makers understand how commutes affect air quality and determine how to best manage or build future roads.

IV. VEHICLE CLASSIFIER AND TRAFFIC FLOW ANALYZER

TRAFFIC MONITORING

The following describes highway traffic monitoring with VECTOR [5]. The camera-based system is able to detect, track, and identify the type vehicles on the roadway. In addition, it produces traffic measurements, similar to traditional loop detectors, in real time.

A. Vehicle Detection and Tracking

Unlike a loop detector that is a spot sensor, cameras observe a vehicle over a period of time while it travels through the camera’s FOV. Each video frame provides a new view that is used to describe the appearance of a vehicle and its dynamics.

The VECTOR systems utilizes a single camera to monitor both directions of a busy four-lane highway. Vehicles are detected as moving regions using background subtraction and tracked using a global nearest neighbor optimization that accounts for both the dynamics and appearance. The trajectory of vehicle \( i \)

\[
F_i = \{f_1, \ldots, f_T\}, \quad \text{with} \quad f_t = [x, y, u, v]^T
\]
is the sequence of positions and velocities that describes the vehicle dynamics. The morphological appearance vector

\[ M_i = [\eta_1, \ldots, \eta_{15}]^T \]  

\[ \{\text{area, breadth, compactness, elongation, perimeter, convex hull perimeter, length, long and short axes of fitted ellipse, roughness, centroid, the four first and second image moments}\} \]

encodes the shape appearance of the particular vehicle.

### B. Vehicle Classification

Using the appearance vector \( M_i \), VECTOR classifies each vehicle into one of the eight different types: sedan, pickup, sport utility vehicle (SUV), van, semi, truck, bike, and merged, as shown in Fig. 3. At a particular instant, \( M_i \) is transformed using Fisher linear discriminant analysis (LDA) and compared with a vehicle database using a weighted \( K \) nearest neighbor (wkNN) technique to produce class weights. The final vehicle label \( L_i \) of a track is determined after iterative refinement with each new video frame.

1) **Feature Transformation:** Appearance features were projected using LDA [29] to better separate vehicle classes in a lower dimensional space.

The size \( N \) training database \( D = \{D_1, \ldots, D_C\} \) is divided into smaller \( N_c \) sized sets for each of the \( C \) classes. The mean of the full data set is \( \mu = (1/N) \sum_{i=1}^{N} M_i \), and the class-specific means are \( \mu_c = (1/N_c) \sum_{i=1}^{N_c} M_i \). The Fisher maximization criteria leads to the generalized eigenproblem \( S_B \omega = \lambda S_W \omega \), where \( S_B \) is the between class scatter matrix and \( S_W \) is the within class scatter matrix as follows:

\[ S_B = \sum_{c=1}^{C} N_c(\mu_c - \mu)(\mu_c - \mu)^T \]  
\[ S_W = \sum_{c=1}^{C} \sum_{M_i \in D_c} (M_i - \mu_c)(M_i - \mu_c)^T. \]  

An LDA projection matrix is obtained by retaining the top \( \rho = 5 \) eigenvectors that is used to obtain a compact representation for vehicle appearance, i.e.,

\[ m_t = PM_i = [\omega_1, \ldots, \omega_1, \ldots, \omega_2]M_i. \]  

2) **Detection Classification:** After projection, a vehicle is compared with a database using a modified NN classifier. The wkNN rule has more robustness to noise and outliers by utilizing a soft class assignment, i.e.,

\[ w_c(t) = \frac{1}{\sum_{t=1}^{T} ||m_t - m(t)||}. \]  

The weight \( w_c(t) \) indicates the likelihood of \( m(t) \) belonging to class \( c \) based on the similarity to the \( K = 5 \) closest training examples.

3) **Track-Based Classification Refinement:** A track-based refinement scheme is adopted to exploit information redundancy from multiple frames for improved classification. Given \( T \) images of a vehicle during tracking, the track class label is found by maximum likelihood estimation as

\[ L_i = \arg\max_{c} \sum_{t=1}^{T} p(m(t)|c) \]  
\[ = \arg\max_{c} \sum_{t=1}^{T} \ln w_c(t). \]  

Using (7), the label is refined with each new frame by calculating the detection weight (6) and iteratively updating the log of the class distribution to leverage any additional appearance evidence.

### C. Traffic Statistics

Using trajectory information, the time series of fundamental highway usage parameters, analogous to those obtained from conventional loop detectors, is collected in real time. The VECTOR system delivers average speed (in miles per hour), density (\# vehicles/distance), and flow (\# vehicles/time) in 30-s intervals, as shown in Fig. 4.

1) **Directional Measurements:** The highway speed [see Fig. 4(a)] is directly measured as the average velocity of all vehicles seen in the 30-s interval. The roadway is calibrated based on ground plane homography to convert pixels/second image tracking into miles per hour. The density indicates how crowded a roadway is and is computed by counting the number of vehicles in the camera view normalized by the roadway length. A loop detector cannot directly measure density because it is a spot sensor; instead, density is inferred based on occupancy and flow. Traffic flow is a count of the number of passing vehicles in a 30-s time interval. VECTOR produces the flow statistic by counting vehicle tracks as they exit the camera FOV in a manner similar to loop detectors.

2) **Lane-Level Measurements:** In Fig. 4(a), the traffic in northbound and southbound directions are compared during daylight hours. In this section of road, the southbound traffic is affected during the evening commute hours of 14:00–16:00. The causes can be further investigated by focusing on the lane-level traffic measurements shown in Fig. 4(b). During tracking, the lane number is determined based on position in the image. The density of vehicles in the fast lane (lane 1) dramatically

![Fig. 3: Sample images from each vehicle class. (a) Sedan. (b) Pickup. (c) SUV. (d) Van. (e) Semi. (f) Truck. (g) Bike. (h) Merged.](image-url)
increases during evening and results in a significant drop in
driving speed.

3) Vehicle-Level Measurements: Since VECTOR is based on
video technology, rich contextual information not obtained with
loop detectors can be extracted to further study traffic condi-
tions. The loop-like traffic statics, speed, density, and flow are fur-
ther categorized based on the vehicle type. Fig. 4(c) shows the
proportion of different vehicle types on the road over the course
of a day. As expected, the vehicle-specific speed measurements
reflect lower speeds for the large classes (semi and truck).

V. EMISSION/ENERGY ESTIMATION

Current emphasis on environmental issues, such as air pollu-
tion, greenhouse gases, and energy consumption has fueled re-
search in areas such as “green” vehicle technologies, alternative
fuels, and ITS. This has resulted in the commercial success of
hybrid automobiles and the reintroduction of consumer electric
vehicles as a way to curb emissions and energy consumption.
ITS advances in traffic monitoring and data collection help
improve traffic characterization and management; however, it
is still unclear exactly how the transportation network affects
emissions and how emissions contributions can be character-
ized by location, time, or mobile source.

It is generally difficult to measure pollution from specific
mobile sources under real-world conditions due to the mixing
and dispersion of emissions and contributions from additional
sources such as other vehicles, nearby factories, and secondary
pollutants. These issues are further influenced by environmental
factors, such as meteorological conditions and topographic
features. Some methods of emission measurements, such as
tunnel studies, remote sensing, and portable emission moni-
toring systems, address these problems to some extent. These
methods, however, either do not isolate emissions from specific
vehicles, are very limited in testing locations, or cannot be
practically applied to large sets of vehicles. A method to better
estimate emissions attributed to transportation and even specific
vehicle classes in real time is useful for traffic management and
policy makers, and health organizations.

Using a vision-based traffic management system, which is
able to accurately track individual vehicles (collecting dynamic
driving patterns) and determine their type, it is possible to
estimate the emissions and energy consumption from specific
vehicles on the road. By accumulating emissions data over time,
a real-time map can be formed to indicate the level of pollutants
on our roadways.

A. Vehicle Class Emission Modeling

To accurately determine the amount of emissions or fuel
usage from a particular vehicle, it is necessary to know cer-
tain vehicle characteristics such as weight, fuel type, engine
displacement, after-treatment technology, vehicle model year,
vehicle age, and how the vehicle is being operated (the driving
profile). Unfortunately, it is not possible to determine many of
these vehicle characteristics using conventional traffic cameras.
The resolution of these setups along with the vast number of
vehicles on the road with varying characteristics makes this
level of data collection almost impossible without the use of
other identifying techniques such as radio-frequency tags or
license plate recognition. As shown earlier, it is, however,
possible to distinguish between different classes of vehicles
using conventional traffic cameras. Each class of vehicles has
different emission properties that are generally related to vehi-
cle size and type.

At the current time \( t \), an instantaneous emission value
\( E_{\text{pol}}(t) \) for pollutant \( \text{pol} \) can be estimated for each vehicle
based on the vehicle class \( L \) and dynamic profile \( F(t) \) as follows:

\[
E_{\text{pol}}(t) = h_{F}(L, F(t)).
\]

The functional mapping \( h_{F} \) specifies how emissions are ob-
tained from the vehicle information and must be specified or
modeled.

B. VSP Approach

There are various approaches to estimating vehicle emis-
sions depending on the scope of the analysis and the available
data. Traditional emission modeling techniques utilize average
speed-based emission rates for estimation. One of the funda-
mental drawbacks of this modeling approach is that speed alone
is not a good predictor of emissions since speed under various
levels of acceleration will result in a wide range of emissions.
Acceleration is an important factor in the estimation of vehicle
load, which is well correlated with fuel use and consequently
emissions. To take advantage of this additional level of detail, VSP [6] was used as the basis for emission rates. VSP is defined as the instantaneous power to move a vehicle per the mass of the vehicle. The calculation for VSP in kilowatts/metric tonne is based on the following equation, simplified from the power demand terms for a moving vehicle:

\[ VSP(t) = v(t) (1.1a(t) + g \sin(\theta) + gC_r) + \frac{\rho_0 C_d A_f v^3(t)}{2M} \]  

(9)

where

\[ v = \text{speed [m/s]}; \]
\[ a = \text{acceleration [m/s}^2]; \]
\[ g = \text{gravity} = 9.8 [\text{m/s}^2]; \]
\[ \theta = \text{grade [radians]}; \]
\[ C_r = \text{coefficient of rolling friction}; \]
\[ \rho_0 = \text{density of air} = 1.2 [\text{kg/m}^3] \text{ at sea level and 20 °C}; \]
\[ C_d = \text{coefficient of aerodynamic drag}; \]
\[ A_f = \text{frontal area [m}^2]; \]
\[ M = \text{mass [kg]}. \]

The vehicle dynamic information \([v(t), a(t)]\) is encoded in the trajectory \(F(t)\). The parameter values in Table I are the approximate values used for the seven VECTOR vehicle types (merged is excluded) based on the NCHRP 25-11 [30] data set and values found in literature.

Using the VSP approach, emissions are estimated by modifying (8) as

\[ E_{pol} = h(L, VSP(t)) \]  

(10)

where class \(L\) represents the VECTOR vehicle type categories (7) and \(VSP(t)\) encodes the vehicle emission-dynamics relationships. The mapping \(h\), to produce emission values [31], is modeled based on CMEM and MOVES, as described in the following.

C. Vehicle-Specific Emission Tables

After accounting for the vehicle type and driving dynamics with the VSP, the mapping \(h\) from VSP to a particular pollutant emissions was determined. The VSP mapping is generated using two different simulation models: CMEM [7] and the EPA’s MOVES [32]. Emission tables developed for this project provide instantaneous emission rates for VSP values between 0 and 40 kW/T and can be conveniently applied both in real time and in post-processing. For each vehicle and at each time step, a VSP value is calculated using (9) with vehicle-class-specific constants of Table I. (An alternative, more computationally intensive approach would be to use individual vehicle trajectories as input to CMEM.)
The MOVES modeling methodology is based on VSP-binned emission rates. It is applicable at the microscale level for conformance and emissions of future vehicle models and applications of new technology (e.g., after treatment devices).

VSP and emission values are calculated for each CMEM vehicle category using the Federal Test Procedure 75, US06 Supplemental Federal Test Procedure, and Modal Emission Cycle [30] driving schedules. Vehicle population data from CARB’s The EMission FACtors (EMFAC) model [3] model for San Diego County and calendar year 2010 is used to approximate fleet distributions for CMEM categories. CMEM categories are further grouped into the VECTOR vehicle classes for compositing. Fig. 5 shows compositing results for the VECTOR pickup class. In this figure, the light green lines show VSP emission results for individual CMEM vehicle categories within the VECTOR pickup class, and the black line shows the weighted composited VSP-based emission values for the VECTOR pickup class. The emission values are binned in 1 kW/T bins, and in the figure, the first bin represents VSP values between 0 and 1 kW/T.

In addition to the VECTOR sedan, pickup, and semi classes, specific van and SUV categories were developed to model emissions from these two vehicle types with the CMEM model. To determine van and SUV CMEM categories, individual van and SUV vehicles from the NCHRP 25–11 database from the original CMEM project were identified (20 SUV vehicles and 37 vans), calibrated, and modeled using CMEM.

2) EPA MOVES Model: The remaining two VECTOR categories, i.e., truck and motorcycle, are not supported by the CMEM model and are instead modeled using the 2010 MOVES database, which is the EPA’s latest mobile source emission model. The VECTOR truck category is a broad category and encompasses a range of visually similar vehicle types such as buses, garbage trucks, and medium-heavy trucks. For the most part, these vehicles are large diesel-engine-driven vehicles, and for this paper, this class was approximated as an urban bus according to the EPA’s approximation for 1996–2006 class 48 vehicles from heavy-heavy duty (HHD) vehicles [32]. The motorcycle class is taken directly from the motorcycle base 48 vehicles from heavy-heavy duty (HHD) vehicles [32].

The vehicle classifier was only evaluated during the daylight hours of a single day (approximately 5:00–19:00) because the detection and tracking does not work at night due to poor lighting and headlight reflection on the road surface. Each hour, a 5-min video clip was saved for manual annotation. Both the type of vehicle and the lane of travel were recorded, resulting in 6491 total tracks.

A. Visual Vehicle Type Classification

The vehicle classifier was only evaluated during the daylight hours of a single day (approximately 5:00–19:00) because the detection and tracking does not work at night due to poor lighting and headlight reflection on the road surface. Each hour, a 5-min video clip was saved for manual annotation. Both the type of vehicle and the lane of travel were recorded, resulting in 6491 total tracks.

The percentage of the tracks classified as the correct vehicle type was 78.44%. The full-day confusion matrix is shown in Table II. The classifier has difficulties with the van and truck classes because they are quite similar in appearance to SUV and semi classes, respectively. Table III gives the number of test vehicles (count) and classification accuracy for each hour of the experiment. The results from a single 5-min clip is generally in the 80% range, except between 08:00–11:00. There is a significant performance drop during these morning hours due to adverse lighting conditions, which caused large cast shadows.

### Table II

| CLASSIFICATION CONFUSION MATRIX (78.4% ACCURACY ON 6491 TEST TRACKS) |
|---|---|---|---|---|---|---|---|
| **sedan** | **pickup** | **suv** | **van** | **semi** | **truck** | **bike** | **merged** |
| **sedan** | 2726 | 127 | 202 | 55 | 0 | 0 | 1 | 0 |
| **pickup** | 40 | 374 | 52 | 24 | 0 | 14 | 0 | 4 |
| **suv** | 411 | 113 | 1147 | 172 | 0 | 3 | 0 | 4 |
| **van** | 15 | 11 | 54 | 83 | 0 | 6 | 0 | 7 |
| **semi** | 0 | 0 | 0 | 0 | 26 | 1 | 1 | 1 |
| **truck** | 1 | 5 | 1 | 11 | 36 | 0 | 0 | 0 |
| **bike** | 1 | 0 | 0 | 0 | 0 | 18 | 0 | 0 |
| **merged** | 7 | 7 | 5 | 7 | 6 | 10 | 3 | 677 |
| **% correct** | 85.2 | 58.7 | 78.5 | 24.0 | 65.0 | 39.6 | 85.7 | 97.7 |
TABLE III

<table>
<thead>
<tr>
<th>time</th>
<th>count</th>
<th>sedan</th>
<th>pickup</th>
<th>suv</th>
<th>van</th>
<th>semi</th>
<th>truck</th>
<th>bike</th>
<th>merged</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>06:21</td>
<td>405</td>
<td>94.9</td>
<td>59.5</td>
<td>81.9</td>
<td>31.6</td>
<td>50.0</td>
<td>33.3</td>
<td>0</td>
<td>97.5</td>
<td>81.5</td>
</tr>
<tr>
<td>07:19</td>
<td>497</td>
<td>96.2</td>
<td>32.5</td>
<td>83.2</td>
<td>06.7</td>
<td>66.7</td>
<td>25.0</td>
<td>100</td>
<td>98.0</td>
<td>84.7</td>
</tr>
<tr>
<td>08:17</td>
<td>530</td>
<td>61.2</td>
<td>33.3</td>
<td>91.9</td>
<td>14.3</td>
<td>50.0</td>
<td>50.0</td>
<td>100</td>
<td>96.7</td>
<td>67.6</td>
</tr>
<tr>
<td>09:15</td>
<td>444</td>
<td>53.2</td>
<td>38.7</td>
<td>82.3</td>
<td>53.9</td>
<td>37.5</td>
<td>23.1</td>
<td>100</td>
<td>96.8</td>
<td>63.7</td>
</tr>
<tr>
<td>10:13</td>
<td>357</td>
<td>36.8</td>
<td>26.7</td>
<td>77.3</td>
<td>26.7</td>
<td>71.4</td>
<td>40.0</td>
<td>100</td>
<td>93.9</td>
<td>51.0</td>
</tr>
<tr>
<td>11:11</td>
<td>417</td>
<td>63.4</td>
<td>47.2</td>
<td>90.4</td>
<td>28.0</td>
<td>66.7</td>
<td>33.3</td>
<td>-</td>
<td>89.6</td>
<td>68.6</td>
</tr>
<tr>
<td>12:09</td>
<td>432</td>
<td>86.0</td>
<td>71.7</td>
<td>82.6</td>
<td>48.0</td>
<td>50.0</td>
<td>37.5</td>
<td>100</td>
<td>96.9</td>
<td>80.8</td>
</tr>
<tr>
<td>13:08</td>
<td>393</td>
<td>95.6</td>
<td>76.3</td>
<td>83.5</td>
<td>39.1</td>
<td>100</td>
<td>50.0</td>
<td>-</td>
<td>97.9</td>
<td>87.0</td>
</tr>
<tr>
<td>14:06</td>
<td>449</td>
<td>96.9</td>
<td>77.8</td>
<td>84.2</td>
<td>18.2</td>
<td>-</td>
<td>66.7</td>
<td>100</td>
<td>94.9</td>
<td>86.2</td>
</tr>
<tr>
<td>15:04</td>
<td>492</td>
<td>96.0</td>
<td>76.4</td>
<td>81.9</td>
<td>23.1</td>
<td>100</td>
<td>99.1</td>
<td>100</td>
<td>100</td>
<td>85.4</td>
</tr>
<tr>
<td>16:02</td>
<td>553</td>
<td>97.1</td>
<td>66.2</td>
<td>76.0</td>
<td>24.0</td>
<td>100</td>
<td>55.6</td>
<td>100</td>
<td>100</td>
<td>85.7</td>
</tr>
<tr>
<td>17:00</td>
<td>630</td>
<td>99.1</td>
<td>65.5</td>
<td>62.0</td>
<td>03.6</td>
<td>-</td>
<td>0</td>
<td>100</td>
<td>94.5</td>
<td>83.0</td>
</tr>
<tr>
<td>17:45</td>
<td>297</td>
<td>89.0</td>
<td>75.9</td>
<td>52.8</td>
<td>10.0</td>
<td>-</td>
<td>100</td>
<td>67.0</td>
<td>97.7</td>
<td>76.0</td>
</tr>
<tr>
<td>18:45</td>
<td>382</td>
<td>96.0</td>
<td>57.9</td>
<td>73.9</td>
<td>10.5</td>
<td>-</td>
<td>100</td>
<td>50.0</td>
<td>97.9</td>
<td>84.6</td>
</tr>
<tr>
<td>19:43</td>
<td>222</td>
<td>95.0</td>
<td>77.8</td>
<td>78.5</td>
<td>0</td>
<td>100</td>
<td>100</td>
<td>-</td>
<td>100</td>
<td>86.5</td>
</tr>
</tbody>
</table>

B. Vision-Based Traffic Statistics

Although real-time data are needed to understand current conditions, historical measurements provide the data for model-

ing and provides a deeper understanding of higher order effects. The observed flow and speed over a given week are shown in Fig. 6. The differences between weekday and weekend traffic patterns are quite clear. During the weekdays, there is a large increase in demand between the evening commute hours of 15:00 and 17:00. The increased flow rate causes congestion and results in a large drop in speed. On the weekdays, there is a 50% decrease in average speed during the commute hours, whereas the weekends show no significant speed difference. By storing measurements in a database, they can be utilized to learn and model variations in traffic patterns and behavior.

During the daylight test, 98% of vehicles traveling south (closest to the camera) were identified in the correct lane. The northbound direction, which is much further away and suffers from more perspective distortion, only had 93% total accuracy in lane assignment. The northbound 3-lane had the poorest performance at 84.4%. This implies high lane-level accuracy is possible with the proper camera-road configuration.

C. Real-Time Vehicle Emission Aggregation

Using Table I along with VSP-based emission profiles (highlighted in Fig. 5), the emissions from each vehicle were calculated at 10 Hz and archived. The 10-Hz update rate was chosen as a compromise to smooth variability between VECTOR’s high 30-fps rate and the natural 1-Hz operation of the microsimulation model, which obscures the trajectory profile. Emission measurements were aggregated into 30-s increments before archival to provide more stable timescales that match PeMS loop detector rates.

Fig. 7(a) shows the emissions (CO$_2$, CO, HC, and NO$_x$) rates (in grams per second) in the southbound direction of the highway. The HC and NO$_x$ rates are scaled 50× and CO by 3× for plotting purposes since the CO$_2$ rate is much higher. The time-series plots have a number of spikes that should not come as a surprise because of the nature of traffic. During a particular 30-s time interval, the number of and types of cars and driving style, which greatly impacts emission production, is variable. In Fig. 7(b), the emissions are aggregated over a longer 5-min time period. The emission measurements are significantly more stable at this timescale and provides a better indication of the daily patterns. Around the 16:00–17:00 time period, there is a drop in emissions due to congestion. At this time, vehicles move slower, and the VSP is greatly influenced by speed.
Fig. 7. Highway emission estimates for 30-s and 5-min aggregates. (a) 30-s CalSentry. (b) 5-min CalSentry. (c) 30-sec PeMS. (d) 5-min PeMS.

1676 IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 13, NO. 4, DECEMBER 2012

Further work will need to take into account idling emissions.

The emissions measurements generated by CalSentry are compared with loop-detector-based estimates, shown in Fig. 7(c) and (d). These plots were generated by using the PeMS speed measurement (no acceleration) in (9). The PeMS flow value was divided into the VECTOR vehicle classes based on the EMFAC registration distribution resulting in fixed ratios of all vehicles at all times. The 5-min aggregates are of similar scale and have the drop-off during the evening commute. However, notice that the loop data produces a much smaller gap between the CO and NO\textsubscript{x} rates, which requires further investigation.

CalSentry results are generally more variable and result in slightly higher estimated emission rates than the loop estimates solely based on average speed.

D. Future Work

Future work is needed to evaluate the validity of the emission estimates, which will require more sophisticated techniques such as tunnel measurements or instrumented vehicles. Although it is expected that the accuracy of emission estimates with the CalSentry system will be better than loop-based estimates due to the inclusion of acceleration terms and that the system will significantly improve the quality of on-road vehicle emission estimates, there are several factors that impact prediction accuracy. These factors include misclassification of vehicle categories, lumping of certain vehicle types (e.g., truck category), unknown vehicle operating weights (particularly for heavy duty vehicles), unknown vehicle conditions (bad catalysts, tampered emission controls, engine malfunctions, age of the vehicle), and the representativeness of emission factors that will vary based on the test-vehicle sample size used to develop them. However, these same factors impact loop-based estimation, and loops typically do not benefit from driving profiles or diverse vehicle classification.

VII. CONCLUDING REMARKS

This paper introduced the first vision-based system for traffic monitoring and emission/energy measurement. The CalSentry system integrates live video processing, emissions modeling, and historical data into a visualization framework for providing real-time emissions. Real-time vision-based tracking was used to obtain dynamics measurements and vehicle-type classification. A VSP-based approach, utilizing CMEM and MOVES
emission models, converted the vehicle class and dynamic driving profile into CO₂, CO, HC, and NOₓ emission estimates. A public website provided a color-coded map based on the current emission conditions on a highway link that could be extended for wider area coverage and used for policy decisions and real-time traffic management.

ACKNOWLEDGMENT

The authors would like to thank the reviewers for their insightful comments and all their colleagues and students at CVRR and CE-CERT who contributed to this system.

REFERENCES


Brendan Tran Morris (M’10) received the B.S. degree from the University of California, Berkeley, in 2002 and the Ph.D. degree from the University of California, San Diego, in 2010. He is currently an Assistant Professor of electrical and computer engineering with the University of Nevada, Las Vegas. His main research interests include real-time sensing and processing for understanding environments and situations with emphasis in transportation. His current research interests include unsupervised machine learning for recognizing and understanding activities; real-time measurement, monitoring, and analysis; and driver assistance and safety systems.

Dr. Morris received the IEEE Intelligent Transportation Systems Society Best Dissertation Award in 2010 for his dissertation research on “Understanding Activity from Trajectory Patterns.”

Cuong Tran (M’12) received the B.S. degree in computer science from Hanoi University of Technology, Hanoi, Vietnam, in 2004 and the M.S. and Ph.D. degrees in computer science from the University of California, San Diego, (UC San Diego) in 2008 and 2011, respectively. He is currently a Researcher with the Computer Vision and Robotics Research Laboratory, UC San Diego. His research interests include vision-based analysis of human posture and activity at different levels of detail for interactive applications, intelligent driver assistance, advanced human–machine interfaces, and human behavior prediction.

Dr. Tran is a Fellow of Vietnam Education Foundation.
George Scora received the B.S. degree in environmental engineering and the M.S. and Ph.D. degrees in chemical and environmental engineering from the University of California, Riverside (UCR), Riverside, in 1996, 2005, and 2012, respectively.

As a student, he was involved in work with the Office of Transportation and Air Quality, U.S. Environmental Protection Agency, Ann Arbor, MI, which focused on characterizing emissions for both light-duty and heavy-duty vehicles for the Motor Vehicle Emission Simulator model. For 15 years, he has been with the Transportation Systems Research Group, College of Engineering Center for Environmental Research and Technology, UCR, where he is currently a Project Scientist. He has been heavily involved in the development of UCR’s Comprehensive Modal Emissions Model. His research interests include vehicle emission modeling, traffic simulation, intelligent transportation systems, and other transportation-related energy and air quality issues.

Mohan Manubhai Trivedi (F’11) received the B.E. (Hons.) degree in electronics from the Birla Institute of Technology and Science, Pilani, India, in 1974, and the M.S. and Ph.D. degrees in electrical engineering from Utah State University, Logan, in 1976 and 1979, respectively.

He is currently a Professor of electrical and computer engineering and a Founding Director with the Computer Vision and Robotics Research Laboratory and Laboratory for Intelligent and Safe Automobiles, University of California, San Diego. He also serves as a consultant to industry and government agencies in the U.S. and abroad, including the National Academies, major automobile manufacturers and research initiatives in Asia and Europe. He and his team are currently working on research in machine and human perception, machine learning, human-centered multimodal interfaces, intelligent transportation, driver assistance, and active safety systems.

Prof. Trivedi is a Fellow of the International Association of Pattern Recognition (for contributions to vision systems for situational awareness and human-centered vehicle safety) and SPIE (for distinguished contributions to the field of optical engineering).

Matthew J. Barth (SM’00) received the B.S. degree in electrical engineering/computer science from the University of Colorado, Boulder, in 1984, and the M.S. and Ph.D. degrees in electrical and computer engineering from the University of California, Santa Barbara, in 1986 and 1990, respectively.

From 1990 to 1991, he conducted research in robot perception and navigation with Osaka University, Osaka, Japan, as a Visiting Research Scholar. He is currently a Professor with the Center for Environmental Research and Technology, University of California, Riverside. His research interests include intelligent transportation systems, transportation/emission modeling, vehicle activity analysis, electric vehicle technology, robotics, and advanced sensing and control.

Dr. Barth is currently the Vice President of Conferences of the IEEE Intelligent Transportation System Society (ITS) and an Associate Editor for the IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS. He has served as a member of the ITS Society Board of Governors and as a Program Chair for the ITSC-08, ITSC-09, and FISTS 2011 conferences. He is also active in several committees of the Transportation Research Board including the Transportation Air Quality Committee and the ITS Committee.