December 15, 2016
RT0977
Computer Science; Mathematics 13 pages

Research Report

Traffic Velocity Estimation from Vehicle Count Sequences

Takayuki Katsuki, Tetsuro Morimura, Masato Inoue

IBM Research - Tokyo
IBM Japan, Ltd.
19-21, Nihonbashi Hakozaki-cho
Chuo-ku, Tokyo 103-8510 Japan

Traffic Velocity Estimation from Vehicle Count Sequences

Takayuki Katsuki, Tetsuro Morimura and Masato Inoue

***Abstract*—Traffic velocity is a fundamental metric for inferring traffic conditions. This paper proposes a new velocity estimation approach from temporal sequences of vehicle count that does not require tracking any vehicles or using any labeled data. It is useful for measuring traffic velocities with low quality and inexpensive sensors such as web cameras in general use. We formalize the task as a density estimation problem by introducing a new model for temporal sequences of vehicle counts wherein the correlation between the sequences is directly related to the traffic velocity. We also derive a sampling-based algorithm for the density estimation. We show the effectiveness of our method on artificial and real-world datasets.**

***Index Terms*—Intelligent Transportation Systems, Bayes procedures, Unsupervised learning, Velocity measurement**

I. Introduction

EFFICIENT control of traffic and city planning for better traffic flow are keys to economic growth and improving our lives. Intelligent Transportation Systems (ITS) offer such solutions. Traffic monitoring is a fundamental part of ITS. In this paper, we tackle an important part of traffic monitoring, which is obtaining the traffic velocity. The traffic velocity is a basic measure of traffic congestion and is also useful information for identifying the travel time [1].

In order to measure the traffic velocity, two basic approaches have been proposed in the literature: (i) direct velocity sensing and (ii) indirect velocity estimation. The first approach attempts to sense velocity directly. Most of the existing studies on this approach track vehicles or track feature points in time-series of observations that correspond to the vehicles' movements by using techniques for determining identical vehicles, such as the Global Positioning System (GPS) [2], [3], video-based vehicle classification [4], [5], and video-based feature tracking including tracking of edges and lines characteristic of vehicles [6], [7] and optical flow [8]–[10], as well as targeted recognition of windshields [11] and headlights [11], [12]. Once the vehicles have been tracked in the temporal sequences, we can obtain their moving distances and can compute the traffic velocity by dividing the distance by the elapsed time. Moreover, there are variants of this approach based on matching between consecutive frames, e.g., tracking pixels in consecutive observation images using cross-correlation of an image feature [13], [14] or using matching of intensity profiles [15]. These approaches are robust against variations in the expected conditions of a roadway scene and are computationally relatively inexpensive. In most cases, such video-based techniques require camera calibrations, as they need to find the correct coordinate transform for obtaining the traffic velocity. An algorithm to detect scene changes was proposed in [16]; this algorithm can determine whether a camera has to be re-calibrated for video-based velocity estimation. On the other hand, some algorithms can use un-calibrated cameras for the velocity estimation. They use parameters derived from distributions of known vehicle lengths [17], [18], use an estimation of the camera's position relative to the roadway [19], or use a spatio-temporal map [20].

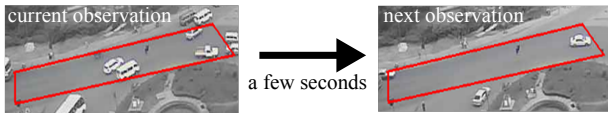


Fig. 1: At low sampling rates, only a few vehicles appear in consecutive observations [40].

The second approach (indirect velocity estimation) uses traffic information other than the traffic velocity itself for estimating the traffic velocity. The loop detector can obtain the traffic density, the traffic flow, and in some cases, the relationship between this information on neighboring detectors, and then we can estimate the velocity from them by using the models in [21]–[26]. If we use only the traffic density as the observation, we estimates the traffic velocity from the traffic density with a regression model. The regression model is determined by solving a regression problem between the traffic density and velocity on a labeled training dataset. Many regression models have been proposed, such as the linear model [27], [28], log-linear model [29], exponential model [30]–[32], bell-shaped curve model [33], and stochastic model [34], [35]. Extracting the traffic density from data is easier than sensing the traffic velocity directly. We can use many raw features for representing the traffic density, such as, for image features, the local variances of pixels [36] and the total area that may correspond to moving objects [37]–[39], and they do not require high-quality observations with high sampling rates. Here, the task of velocity regression from the traffic density is a one-shot estimation for a single observation and does not use sequences of consecutive observations. Accordingly, this approach works for any sampling rate.

However, the requirements of these approaches are often costly. This is particularly the case in cities in developing countries [37], [41]–[44]. The first approach requires that

T. Katsuki and T. Morimura are with the Analytics & Optimization Group, IBM Research – Tokyo, Tokyo 103-8510, Japan. E-mail: {kats,tetsuro}@jp.ibm.com.

Masato Inoue is with the Department of Electrical Engineering and Bioscience, Graduate School of Advanced Science and Engineering, Waseda University, Tokyo 3–4–1, Japan. E-mail: (see http://www.eb.waseda.ac.jp/m_inoue/).

many vehicles be identified in consecutive observations. Their feasibilities are sensitive to the quality and the sampling rate of the sensor. When using sensors with low sampling rates, such as web cameras, instead of expensive infrastructures such as the special-purpose close-view cameras used in the first approach, the number of vehicles that appear in consecutive observations is small, as shown in Fig. 1 [40]. The velocity regression approach from the traffic density enables us to use a variety of sensors whose sampling rates are far too slow for the direct sensing approach, including inexpensive and non-intrusive ones such as web cameras or mobile phones equipped with video and audio sensors [37], [41], [42], but we need to translate the obtained traffic density into a traffic velocity with a regression model and labeled training dataset. The dataset involves labeling a lot of training data, which is time-consuming and costly. We need a lightweight approach for traffic monitoring.

This paper proposes a new approach in which the traffic velocity is estimated only from observed temporal-sequences of the numbers of vehicles on a certain road area *without tracking any vehicles or using any labeled training data*. We use the fact that the some proportion of vehicles in two or more consecutive observations will be the same vehicles. The proportion will increase as the traffic velocity v decreases, and it directly represents the correlation between the numbers of vehicles in the consecutive observations. On the basis of the above fact, we first propose an observation model for observations conditioned on the traffic velocity v. Then, we estimate the traffic velocity through the density estimation of the model given the observations. This estimation task is an unsupervised one without using any labeled training data. Since our method does not need to track any vehicles, it can work with low quality and inexpensive sensors with low sampling rates, such as one observation every several seconds.

We formulate the velocity estimation task in Section II of this paper and propose the velocity estimation model in Section III. In Section IV, we describe a sampling-based approach for estimating the traffic velocity v. In Section V, we evaluate the proposed method using artificial and real-world datasets. We discuss the method in Section VI and conclude in Section VII.

II. Task and Goal

Let us define the task. We repeatedly observe the numbers of vehicles $\boldsymbol{x} \equiv [x_1, x_2, ..., x_N]^\top \in \mathbb{N}^N$ on a certain road area at time $\boldsymbol{t} \equiv [t_1, t_2, ..., t_N]^\top \in \mathbb{R}^N$ $(t_1 < t_2 < ... < t_N)$, as shown in Fig. 2, where $\mathbb{R}$ and $\mathbb{N}$ respectively denote the real number field and the set of natural numbers including zero and the superscript $\top$ denotes the transpose. All vectors in this paper are column vectors. Note that the time intervals of $\boldsymbol{t}$ are generally different from each other. For the length of the road area in which $\boldsymbol{x}$ is observed, $L > 0$ is known, and the road area has no intersections or branches. We refer to the road area as the observation area.

Our goal is to estimate the average traffic velocity $v \geq 0$ throughout the observations only from the available data $\boldsymbol{x}$ without tracking vehicles or using any labeled training data.

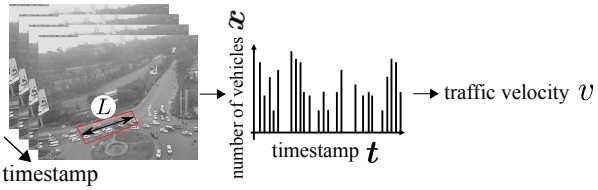


Fig. 2: Outline of the traffic velocity estimation problem.

In the probabilistic formulation, v is found through the posterior distribution $p(v|\boldsymbol{x})$, which represents the probability distribution of the traffic velocity v given the numbers of vehicles $\boldsymbol{x}$. From the Bayesian perspective, we find the optimal estimator for the traffic velocity, v^*, to be the posterior mean,

$$v^* \equiv \int v p(v|\boldsymbol{x}) \mathrm{d}v. \tag{1}$$

We will discuss the validity of the estimator v^* in Section VI-A.

III. Model

The estimator of the velocity v is found through the posterior distribution $p(v|\boldsymbol{x})$ from Eq. (1). The posterior $p(v|\boldsymbol{x})$ for v can be decomposed into an observation model for the number of vehicles $p(\boldsymbol{x}|v)$ that is conditioned on the average traffic velocity v and the prior model for the velocity $p(v)$:

$$p(v|\boldsymbol{x}) = \frac{p(\boldsymbol{x}|v)p(v)}{\int p(\boldsymbol{x}|v)p(v)\mathrm{d}v}. \tag{2}$$

We define the observation model $p(\boldsymbol{x}|v)$ conditioned on v and the prior model $p(v)$ in the following subsections.

A. Observation Model for Temporal-sequences of Numbers of Vehicles Conditioned on Traffic Velocity

We derive the observation model $p(\boldsymbol{x}|v)$ for $\boldsymbol{x}$ conditioned on the average traffic velocity v by considering the proportion of the number of vehicles that are in consecutive observations. If the time interval between t_n and t_{n+1} is not too large and the length of the observation area L is not too small, some of the vehicles will be the same in consecutive observations. In particular, the proportion of identical vehicles is supposed to become large when the average traffic velocity v is small. Conversely, it becomes small when the velocity is large. Also, the strength of the correlation between the consecutive observations in $\boldsymbol{x}$ will increase with this proportion. We use these relationships between the proportion of identical vehicles, the average traffic velocity v, and the correlation in $\boldsymbol{x}$ to derive the observation model.

First, we place two assumptions on the observed traffic during the period of N observations, $t_N - t_1$: (i) all the vehicles have a common constant velocity, v, toward the observation area, and (ii) the positions of the vehicles at time zero are independent and identically distributed. These assumptions might be strong but can be applicable for many cases when the period $t_N - t_1$ is not so long. We will examine and discuss these assumptions in Sections V and VI.

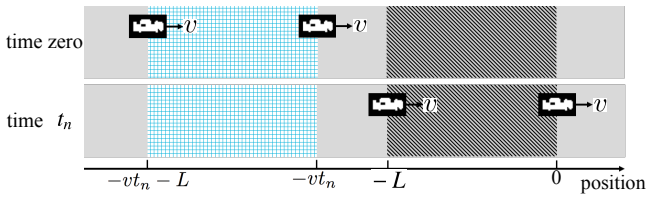


Fig. 3: Positions of observed vehicles at time zero and t_n.

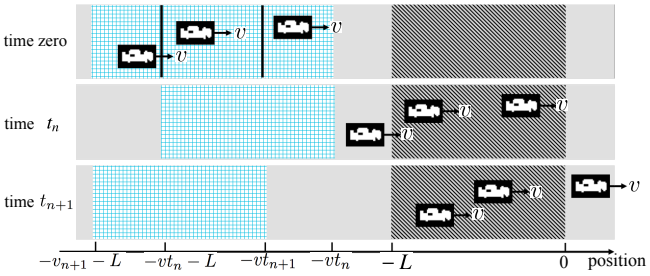


Fig. 4: Positions of observed vehicles at time zero, t_n and t_{n+1}. In this case, there are overlapping area between the consecutive n-th and n+1-th observations in $[-vt_n - L, -vt_{n+1})$.

Here, we define the positions of the vehicles as the distance from the front of the observation area, which is regarded as the zero position (see Fig. 3). The observation area is defined as $[-L, 0)$, as shown in the hatched area in Fig. 3, where $[\bullet]$ denotes a closed interval and $(\bullet)$ denotes an open interval. If a vehicle is located at the position $-y$ at time zero, the assumption (i) indicates that the vehicle is observed during time $[(y-L)/v, y/v)$. Accordingly, the vehicles that will be observed at time t_n in the n-th observation should be located in the area $[-vt_n - L, -vt_n)$ at time zero, as shown in the check-pattern area in Fig. 3.

Since the area in which the vehicles in the $n+1$-th observation can exist at time zero is $[-vt_{n+1} - L, -vt_{n+1})$, when $-vt_n - L < -vt_{n+1}$, the areas for the consecutive n-th and n+1-th observations partially overlap each other. This overlapping area can be defined as

$$[\max(-vt_{n+1}, -vt_n - L), \min(-vt_{n+1}, -vt_n - L)), \quad (3)$$

where vehicles that are in this area at time zero are observed at both times in the n-th and $n+1$-th observations and *are not observed individually in each of the n-th and $n+1$-th observations*. Note that if there is no overlapping area between the consecutive n-th and n+1-th observations, Eq. (3) becomes an empty set. We denote the overlapping area as $[-vt_n - L, -vt_{n+1})$ in Fig. 4. From Eq. (3), the length of the overlapping area can be written by v as

$$\begin{aligned} &|[\max(-vt_{n+1}, -vt_n - L), \min(-vt_{n+1}, -vt_n - L))| \\ &\quad = \max\Big(0, \min(-vt_{n+1}, -vt_n - L) \\ &\qquad\qquad - \max(-vt_{n+1}, -vt_n - L)\Big), \end{aligned} \quad (4)$$

where we make the length to be zero for when Eq. (3) is the empty set.

While the above overlapping area and its length is defined between the consecutive n-th and $n+1$-th observations, it

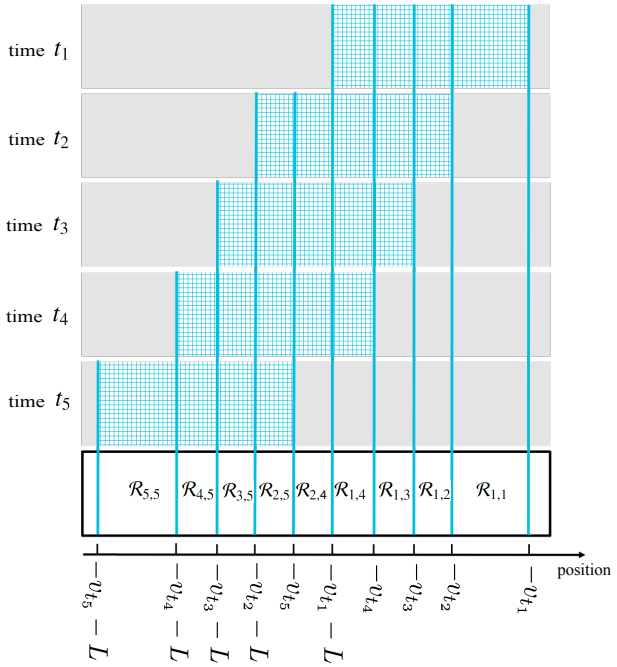


Fig. 5: Example of $\mathcal{R}_{j,k}$s and $L_{j,k}(v)$s in the case of $N = 5$. For example, $L_{3,5}(v) = -vt_2 - L - (-vt_3 - L)$. The check-pattern areas represent the areas at which the vehicles that will be observed at time t_n in n-th observation should be located at time zero.

can be naturally generalized to the overlapping area between the j-th to k-th consecutive observations in $\boldsymbol{x}$, $\mathcal{R}_{j,k}$ $(j, k \in 1, 2, ..., N,$ and $j \leq k)$. The vehicles, which are in this area at time zero, are in all of the j-th to k-th observations and are not observed in the other observations in $\boldsymbol{x}$. This area $\mathcal{R}_{j,k}$ can be defined as

$$\mathcal{R}_{j,k} \equiv \bigcap_{n=1}^{N} \begin{cases} [-vt_n - L, -vt_n) & (\text{if } j \leq n \leq k) \\ \overline{[-vt_n - L, -vt_n)} & (\text{elsewhere}) \end{cases}, \quad (5)$$

where $\bigcap_{n=1}^{N}$ denotes intersection over $n = 1$ to $n = N$ and $\overline{\bullet}$ denotes exclusion of $\bullet$. Figure 5 shows an example of $\mathcal{R}_{j,k}$s in the case of $N = 5$. In Eq. (5), since $[-vt_n - L, -vt_n)$ represents the check-pattern area in Fig. 5 at each time, $\mathcal{R}_{j,k}$ can be computed as the intersection over the corresponding check-pattern areas:

$$\mathcal{R}_{j,k} = [\max(-vt_{k+1}, -vt_j - L), \min(-vt_k, -vt_{j-1} - L)). \quad (6)$$

where $t_0 \equiv -\infty$, $t_{N+1} \equiv \infty$. Note that, from the definition in Eqs. (5) and (6), some of the intervals $\mathcal{R}_{j,k}$ may also be the empty set, such as $\mathcal{R}_{1,5}$, $\mathcal{R}_{2,2}$, $\mathcal{R}_{2,3}$, $\mathcal{R}_{3,3}$, $\mathcal{R}_{3,4}$, and $\mathcal{R}_{4,4}$ in Fig. 5. The area $\mathcal{R}_{n,n}$ represents the area in which the vehicles observed only in the n-th observation exist at time zero. When $N = 2$, Eq. (3) and Eq. (6) are identical. Additionally, all the ranges $\mathcal{R}_{j,k}$ are mutually exclusive. The length of the

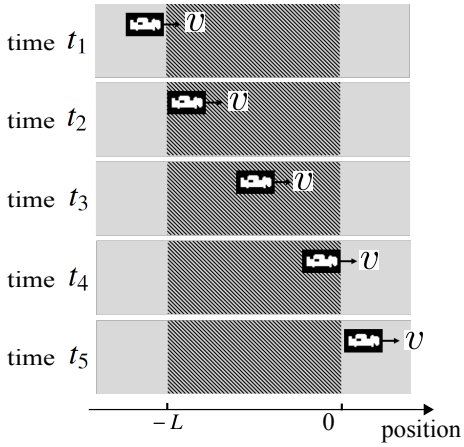


Fig. 6: Example in which $c_{2,4} = 1$ and the other components in $\boldsymbol{c}$ are 0 in the case of $N = 5$.

overlapping area can be written by v as

$$\begin{aligned} L_{j,k}(v) &\equiv |\mathcal{R}_{j,k}| \\ &= \max\Big(0, \min(-vt_k, -vt_{j-1} - L) \\ &\quad - \max(-vt_{k+1}, -vt_j - L)\Big), \end{aligned} \tag{7}$$

where we also make the length $L_{j,k}(v)$ to be zero for when Eq. (6) is the empty set.

Then, we introduce a random variable $c_{j,k}$ $(j, k \in 1, 2, ..., N, \text{ and } j \leq k)$, which denotes the number of vehicles in the mutually exclusive area $\mathcal{R}_{j,k}$ at time zero and is a decomposition of $\boldsymbol{x}$. Figure 6 shows an example where $c_{2,4} = 1$ and the other components in $\boldsymbol{c}$ are 0 in the case of $N = 5$. The observation variables $\boldsymbol{x}$ can be determined uniquely in terms of $\boldsymbol{c} \equiv [c_{1,1}, ..., c_{1,N}, c_{2,2}, ..., c_{2,N}, ..., c_{N,N}]^\top$:

$$x_n = \sum_{1 \leq j \leq n \leq k \leq N} c_{j,k}. \tag{8}$$

For simplicity, we introduce an $N \times \frac{1}{2}N(N+1)$ matrix $\boldsymbol{D}$ that corresponds to the above summation and satisfies

$$\boldsymbol{x} = \boldsymbol{D}\boldsymbol{c}. \tag{9}$$

Since the position of a vehicle at time zero is random from assumption (ii), the probability, which is that the vehicle appears in the area $\mathcal{R}_{j,k}$ at time zero when the velocity is v, is given by the length of the area $L_{j,k}(v)$ divided by the total length of the roads L_{total} in the possible road area, *i.e.*, $L_{j,k}(v)/L_{\text{total}}$. Because all the areas $\mathcal{R}_{j,k}$ are mutually exclusive, the event that this vehicle does not appear among N observations is a complementary event, and its probability is given by $1 - \sum_{1 \leq j \leq k \leq N} L_{j,k}(v)/L_{\text{total}}$. Thus, the random vector $\boldsymbol{c}$ obeys a multinomial distribution using these probabilities and the total number of vehicles M_{total} in the possible road area:

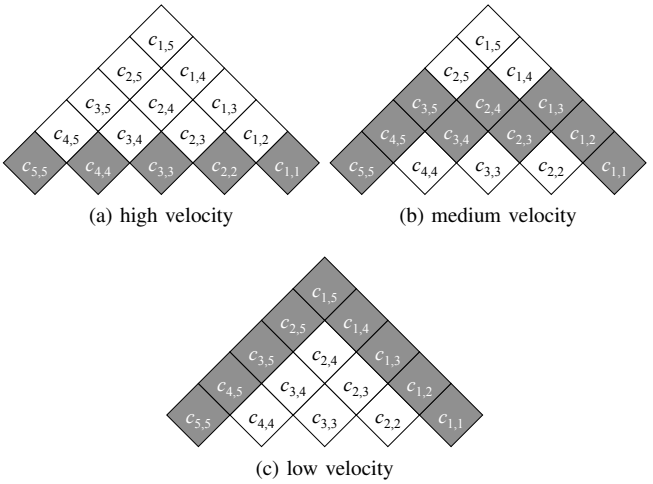


Fig. 7: Examples of $p(\boldsymbol{c}|\, v, M_{\text{total}})$ where v is low, medium, and high. The gray regions have non-zero probability.

$$\begin{aligned} &p(\boldsymbol{c}|\, v, M_{\text{total}}) \\ &= \frac{M_{\text{total}}}{\left(M_{\text{total}} - \sum_{1 \leq j \leq k \leq N} c_{j,k}\right) \prod_{1 \leq j \leq k \leq N} c_{j,k}} \\ &\quad \times \left(1 - \sum_{1 \leq j \leq k \leq N} \frac{L_{j,k}(v)}{L_{\text{total}}}\right)^{M_{\text{total}} - \sum_{1 \leq j \leq k \leq N} c_{j,k}} \\ &\quad \times \prod_{1 \leq j \leq k \leq N} \left(\frac{L_{j,k}(v)}{L_{\text{total}}}\right)^{c_{j,k}}. \end{aligned} \tag{10}$$

Figure 7 shows examples of $p(\boldsymbol{c}|\, v, M_{\text{total}})$, which illustrate typical cases of v. The $c_{j,k}$s depicted as the gray regions have a non-zero probability for each of the high, medium, and low velocity cases. We can see that this distribution $p(\boldsymbol{c}|\, v, M_{\text{total}})$ for $\boldsymbol{c}$ dramatically changes with the velocity.

Since L_{total} and M_{total} are usually huge, we consider the large limit of them while keeping their ratio constant, *i.e.*, $M_{\text{total}}/L_{\text{total}} = M/L$, where M is a newly introduced parameter. Both $M_{\text{total}}/L_{\text{total}}$ and M/L mean the vehicle density per unit road length. After taking this limit, $p(\boldsymbol{c})$ becomes

$$\begin{aligned} p(\boldsymbol{c}|\, v, M) &= \prod_{1 \leq j \leq k \leq N} \text{Poisson}\left(c_{j,k}; q_{j,k}\right), \\ \text{where} \quad q_{j,k} &\equiv \frac{L_{j,k}(v)}{L} M. \end{aligned} \tag{11}$$

Note that by taking the limit, the $c_{j,k}$s become independent from one another. The v and M are unknown parameters to be estimated. Within our Bayesian framework, we introduce their prior distributions in the next subsection.

Finally, through marginalization over $\boldsymbol{c}$ using Eqs. (9) and (11), the observation model for $\boldsymbol{x}$ can be written as

$$p(\boldsymbol{x}|\, v, M) \equiv \sum_{\boldsymbol{c}} \delta_{\boldsymbol{x}, \boldsymbol{D}\boldsymbol{c}} p(\boldsymbol{c}|\, v, M), \tag{12}$$

where δ denotes the Kronecker delta function. However, in this complicated case, the discrete marginalization in Eq. (12)

is computationally infeasible. As an alternative, $p(\boldsymbol{x}|v,M)$ is approximated as a Gaussian distribution in Section IV.

B. Prior Distributions

We introduce prior distributions for v and M:

$$p(v,M) \equiv \mathrm{InverseGamma}(v;a_v,b_v) \times \mathrm{InverseGamma}(M;a_M,b_M), \tag{13}$$

where InverseGamma represents the inverse gamma distribution (see the Appendix for the explicit definition) and the parameters a_v, b_v, a_M, and b_M are treated as input parameters given as part of the model. See the Hyperparameter subsection in the Experimental Results section for these parameters we actually used. The reason we chose inverse gamma distributions for v and M is that they are defined as positive variables and play a role similar to that of the variance parameter in a Gaussian distribution, where the inverse gamma distribution is widely used as the conjugate prior distribution for the variance parameter.

C. Joint Distribution

From Eqs. (12) and (13), we explicitly construct the joint distribution of all random variables:

$$p(\boldsymbol{x},v,M) \equiv p(\boldsymbol{x}|v,M)p(v,M). \tag{14}$$

All marginal and conditional distributions including the posterior $p(v|\boldsymbol{x})$ can be derived in terms of this joint distribution.

IV. Method

Although the estimator, the posterior mean, in Eq. (1) is derived from the joint distribution in Eq. (14), an exact analytical solution is computationally infeasible. We will thus use an approximate inference method to compute the posterior mean.

A. Approximate Marginalization in Observation Model

Since the discrete marginalization in Eq. (12) is computationally infeasible, as stated in Section III, $p(\boldsymbol{x}|v,M)$ is approximated as a Gaussian distribution with the same mean and covariance ignoring cumulants higher than second-order:

$$p(\boldsymbol{x}|v,M) \simeq \mathrm{Gauss}(\boldsymbol{x};\boldsymbol{\mu},\boldsymbol{\Sigma}), \quad \text{where} \tag{15}$$

$$\boldsymbol{\mu} \equiv \mathrm{E}_{p(\boldsymbol{x}|v,M)}[\boldsymbol{x}] \quad \text{and} \tag{16}$$

$$\boldsymbol{\Sigma} \equiv \mathrm{Var}_{p(\boldsymbol{x}|v,M)}[\boldsymbol{x}]. \tag{17}$$

$\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ can be exactly calculated as

$$\boldsymbol{\mu} = \mathrm{E}_{p(\boldsymbol{c}|v,M)}[\boldsymbol{D}\boldsymbol{c}] = \boldsymbol{D}\mathrm{E}_{p(\boldsymbol{c}|v,M)}[\boldsymbol{c}] = M\boldsymbol{i}, \tag{18}$$

$$\boldsymbol{\Sigma} = \mathrm{Var}_{p(\boldsymbol{c}|v,M)}[\boldsymbol{D}\boldsymbol{c}] = \boldsymbol{D}\mathrm{Var}_{p(\boldsymbol{c}|v,M)}[\boldsymbol{c}]\boldsymbol{D}^\top = \boldsymbol{D}(\mathrm{diag}\ \boldsymbol{q})\boldsymbol{D}^\top, \tag{19}$$

where $\boldsymbol{i}$ represents a vector of all ones, $\boldsymbol{q} \equiv [q_{1,1},...,q_{1,N},q_{2,2},...,q_{2,N},...,q_{N,N}]^\top$ and $(\mathrm{diag}\ \boldsymbol{q})$ denotes a diagonal matrix whose diagonal elements are $\boldsymbol{q}$. We will examine this approximation in Section V-B. Note that $\boldsymbol{D}$ is defined as the operation $\sum_{1\le j\le n\le k\le N}$ for each element of an $\frac{1}{2}N(N+1)$-dimensional vector, as shown in Eq. (9). In Eq. (18), each element of the vector $\boldsymbol{D}\mathrm{E}_{p(\boldsymbol{c}|v,M)}[\boldsymbol{c}]$ represents the sum of all expectations of the corresponding $c_{j,k}$, which means $\sum_{1\le j\le n\le k\le N}\mathrm{E}_{p(c_{j,k}|v,M)}[c_{j,k}]$ and is always M. Thus, the mean value of the model has no information on the velocity, but the covariance matrix does. We can hence obtain the likelihood of the traffic velocity from the covariance matrix, which represents the correlation between elements of $\boldsymbol{x}$.

Algorithm 1 Sampling Procedure for Velocity Estimation

1: Initialize the values of v and M with their prior distributions
2: **repeat**
3: $\quad\xi_v \leftarrow \mathcal{U}(\xi_v|0,p(\boldsymbol{x},v^{(\tau-1)},M^{(\tau-1)}))$
4: $\quad$Sample $v^{(\tau)}$ uniformly from part of the slice $S_v = \{v^{(\tau)} : \xi_v < p(\boldsymbol{x},v^{(\tau)},M^{(\tau-1)}))\}$
5: $\quad\xi_M \leftarrow \mathcal{U}(\xi_M|0,p(\boldsymbol{x},v^{(\tau)},M^{(\tau-1)}))$
6: $\quad$Sample $M^{(\tau)}$ uniformly from part of the slice $S_M = \{M^{(\tau)} : \xi_M < p(\boldsymbol{x},v^{(\tau)},M^{(\tau)}))\}$
7: **until** a stopping condition is met.
8: **return** $v^{(1)},v^{(2)},...,v^{(T)}$

B. Approximate Posterior Mean Inference for Estimating the Traffic Velocity

Our goal is to obtain the posterior mean v^* from the above joint distribution in Eq. (14) with the model of Eqs. (15) and (13) given the observations $\boldsymbol{x}$. However, although we can compute the joint distribution $p(\boldsymbol{x},v,M)$, we cannot analytically compute the posterior $p(v|\boldsymbol{x})$.

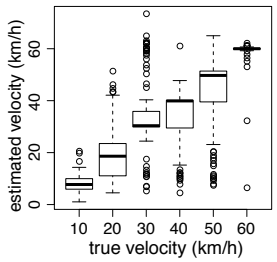
Instead, we derive a sampling-based approximation of the posterior mean v^* by using the Markov chain Monte Carlo (MCMC) method. Given the joint distribution $p(\boldsymbol{x},v,M)$ and observations $\boldsymbol{x}$, we can take T samples for v, $\{v^{(\tau)}\}_{\tau=1}^T$, from the posterior $p(v|\boldsymbol{x})$ with the MCMC method, but without explicitly computing the posterior $p(v|\boldsymbol{x})$. Then, we use the empirical mean of $\{v^{(\tau)}\}_{\tau=1}^T$ as an approximation of v^*:

$$v^* = \int vp(v|\boldsymbol{x})\mathrm{d}v \simeq \frac{1}{T}\sum_{\tau=1}^{T} v^{(\tau)}. \tag{20}$$

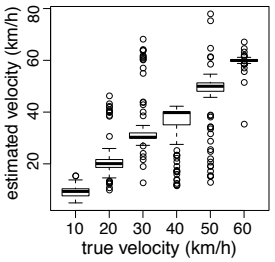
C. Slice Sampling for Traffic Velocity Estimation

For sampling $\{v^{(\tau)}\}_{\tau=1}^T$ in Eq. (20), we use slice sampling [45]. Since we have two different random variables v and M, we repeatedly sample v and M in turn in the same manner as Gibbs sampling [45]–[47] and obtain samples only for v from $p(v|\boldsymbol{x})$ by ignoring samples for M. For the stopping condition, we simply use the number of iterations. See the Hyperparameter subsection in the Experimental Results section for the iteration number we actually used.

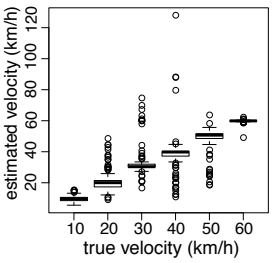
Algorithm 1 shows the sampling procedure for the velocity estimation task. Here, ξ_v and ξ_M are auxiliary variables for v and M, respectively, which are required in the sampling scheme of slice sampling, $\bullet \leftarrow \circ$ denotes that a sample from a distribution $\circ$ is substituted into $\bullet$, and $\mathcal{U}$ denotes a uniform distribution (see the Appendix for the explicit definition). For



(a) average number of vehicles $M_{\mathrm{art}} = 1$

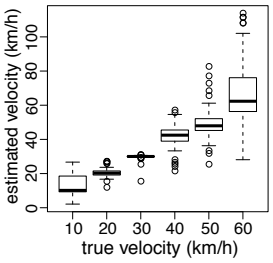


(b) average number of vehicles $M_{\mathrm{art}} = 10$

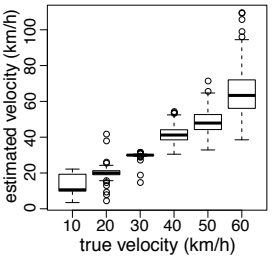


(c) average number of vehicles $M_{\mathrm{art}} = 100$

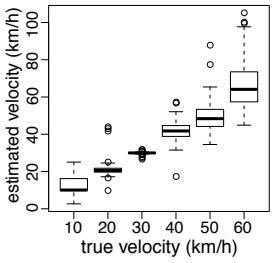
Fig. 8: True velocity and velocity estimated using the proposed method for the time interval $\Delta t = 1$ (second) in the artificial instance. Note that we have used the Tukey boxplot [48].



(a) average number of vehicles $M_{\mathrm{art}} = 1$



(b) average number of vehicles $M_{\mathrm{art}} = 10$



(c) average number of vehicles $M_{\mathrm{art}} = 100$

Fig. 9: True velocity and estimated velocity using the proposed method for the time interval $\Delta t = 4$ (second) in the artificial instance. Note that we have used the Tukey boxplot [48].

efficiency, we use the "stepping out procedure" in Steps 3 and 5 and use the "shrinkage procedure" in Steps 4 and 6, as described in [45].

V. Experimental Results

We examined the proposed approach in numerical experiments. First, we generated artificial datasets to study the performance of the approximate inference method (Section V-B) and to examine the robustness of the proposed method in a situation where each vehicle has a different velocity (Section V-C). Next, to see if it could deal with actual traffic, we applied it to the temporal-sequences of the numbers of vehicles extracted from real-world web-camera images (Section V-D) and to publicly available traffic datasets (Section V-E).

A. Hyperparameters for the Model

For the Bayesian inference, we set the hyperparameter values in Eq. (13) to be as non-informative as possible and to have a quite flat distribution:

$$a_v = a_M = 10^{-4},\ b_v = v_{\mathrm{legal}} \times 10^{-4},\ b_M = \mu_{\boldsymbol{x}} \times 10^{-4}, \tag{21}$$

$$\text{where} \quad \mu_{\boldsymbol{x}} \equiv \frac{1}{N}\sum_{n=1}^{N} x_n.$$

For the prior means of v and M, we respectively used the legal speed limit v_{legal} on each road and the naively computed a sample mean of $\boldsymbol{x}$. The number of effective prior observations of the inverse gamma distribution within the Bayesian framework is equal to twice the value of parameter a. These settings were considered sufficiently non-informative. For fairness, we used this hyperparameter setting for all of the following instances. Also, the number of iterations of the slice sampling was 1000. A preliminary analysis indicated that using more iterations, such as 100000, did not improve the accuracy much.

B. Experiment on Artificial Dataset

In preparing the artificial validation dataset, we randomly generated $\boldsymbol{c}$ at constant time intervals from the model in Eq. (11) and computed the observations $\boldsymbol{x}$ as the sum of the corresponding $\boldsymbol{c}$. Using the model in Eq. (11) for generating the dataset and using the approximate model in Eq. (15) for estimating the velocity, we studied the validity of the Gaussian approximation for the approximate model in Eq. (15) and the performance of the approximate posterior mean inference.

Using Eq. (11), we generated $\boldsymbol{x}$ with the following settings: time intervals of $\boldsymbol{t}$, $\Delta t = \{1, 4\}$ (second), length of the observation area $L = 100$ (m), average number of vehicles on the road $M_{\mathrm{art}} = \{1, 10, 100\}$, traffic velocity $v = \{10, 20, 30, 40, 50, 60\}$ (km/h), and number of input observations $N = 50$. We set the legal speed limit to the one in Japan, $v_{\mathrm{legal}} = 60$ (km/h). For each of these settings, we repeatedly evaluated the proposed method in 100 experiments, where we used a different random seed in each experiment. Thus, the total number of experiments was $2 \times 3 \times 6 \times 100 = 3600$.

Figures 8 and 9 compare the true velocity with the estimated velocity using the proposed method. We can see that the overall performance of our method is good. Our method performed consistently for different values of M, even $M = 1$, which

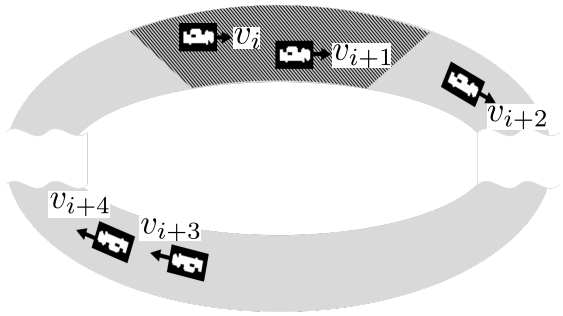


Fig. 10: Traffic simulation for validation.

is a difficult setting for this kind of Gaussian approximation (Eq. (15)). The results show that the Gaussian approximation and approximate MCMC inference method worked well. They are also non-trivial because the density-velocity regression approach cannot work in a scenario where the average number of vehicles is independent from the traffic velocity.

C. Experiment on Simulated Traffic

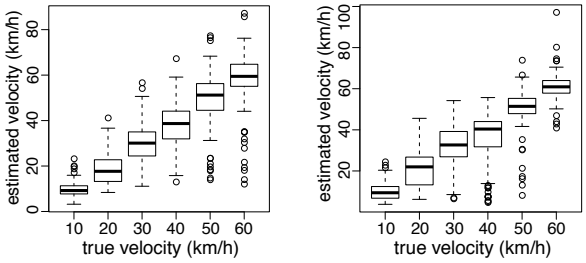
Next, we evaluated the proposed method on simulated traffic data, where the observations $\boldsymbol{x}$ were obtained from our simple traffic simulation. We examined its robustness in a situation where each vehicle had a different velocity.

We simulated the traffic as follows (see Fig. 10). First, we distributed M_{sim} vehicles at random positions on a virtual circuit whose total length was L_{sim}. The parameters M_{sim} and L_{sim} were respectively equal to one-thousandth of the total number of vehicles in Japan and the total road length in Japan. The vehicles moved at different velocities that were generated from a uniform distribution $\mathcal{U}(x|\tilde{v}-10, \tilde{v}+10)$, where the average of the true velocities was set as $\tilde{v} = \{10, 20, 30, 40, 50, 60\}$ (km/h). This means the vehicles had their own velocities. We repeatedly obtained the numbers of vehicles $\boldsymbol{x}$ from a certain road in the virtual circuit at constant time intervals. We experimented with time intervals of $\boldsymbol{t}$ $\Delta t = \{1, 2, 4\}$ (second), a length of the observation area $L = 100$ (m), and $N = 50$ input observations. We set the legal speed limit in this experiment to the legal speed limit in Japan, *i.e.*, $v_{\mathrm{legal}} = 60$ (km/h). For each of these settings, we evaluated the proposed method in 100 experiments, where we used a different random seed in each experiment. Thus, the total number of experiments was $3 \times 6 \times 100 = 1800$.

Figure 11 compares the true average velocity with the estimated velocity using the proposed method. Even when each vehicle had a different velocity, we can see that the proposed method estimated the average traffic velocity as well as it did on the artificial dataset in Section V-B.

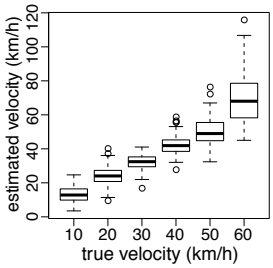
D. Experiment using Real-world Web-camera Images

We demonstrated the utility of our approach by using it to estimate the traffic velocity from web-camera images captured in Tokyo, Japan for the city traffic monitoring scenario [37], [41], [49] (see Fig. 12). The specific location is at 35.651054N, 139.799986E. The image size was 640×480 pixels. The legal



(a) time interval $\Delta t = 1$ (second)

(b) time interval $\Delta t = 2$ (second)



(c) time interval $\Delta t = 4$ (second)

Fig. 11: True average velocity and velocity estimated using the proposed method in the simulation instance. Note that we have used the Tukey boxplot [48].

Fig. 12: Real-world web-camera image used in our experiment.

speed limit on this road was $v_{\mathrm{legal}} = 60$ (km/h). The frame rate of the web camera was about one image per second. Since the sampling rate of the web camera was low, this was a good application for our approach. The dataset contained images captured for 17 minutes, and the images in the set totaled about 1000.

For $\boldsymbol{x}$, we simply used the temporal-sequences of the numbers of vehicles in the images, which were extracted from the images by using the method described in [37], [41], [49], [50]. This method works for any frame rate, even for still images, is robust even when image quality is poor and is almost calibration-free because it does not recognize individual vehicles, but rather estimates their number from the image features. The road length L can easily be obtained because we can estimate the size of the vehicles and the length of the road in the images by using the methods described in [37], [41], [49], [50]. In particular, L was computed by referring

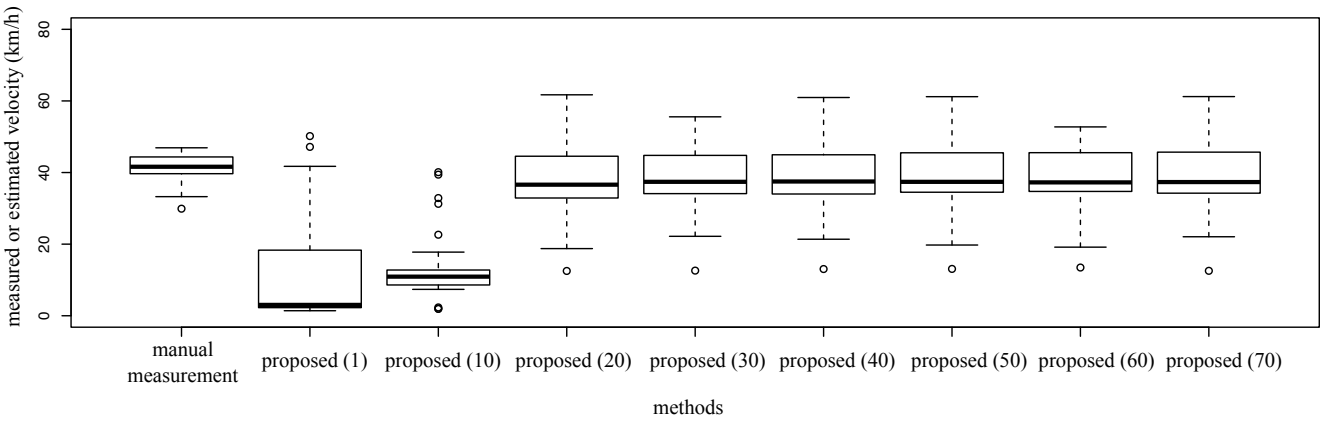


Fig. 13: Manually measured velocity and velocity estimated using the proposed method with different prior means. In this figure, proposed (1) means the proposed method with the prior mean 1 (km/h) for the traffic velocity. Note that we have used the Tukey boxplot [48].

to the typical sizes of vehicles in the real world. We used the timestamps attached to the images as $\boldsymbol{t}$ and input $N = 60$ consecutive images (corresponding to about one minute) for each estimation. We estimated the traffic velocities about 17 times. To create the validation test dataset, we manually measured traffic velocities each minute by using a radar speed gun at the roadside. The average traffic velocity was 41 (km/h).

Since we have only one traffic situation in this real-world traffic scenario, in which the average traffic velocity was 41 (km/h), we checked to ensure that the hyperparameter setting for the traffic velocity does not accidentally become the best for estimating this velocity. We tested the following hyperparameter settings: $b_v = \{1, 10, 20, 30, 40, 50, 60, 70\} \times 10^{-4}$ and $a_v = 10^{-4}$ for the traffic velocity v. This means that the prior mean b_v/a_v for v takes over $\{1, 10, 20, 30, 40, 50, 60, 70\}$ (km/h). In our Bayesian framework, the traffic velocity estimation is generally difficult when this prior mean is significantly different from the true velocity.

Figure 13 compares our estimation results of each prior mean with the manually measured velocities in the 17 minutes of observations. We can see that our method can estimate a reasonable velocity over the different prior mean settings of v, except for the settings in which the prior means are 1 and 10 (km/h). In these cases (in which the prior means are 1 and 10 (km/h)), the true velocity is more than four times larger than the prior mean. Since it is not difficult to set the prior mean in a range that is the true velocity plus or minus 30 (km/h), this result shows that the dependence of the hyperparameters is small enough.

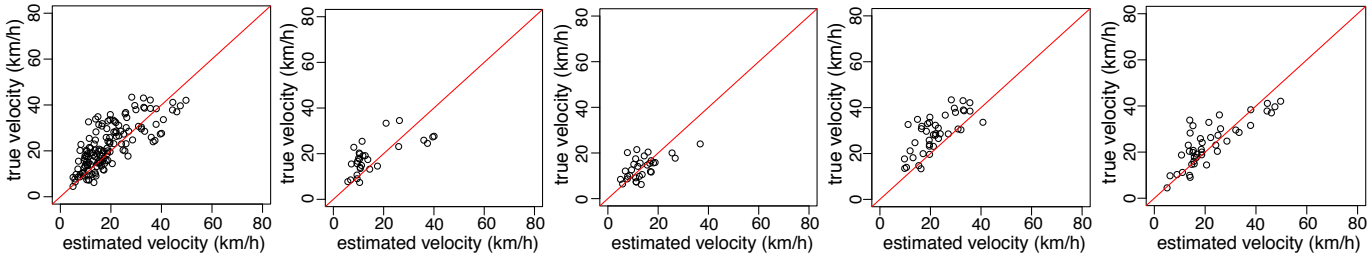
E. Experiment using NGSIM Data

Finally, we evaluated the proposed method using the publicly available Next Generation Simulation (NGSIM) datasets collected by the United States Department of Transportation Federal Highway Administration [51]. The NGSIM datasets consist of real-world vehicle trajectory data collected using digital video cameras at several locations in the United States. *The proposed method only used the number of vehicles every second in the NGSIM datasets; it did not use the original speed information.* The original speed information was used only for evaluating the estimation accuracy of the proposed method.

We used four vehicle trajectory datasets: Peachtree, Lankershim, US-101, and I-80. Each of them was collected at different locations. The Peachtree and Lankershim datasets were collected on local roads, which means that vehicles had relatively lower velocities and the legal speed limit was $v_{\mathrm{legal}} = 56.3269$ (km/h). The US-101 and I-80 datasets were collected on freeways, which means that vehicles had relatively higher velocities and the legal speed limit was $v_{\mathrm{legal}} = 88.5137$ (km/h).

We used the initial 100 (m) areas without intersections or branches at each location as the observation area, which means the length of the observation area was $L = 100$ (m). We used the numbers of vehicles in the trajectories on the observation area at each time and *thinned them to one observation per second for making the observations $\boldsymbol{x}$ in our problem setting*. We input $N = 60$ (corresponding to about one minute) consecutive observations for each estimation. Since the Peachtree and Lankershim datasets contain vehicle trajectories for about 30 minutes, the number of estimations for each of them was about 30. Similarly, since the US-101 and I-80 datasets contain vehicle trajectories for 45 minutes, the number of estimations for each of them was about 45. Thus, the total number of estimations for these datasets was about 150. We used the timestamps attached to the trajectories as $\boldsymbol{t}$.

We compared the estimated velocities with the true ones for each location, as shown in Fig. 14. We can see that the proposed method can estimate a reasonable velocity at every location with the different true velocity. For the US-101 dataset, the estimation results seem to have bias. We will discuss the bias in Section VI-B. Additionally, Fig. 15



(a) All (b) Peachtree (c) Lankershim (d) US-101 (e) I-80

Fig. 14: Relationship between true NGSIM velocity and velocity estimated by the proposed method. The estimate is good if the dot is close to the identity line (red line).

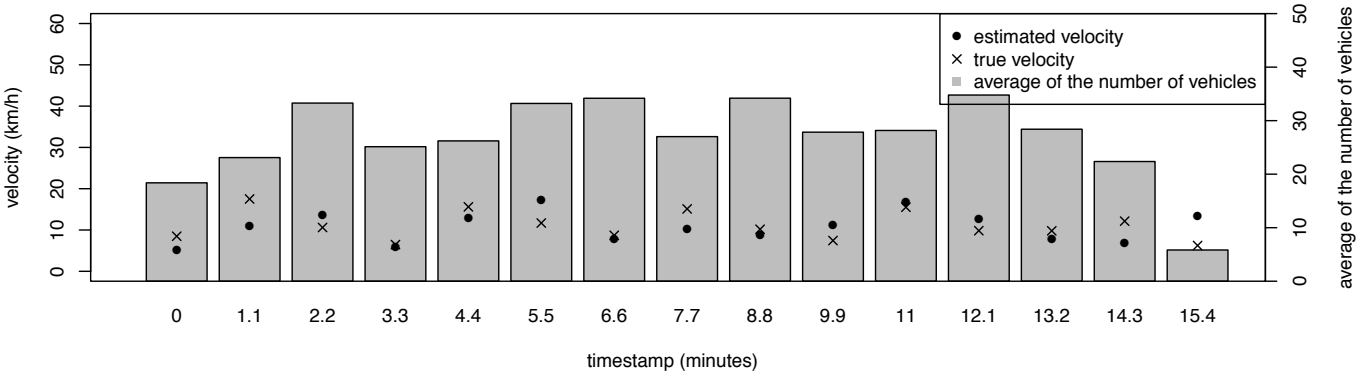


Fig. 15: Time-series of velocities estimated by the proposed method and the corresponding true velocities. The average number of vehicles for the corresponding timestamp is indicated by the bar chart.

shows typical time-series of the true and estimated velocities, together with average number of vehicles. We can see that the proposed method can estimate a reasonable velocity regardless of the average number of vehicles. In addition, Fig. 16 shows the relationship between the estimation error and standard deviation of vehicle velocity. Although we assume that all vehicles have a common velocity, the standard deviation of the velocity, which reflects variation of the velocities over vehicles and time, including situations such as overtaking and the existence of multiple lanes, did not affect the quality of the estimation of the real-world NGSIM data, as well as, the simulated traffic of Section V-C.

VI. Discussion

We discuss the proposed approach in this section. First, we give the validity of our velocity estimation approach (Section VI-A). Next, we show the validation of the limitations of the approach (Section VI-B). Finally, we discuss other applications of the approach (Section VI-C).

A. Validity of Velocity Estimation Method from Temporal-sequences of Vehicle Counts

From our experiments shown in Section V, without tracking any vehicles or using any labeled training data, we confirmed that the proposed observation model for the temporal-sequences of the numbers of vehicles $\boldsymbol{x}$ can properly represent the likelihood of v given $\boldsymbol{x}$. Also, we showed that our approximate estimation method performs consistently and stably well on the simulation and real-world datasets.

From a theoretical perspective, we give the validity of the estimator of the traffic velocity, v^*, which is the posterior mean as defined in Eq. (1), even when there is no labeled training data. First, we consider the objective error function between the estimator v^* and the corresponding true velocity v. Since the velocity is a positive real number, the use of an all-or-none type error, such as the Dirac delta or Kronecker delta function, is nonsensical, whereas the squared L2-norm error is a conventional way of doing so. The squared L2-norm error is defined as

$$\|v^* - v\|^2. \tag{22}$$

Since only the observations, $\boldsymbol{x}$, are available for the estimator, we explicitly express the estimator as a function, $v^*(\boldsymbol{x})$. The objective function to be minimized regarding the estimator is defined as

$$F(v^*(\boldsymbol{x})) \equiv \int \|v^*(\boldsymbol{x}) - v\|^2 p(\boldsymbol{x}, v) \mathrm{d}\boldsymbol{x}\mathrm{d}v. \tag{23}$$

This is because we prefer good estimator performance on average over various observations, $\boldsymbol{x}$, and the corresponding

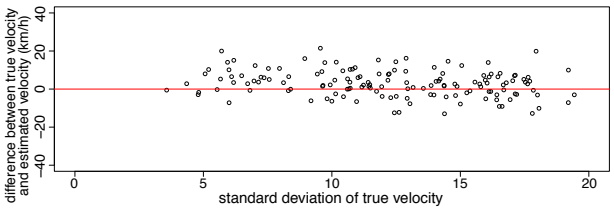


Fig. 16: Relationship between the estimation error and standard deviation of velocity. The estimate is good if the dot is close to the line, which means the difference is zero (red line).

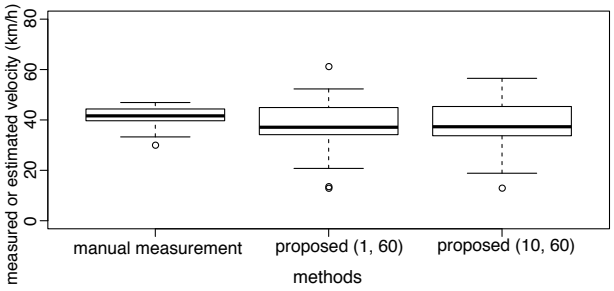


Fig. 17: Manually measured velocity and velocity estimated using the proposed method with different prior means. The initial sample value was always the legal speed limit, 60 (km/h). In this figure, proposed (1, 60) means the proposed method with the prior mean 1 (km/h) and the initial sample value 60 (km/h) for the traffic velocity. Note that we have used the Tukey boxplot [48].

true traffic velocities, v. Here, we assume that the occurrence rates of v and $\boldsymbol{x}$ exactly coincide with those of the model introduced in Section III. Accordingly, we can explicitly compute the optimal estimator of the velocity for the objective function as the posterior mean,

$$v^*(\boldsymbol{x}) = \underset{v^*(\boldsymbol{x})}{\operatorname{argmin}} F(v^*(\boldsymbol{x})) = \int v p(v|\boldsymbol{x}) \mathrm{d}v. \qquad (24)$$

This is a well-known result in the Bayesian framework; that is, the posterior mean coincides with the minimum mean square error estimator.

We used slice sampling in the implementation of our method. Gibbs sampling, which is one of the most common MCMC methods, is not applicable to our model because we cannot analytically compute the conditional distributions required in its sampling procedure. Also, an efficient implementation of the Metropolis algorithm, which is also a common MCMC methods, is difficult because it is almost impossible to prepare appropriate proposal distributions for both of the random variables v and M. We prefer slice sampling because it does not require such analytical modeling or sensitive setting of the proposal distributions [45], [46].

With regard to the initial sample value for the traffic velocity in our slice sampling procedure, we used the prior mean of the traffic velocity. Here, we recommend that the initial sample value for the traffic velocity be set sufficiently high regardless of the prior mean value. The legal speed limit is good enough for the recommended setting. To make sure, by using the real-world dataset in Section V-D, we examined whether the proposed method worked well when we set the initial sample value to the legal speed limit regardless of the prior mean value. We tested it in settings in which the prior means were 1 and 10 (km/h) and the initial sample value was always the legal speed limit, 60 (km/h). Although we failed to estimate the correct traffic velocity by using the prior mean as the initial sample value in these settings for the prior mean in the experiment in Section V-D, by using the legal speed limit as the initial sample value, we could estimate the correct traffic velocity 41 (km/h), as shown in Fig. 17.

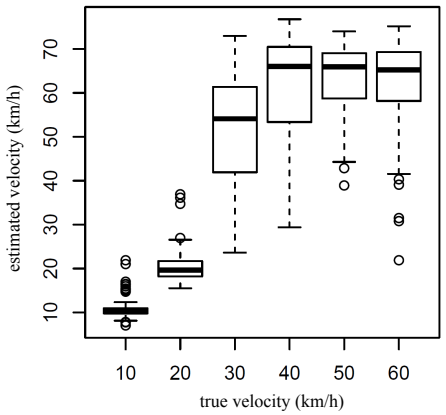


Fig. 18: An extreme setting: true velocity and velocity estimated by the proposed method for a time interval $\Delta t = 10$ (second) in an artificial instance. The average number of vehicles $M_{\mathrm{art}} = 10$. Note that we have used the Tukey boxplot [48]

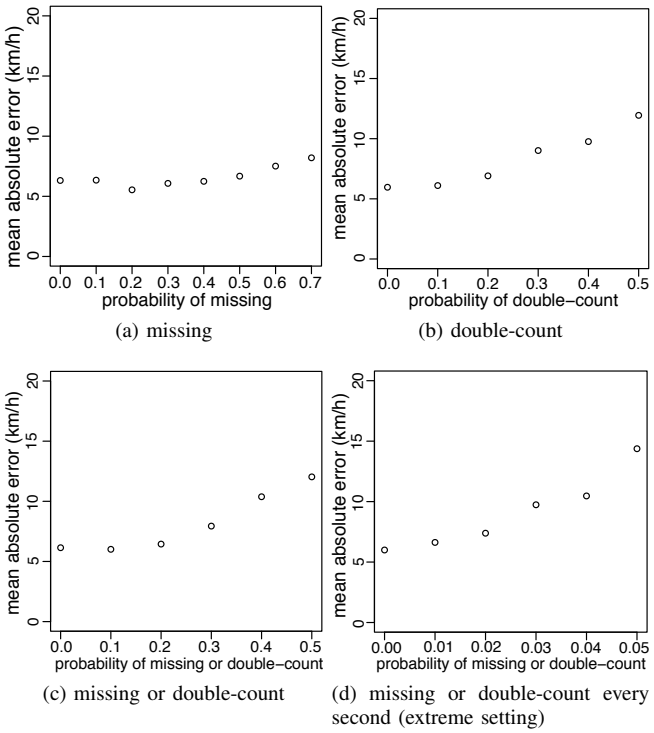
With regard to the number of observations N for $\boldsymbol{x}$, we used $N = 50$ for the artificial and simulation instances and $N = 60$ for the real-world instance. Roughly speaking, the proposed method required more than $N = 10$ observations to get a good estimation. A preliminary analysis indicated that the accuracy did not improve much when more observations N, such as $N = 100$, were used. On the other hand, when we used fewer observations, such as $N = 5$, the proposed method could not estimate an appropriate traffic velocity.

With regard to the calculation cost, the proposed algorithm requires $\mathcal{O}(N^3)$. This cost is determined by the matrix inversion: $\boldsymbol{\Sigma}$ in Eq. (15). Also, since we use the MCMC-based approach, slice sampling, the proposed algorithm requires a large number of iterations for the inference. The use of deterministic algorithms, such as variational Bayes [52], seems to be a promising way of making a more efficient inference.

B. Validation of Limitations of the proposed method

Here, we discuss the limitations of the proposed model, which is caused by our assumptions on the time interval (sampling rate) between observations, the quality of the observation, and the positions of the vehicles at time zero.

When the time interval between observations is quite large, since all of the $L_{j,k}$s for $j, k \in j < k$ are zero in Eq. (7), the likelihood of v given $\boldsymbol{x}$ (Eq. (15)) always takes same value for any v. In this situation, because the posterior distribution becomes equivalent to the prior distribution, the estimated



(a) missing (b) double-count

(c) missing or double-count (d) missing or double-count every second (extreme setting)

Fig. 19: Relationship between probability of artificially added noise to the count and mean absolute error.

traffic velocity converges to the prior mean and we cannot estimate the reasonable traffic velocity. From the definition shown in Eq. (7), the condition under which we can estimate the traffic velocity is

$$\max_{j,k \in j<k} L_{j,k}(v) > 0. \tag{25}$$

From Eq.(25), we can also derive an upper bound of the traffic velocity that the proposed method can deal with:

$$v < \frac{L}{\min_{j \in \{1,2,\ldots,N-1\}}(t_{j+1} - t_j)}. \tag{26}$$

The upper bound is determined by the time interval of $\boldsymbol{t}$ and the road length L. We have seen that the proposed model can estimate a reasonable traffic velocity in most practical settings, including the real-world case study in Section V, where one can obtain, e.g., one image per second and the length of the observed area is $L = 24$ (m). In this section, we examine the limitation of the proposed model in terms of the time interval by using the artificial validation dataset in Section V-B with an extreme setting, that is, the time interval $\Delta t = 10$ (second) and road length $L = 100$ (m). Figure 18 compares the true velocity with the estimated velocity. We can see that the estimated velocity converges to the prior mean, 60 (km/h), from around the upper bound velocity, 36 (km/h), in this extreme setting.

With regard to the quality of the observation variable, that is, the vehicle count $\boldsymbol{x}$, we modeled it probabilistically by taking into consideration statistical noise, as shown in Eqs. (10) to (12). Here, let us examine the influence of the counting error. Using the NGSIM data in Section V-E, we tested the proposed method when some vehicles were constantly missing

Fig. 20: Traffic at intersection captured by web camera in Nairobi [40].

or double-counted through all the observations; this situation corresponds to one in which we cannot find some vehicles because they blend in with the background or one in which some vehicles, such as busses or trucks, are double-counted because they are larger than ordinary vehicles. Figure 19 shows the mean absolute error for the settings in which vehicles are (a) missing (negative noise) in a particular proportion, (b) double-counted (positive noise) in a particular proportion, and (c) missing or double-counted (mixed noise) in a particular proportion. They show that our method is robust against the counting error even when 40 percent of the vehicles are missing or double-counted. The counting error for the double-count has a greater effect on the accuracy since the double-counted vehicles do not become independent. We also tried a more extreme (but unrealistic) setting, where vehicles are miscounted independently in every second during the one minute of the N observations. As shown in Figure 19 (d), the proposed method cannot achieve good accuracy when the noise proportion is more than 5 percent. This is because this type of noise makes the correlation between the observation sequences quite low. Since the velocity estimated by the proposed model is high when the correlation is low, it overestimates the velocity when it is given such fluctuating observations.

To derive the observation model, we assume that the positions of the vehicles at time zero are random and independent from each other. In real world traffic data, the positions of the vehicles are not exactly random and the number of vehicles in the current observation depends on and is correlated to the one of the former observations even if the observations do not have an overlapping area between them. This correlation may cause that our estimation results have an underestimation bias since the strong correlation means larger overlapping area and slower velocity in the proposed model. The experiments using the real-world datasets indicated that such a bias occurred. While the proposed method can estimate velocities with considerable accuracy, the estimated mean values of the velocities are slightly lower than the means of the true velocities in four of the five real-world datasets described in Sections V-D and V-E. However, since this bias is only -2.7 (km/h) on average and is statistically significant (paired t-test, $p \leq 0.05$) in only one of the five datasets (US-101 dataset in the NGSIM datasets), we conclude that the bias is not so large. Solving this problem will be part of our future work.

C. Other Applications of the Proposed Method

While we assume for the proposed model that v is the same for all of the vehicles during the timestamps, our experiment showed that the model is robust enough for situations where each vehicle has a different velocity. Additionally, because the definition in Eqs. (5) and (7) is *invariant with respect to the movement direction of the vehicles*, we can use it in situations where the observations $\boldsymbol{x}$ include vehicles moving inbound as well as outbound without any changes to the model or algorithm. It estimates a common absolute value of the velocity for both lane directions. The directional invariance is preserved even in other multi-directional cases, such as at intersections, as shown in Fig. 20, which shows an image captured in Nairobi, Kenya [40]. The above capabilities also mean that we can use the proposed method for estimating the velocities of crowds, molecules, etc. It can also be used to estimate the velocity separately for each of the three legs of that intersection if we can separately obtain the number of vehicles in each leg.

The Gaussian approximation has another advantage in that we can straightforwardly extend the model so that it can use low-level features by taking $\boldsymbol{x}$ to be real numbers, where the features need to be such that larger values correspond to larger vehicle counts. For example, in the case of analyzing web-camera images, we can use the total area that may correspond to moving objects in an image (TAM) [37]–[39], [41], [49], [50] as $\boldsymbol{x}$. Since we can usually obtain such low-level features more easily than the number of vehicles, this ability is quite useful when we do not have any way to determine the explicit number of vehicles from the raw input data. Here, we examined this extension by using the Web-camera dataset in Section V-D. We tested the proposed method by inputting TAM as $\boldsymbol{x}$. Figure 21 compares manually measured velocity, velocity estimated using the proposed method with TAM as input, and velocity estimated using the proposed method with the vehicle count as input. We can see that the degradation in the estimation accuracy is small when we use the TAM as input of the proposed method.

Regarding future work, it would be interesting to apply the proposed approach to other applications, such as crowds or microscopic images. Also, introducing a more hierarchical model for the hyperparameters may be a way to achieve higher accuracy and robustness within the Bayesian framework. Additionally, we may try deterministic algorithms, such as variational methods, for efficient estimation.

VII. Concluding Remarks

This study tackled the novel task of *traffic velocity estimation from temporal-sequences of the numbers of vehicles without tracking any vehicles or using any labeled training data* for the purpose of traffic monitoring using sensors with low sampling rates. We formulated the task as a density estimation problem by deriving a new model for the temporal-sequences of the numbers of vehicles, which represents the likelihood of the traffic velocity v given temporal-sequences of the numbers of vehicles $\boldsymbol{x}$. For this model, we proposed an efficient approximate method of estimating the posterior mean of the traffic velocity by using slice sampling. In experiments, the proposed approach can estimate a reasonable traffic velocity.

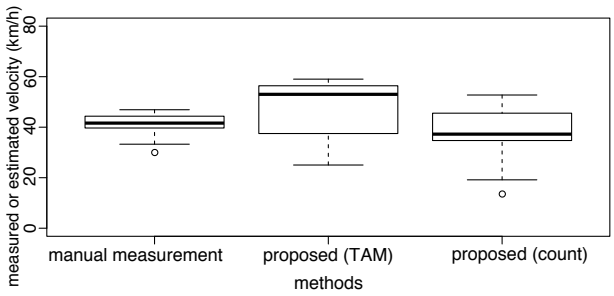


Fig. 21: Manually measured velocity, velocity estimated using the proposed method with the low-level image feature, TAM, as input, and velocity estimated using the proposed method with the vehicle count as input. Note that we have used the Tukey boxplot [48].

Acknowledgments

This research was supported by CREST, JST.

The authors thank AccessKenya.com for permitting us to use their images. We also thank T. Idé, H. Watanabe, S. Yoshihama, M. Teraguchi, M. Tatsubori, H. Muta, T. Suzumura, T. Imamichi, K. Maeda, R. Takahashi, Y. Tsuboi, T. Yoshizumi, and Y. Shinohara for their helpful discussions.

Appendix

A. Definitions of Distributions

Here are the definitions of the uniform, inverse gamma, and Gaussian distributions in terms of their probability density function:

$$\mathcal{U}(x|a,b) \equiv \begin{cases} \frac{1}{b-a}, & a \leq x \leq b, \\ 0, & \text{otherwise}, \end{cases} \quad (x \in \mathbb{R}),$$

$$\mathrm{InverseGamma}(x|a,b) \equiv \frac{b^a}{\Gamma(a)} \left(\frac{1}{x}\right)^{a+1} \mathrm{e}^{-\frac{b}{x}} \quad (x > 0),$$

$$\mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu},\boldsymbol{\Sigma}) \equiv |2\pi\boldsymbol{\Sigma}|^{-\frac{1}{2}} \mathrm{e}^{-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^\top \boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})} \quad (\boldsymbol{x} \in \mathbb{R}^N),$$

where Γ denotes the gamma function. $|\bullet|$ denotes the determinant of the given matrix. The variables in these definitions are not related to the variables that appear in the main text.

References

[1] J. W. Van Lint and N. Van der Zijpp, "Improving a travel-time estimation algorithm by using dual loop detectors," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1855, no. 1, pp. 41–48, 2003.

[2] J. C. Herrera, D. B. Work, R. Herring, X. J. Ban, Q. Jacobson, and A. M. Bayen, "Evaluation of traffic data obtained via gps-enabled mobile phones: The mobile century field experiment," *Transportation Research Part C: Emerging Technologies*, vol. 18, no. 4, pp. 568–583, 2010.

[3] R. Zito, G. d'Este, and M. A. Taylor, "Global positioning systems in the time domain: How useful a tool for intelligent vehicle-highway systems?" *Transportation Research Part C: Emerging Technologies*, vol. 3, no. 4, pp. 193–209, 1995.

[4] J.-Y. Choi, K.-S. Sung, and Y.-K. Yang, "Multiple vehicles detection and tracking based on scale-invariant feature transform," in *Proc. IEEE Intl Conf. Intelligent Transportation Systems*, 2007, pp. 528–533.

[5] A. Kembhavi, D. Harwood, and L. Davis, "Vehicle detection using partial least squares," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 33, no. 6, pp. 1250–1265, 2011.

[6] D. Beymer, P. McLauchlan, B. Coifman, and J. Malik, "A real-time computer vision system for measuring traffic parameters," in *Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 97)*, 1997, pp. 495–501.

[7] Z. Kim and J. Malik, "Fast vehicle detection with probabilistic feature grouping and its application to vehicle tracking," in *Proc. IEEE Intl. Conf. on Computer Vision*, vol. 1, 2003, pp. 524–531.

[8] M. Haag and H.-H. Nagel, "Combination of edge element and optical flow estimates for 3d-model-based vehicle tracking in traffic image sequences," *International Journal of Computer Vision*, vol. 35, no. 3, pp. 295–319, 1999.

[9] S. Indu, M. Gupta, and A. Bhattacharyya, "Vehicle tracking and speed estimation using optical flow method," *International Journal of Engineering Science and Technology (IJEST)*, vol. 3, no. 1, 2011.

[10] J. Lan, J. Li, G. Hu, B. Ran, and L. Wang, "Vehicle speed measurement based on gray constraint optical flow algorithm," *Optik-International Journal for Light and Electron Optics*, vol. 125, no. 1, pp. 289–295, 2014.

[11] K. Robert, "Video-based traffic monitoring at day and night vehicle features detection tracking," in *Proc. IEEE Intl. Conf. on Intelligent Transportation Systems*, 2009, pp. 1–6.

[12] Y.-L. Chen, B. fei Wu, H.-Y. Huang, and C.-J. Fan, "A real-time vision system for nighttime vehicle detection and traffic surveillance," *IEEE Trans. on Industrial Electronics*, vol. 58, no. 5, pp. 2030–2044, 2011.

[13] T. Schoepflin and D. Dailey, "Correlation technique for estimating traffic speed from cameras," *Transportation Research Record: Journal of the Transportation Research Board*, no. 1855, pp. 66–73, 2003.

[14] ——, "Cross-correlation tracking technique for extracting speed from cameras under adverse conditions," *Transportation Research Record: Journal of the Transportation Research Board*, no. 1867, pp. 36–45, 2004.

[15] Y. Cho and J. Rice, "Estimating velocity fields on a freeway from low-resolution videos," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 7, no. 4, pp. 463–469, 2006.

[16] S. Pumrin and D. Dailey, "Roadside camera motion detection for automated speed measurement," in *Intelligent Transportation Systems, 2002. Proceedings. The IEEE 5th International Conference on*. IEEE, 2002, pp. 147–151.

[17] D. J. Dailey and L. Li, "An algorithm to estimate vehicle speed using uncalibrated cameras," *Transportation Research Record 1719*, pp. 27–32, 2000.

[18] D. J. Dailey, F. W. Cathey, and S. Pumrin, "An algorithm to estimate mean traffic speed using uncalibrated cameras," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 1, no. 2, pp. 98–107, 2000.

[19] T. N. Schoepflin and D. J. Dailey, "Dynamic camera calibration of roadside traffic management cameras for vehicle speed estimation," *Intelligent Transportation Systems, IEEE Transactions on*, vol. 4, no. 2, pp. 90–98, 2003.

[20] Y. Malinovskiy, Y.-J. Wu, and Y. Wang, "Video-based vehicle detection and tracking using spatiotemporal maps," *Transportation Research Record: Journal of the Transportation Research Board*, no. 2121, pp. 81–89, 2009.

[21] M. J. Cassidy and R. L. Bertini, "Some traffic features at freeway bottlenecks," *Transportation Research Part B: Methodological*, vol. 33, no. 1, pp. 25–42, 1999.

[22] B. Coifman, "Improved velocity estimation using single loop detectors," *Transportation Research Part A: Policy and Practice*, vol. 35, no. 10, pp. 863–880, 2001.

[23] B. Coifman, S. Dhoorjaty, and Z.-H. Lee, "Estimating median velocity instead of mean velocity at single loop detectors," *Transportation Research Part C: Emerging Technologies*, vol. 11, no. 3, pp. 211–222, 2003.

[24] B. Coifman and M. Cassidy, "Vehicle reidentification and travel time measurement on congested freeways," *Transportation Research Part A: Policy and Practice*, vol. 36, no. 10, pp. 899–917, 2002.

[25] M. Sigari, N. Mozayani, and H. Pourreza, "Fuzzy running average and fuzzy background subtraction: concepts and application," *International Journal of Computer Science and Network Security*, vol. 8, no. 2, pp. 138–143, 2008.

[26] B. Coifman and S. Kim, "Speed estimation and length based vehicle classification from freeway single-loop detectors," *Transportation research part C: emerging technologies*, vol. 17, no. 4, pp. 349–364, 2009.

[27] B. D. Greenshields, J. Thompson, H. Dickinson, and R. Swinton, "The photographic method of studying traffic behavior," in *Highway Research Board Proceedings*, vol. 13, 1934.

[28] B. D. Greenshields, "A study of traffic capacity," in *Highway research board proceedings*. National Research Council (USA), Highway Research Board, 1935.

[29] H. Greenberg, "An analysis of traffic flow," *Operations research*, vol. 7, no. 1, pp. 79–85, 1959.

[30] U. RT, "Speed, volume, and density relationship: quality and theory of traffic flow," *Bureau of Highway Traffic, Yale University*, pp. 141–188, 1961.

[31] J. Del Castillo and F. Benitez, "On the functional form of the speed-density relationship–I: general theory," *Transportation Research Part B: Methodological*, vol. 29, no. 5, pp. 373–389, 1995.

[32] ——, "On the functional form of the speed-density relationship–II: empirical investigation," *Transportation Research Part B: Methodological*, vol. 29, no. 5, pp. 391–406, 1995.

[33] J. S. Drake, J. L. Schofer, and A. D. May Jr, "A statistical analysis of speed-density hypotheses. in vehicular traffic science," *Highway Research Record*, no. 154, 1967.

[34] H. Wang, D. Ni, Q.-Y. Chen, and J. Li, "Stochastic modeling of the equilibrium speed–density relationship," *Journal of Advanced Transportation*, vol. 47, no. 1, pp. 126–150, 2013.

[35] H. Wang, J. Li, Q. Chen, and D. Ni, "Speed-density relationship: From deterministic to stochastic," in *Transportation Research Board 88th Annual Meeting*, 2009.

[36] S. Santini, "Analysis of traffic flow in urban areas using web cameras," in *Fifth IEEE Workshop on Applications of Computer Vision*, 2000, pp. 140–145.

[37] T. Idé, T. Katsuki, T. Morimura, and R. Morris, "Monitoring entire-city traffic using low-resolution web cameras," in *Proceedings of the 20th ITS World Congress, Tokyo*, 2013.

[38] S. Hu, J. Wu, and L. Xu, "Real-time traffic congestion detection based on video analysis," *Journal of Information and Computational Science*, vol. 9, no. 10, pp. 2907–2914, 2012.

[39] X.-D. Yu, L.-Y. Duan, and Q. Tian, "Highway traffic information extraction from skycam mpeg video," in *Proc. IEEE Intl. Conf. on Intelligent Transportation Systems*, 2002, pp. 37–42.

[40] AccessKenya.com, "http://traffic.accesskenya.com/."

[41] T. Idé, T. Katsuki, T. Morimura, and R. Morris, "City-wide traffic flow estimation from limited number of low quality cameras," *Intelligent Transportation Systems, IEEE Transactions on*, pp. — (to appear), 2016.

[42] V. Joshi, N. Rajamani, K. Takayuki, N. Prathapaneni, and L. V. Subramaniam, "Information fusion based learning for frugal traffic state sensing," in *Proceedings of the Twenty-Third international joint conference on Artificial Intelligence*. AAAI Press, 2013, pp. 2826–2832.

[43] N. Radjou, J. Prabhu, and S. Ahuja, *Jugaad innovation: Think frugal, be flexible, generate breakthrough growth*. John Wiley & Sons, 2012.

[44] N. Radjou, "Creative problem-solving in the face of extreme limits," https://www.ted.com/talks/navi_radjou_creative_problem_solving_in_the_face_of_extreme_limits?language=en, 2014.

[45] R. M. Neal, "Slice sampling," *Annals of statistics*, pp. 705–741, 2003.

[46] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, "An introduction to mcmc for machine learning," *Machine learning*, vol. 50, no. 1-2, pp. 5–43, 2003.

[47] C. M. Bishop *et al.*, *Pattern recognition and machine learning*. springer New York, 2006, vol. 1.

[48] J. W. Tukey, *Exploratory data analysis*. Reading, Mass.: Addison-Wesley, 1977.

[49] T. Katsuki, T. Morimura, and T. Idé, "Unsupervised object counting without object recognition," in *Proceedings of the 23rd International Conference on Pattern Recognition (ICPR 2016)*, 2016, pp. — (to appear).

[50] T. Katsuki, T. Morimura, and T. Idé, "Bayesian unsupervised vehicle counting," Technical Report, Tech. Rep., 2013.

[51] United States Department of Transportation, "Ngsim – next generation simulation," http://ops.fhwa.dot.gov/trafficanalysistools/ngsim.htm.

[52] H. Attias and L. W. Ar, "Inferring parameters and structure of latent variable models by variational Bayes," in *Proc. of the Fifteenth Conference on Uncertainty in Artificial Intelligence*. Morgan Kaufmann Publishers, 1999, pp. 21–30.