Volume Estimation using Traffic Signal Event-Based Data from Video-Based Sensors

Xiaofeng Li¹, Yao-Jan Wu¹, and Yi-Chang Chiu¹

Abstract
Traffic volume data is one of the most critical variables for signal retiming. However, collecting traffic volume manually can be time-consuming and costly. In recent years, video-based sensor systems have been applied on signalized intersections for signal timing control. The detectors in video-based sensors generate large amounts of real-time high-resolution event-based data, including signal status and detection status data. The vehicle arrivals for each detection event is a stochastic process and has a relationship with the signal status and the detection duration (time occupancy). Therefore, a modified dynamic hidden Markov model (DHMM) is proposed to estimate vehicular volume by modeling the vehicle arrivals using event-based data collected at signalized intersections. The concept of an additional hidden state is introduced to make the vehicular volume finite by grouping volumes that have only a small probability of occurring into one hidden state. Additionally, a linear regression model is built to estimate the vehicular volume when the output of the DHMM is an additional hidden state. The resulting mean absolute percentage errors of the 15-min estimated volume are 14.1%, 10.3%, and 10.5%, respectively, at three study locations in Tucson, Arizona.

Traffic signals at intersections play a critical role in the road network. As of 2012, over 300,000 intersections in the U.S. have traffic signals installed (1). To mitigate delay and improve traffic movements, signal timings have to be retimed periodically. The traffic volume (turning movement count) data is considered as one of the most important variables in the process of signal retiming. Traditionally, volume data can be collected manually and be used to retime the signal timings, but the data collection process is time-consuming and costly. Therefore, another common practice is to use intelligent transportation system traffic sensors to collect the volume data. In the past decades, inductive-loop detectors have been one of the most common methods to collect volume data. One of the disadvantages of using inductive-loop detectors is that it requires the pavement to be cut when they are installed. This causes a decrease in the pavement life (2). In contrast to the inductive-loop detectors, video-based sensors do not have any negative effects on the pavement. In addition, virtual detectors can be conveniently added and modified in the video-based sensors (3). However, these “virtual-loop detectors” are not implemented in video-based sensors by default. Implementing the virtual-loop detectors requires additional time from traffic engineers and technicians. Additionally, optimizing the layout of these virtual-loop detectors for high-quality data collection is time-consuming.

Different from the virtual-loop detectors, advance-event detectors (AED) are configured in video-based sensors by default because of the need for traffic signal control (green time extension). Traffic signal controllers can collect event-based data from the AED. Recently, event-based data has attracted increased attention and many research studies have been conducted. Researchers have applied event-based data for vehicle classification (4), freeway travel-time estimation (5), and long queue estimation (6, 7). Even though event-based data can supply very rich information to transportation agencies, collecting volume data is still a challenging task. First, the AED is commonly wired together spanning multiple lanes (6–8). Second, the detection area can be extended because of the angle of cameras. The second issue is a problem specific to video-based sensors. The details will be introduced in the problem statement section. The

¹Department of Civil and Architectural Engineering and Mechanics, The University of Arizona, Tucson, AZ

Corresponding Author:
Address correspondence to Yao-Jan Wu: yaojan@email.arizona.edu
above two issues could result in the AED counting multiple consecutive vehicles as one vehicle. The true volume is then underestimated by the AED. Therefore, the goal of this paper is to develop a method for estimating volume data using event-based data collected by video-based sensors. With this proposed method, the volume data can be collected from the existing AEDs directly with large time savings.

The rest of this paper is organized as follows: In the problem statement section, two types of issues in the video-based sensors collecting volume data are explained in more detail. Then the event-based data and ground-truth data are briefly introduced. Next, a method based on the dynamic hidden Markov model (DHMM) is presented. The method performance is evaluated using three types of cases. The conclusions are drawn in the final section.

**Problem Statement**

At signalized intersections, the AEDs are configured to be located upstream from the stop bar for use in green time extension (9). Different types of AEDs have been applied in existing video-based sensors depending on the manufacturer of the sensor. In this study, the AED is configured in the shape of a horizontal bar to detect the vehicles. As shown in Figure 1, the AED, enclosed in a red dashed circle, covers three through-lanes of the west-bound (WB) through-movement. It is common to configure AEDs on multiple lanes because the lanes are contained in the same signal phase. The AED detects the vehicles for extending the green time of the coordinated phases in the same signal phase.

Figure 2a shows how an AED counts two concurrent vehicles on different lanes as one vehicle. Therefore, the count recorded by the AED underestimates the actual volume. Figure 2b shows how the AED is “floating” above all the vehicles, because the AED is drawn on the monitor. Moreover, the camera is angled downward to watch an approach as far as possible. These two factors combine to extend the actual detection area in the horizontal and vertical direction. Therefore, multiple vehicles in one lane could be also counted as one vehicle when vehicles are in the detection area concurrently. Overall, the AED in video-based sensors cannot collect accurate volume data directly.

**Data Description and Collection**

All data was collected from Speedway Boulevard in Tucson, Arizona to train and assess the proposed approach. Speedway Blvd is one of the most important corridors in Tucson, Arizona because it connects between the I-10 freeway and the University of Arizona. Both the event-based and ground-truth volume data were collected between 5:30 a.m. and 8:00 p.m. Three intersections along Speedway Blvd were selected as the study locations, as shown in Figure 3.

**Event-Based Data**

Event-based data is high-resolution data containing a series of events (signal, pedestrian, detector, and controller communication) generated with timestamps in real time. Only two types of event-based data are used in this research: the signal phase change events and the vehicle detection events. The signal events dataset contains the start and end timestamps of the signal status in each signal phase, for example the beginning and ending of the green time of one phase. A sample dataset is shown in Figure 4a. The vehicle detection event dataset only includes two basic events representing vehicles triggering and leaving the AED. The timestamps of each time the AED is triggered on and off are stored in the dataset as shown in Figure 4b. The difference in time between each AED on and off indicates the time interval of vehicles occupying the detector that is referred to as time occupancy. The time period between the AED on and off events is referred to as a detection event duration. Continuous event-based data was collected from the MaxView (10) database in the City of Tucson.

**Ground-Truth Data**

The ground-truth volume data was collected separately for the training and test datasets. Both of the training and test datasets were manually collected from video recordings. The videos were recorded using MaxView’s graphical user interface because the video can show the
vehicles traveling and the status of detectors concurrently. For the testing dataset, the ground-truth vehicular volume data was collected in 15-min intervals. For the training dataset, the number of vehicles passing through the AED was manually recorded for each detection event duration.

Table 1 shows all the periods when both training and test datasets were collected. Note that the training...
dataset was collected at Speedway Blvd and Mountain Ave WB and the test data was collected from all the locations listed in Table 1.

**Methodology**

The proposed method is able to estimate the volume for each cycle using signal status and time occupancy from video-based sensors as input. The vehicular volume varies over time for each detection event duration and has a high correlation with the signal status and the time occupancy. Therefore, the vehicular volume passing the AED should be modeled as a stochastic process. The DHMM is used to estimate the vehicular volume for each event duration, and the resulting volume is then aggregated into 15-min intervals.

Figure 5 illustrates the framework of the proposed volume estimation method. First, the 15-min volume $V$ is initialized as zero at the beginning for each interval. Then the signal status groups (SSGs) and transformation of two signal status groups (TTSSGs) of the $c$th cycle are determined based on the input data. The definitions of SSG and TTSSG will be further explained in the dynamic hidden Markov model section, below. The parameters of the DHMM are selected from the results of the training model and are used to decode all hidden states. If the output has any additional hidden states, the proposed linear model would be used to estimate its vehicular volume. Otherwise, the numbers in hidden states are regarded as the estimated vehicular volumes directly. Both of the vehicular volumes estimated from the additional hidden state and from the numbers in other hidden states are added together as the total volume of the $c$th cycle, $V(c)$.

The process of the proposed estimation method will repeat until the number of cycles reaches the total number of cycles $C$ within 15 min. Then, the final vehicular volume output of this process is aggregated into 15-min intervals.

Three tasks need to be accomplished prior to the volume estimation: 1) determine the finite number of hidden states plus an additional state, 2) calibrate the parameters for the DHMM, and 3) decode the additional hidden state into vehicular volume. Once the training stage is

<table>
<thead>
<tr>
<th>Date</th>
<th>Period</th>
<th>Intersection</th>
<th>Direction</th>
<th>AED location</th>
<th>Through-lanes configuration</th>
<th>Training/test dataset</th>
</tr>
</thead>
<tbody>
<tr>
<td>04/30/2018</td>
<td>4:45–6:45 p.m.</td>
<td>Speedway &amp; Mountain</td>
<td>WB</td>
<td>167 ft</td>
<td>Three lanes turn</td>
<td>Training</td>
</tr>
<tr>
<td>05/01/2018</td>
<td>5:30–9:00 a.m.</td>
<td>Speedway &amp; Mountain</td>
<td>WB</td>
<td>167 ft</td>
<td>Three lanes turn</td>
<td>Test</td>
</tr>
<tr>
<td>05/01/2018</td>
<td>5:45–7:30 p.m.</td>
<td>Speedway &amp; Mountain</td>
<td>WB</td>
<td>167 ft</td>
<td>Three lanes turn</td>
<td>Training</td>
</tr>
<tr>
<td>06/25/2018</td>
<td>10:00 a.m.–5:30 p.m.</td>
<td>Speedway &amp; Campbell</td>
<td>EB</td>
<td>85 ft</td>
<td>Three lanes turn</td>
<td>Test</td>
</tr>
<tr>
<td>06/27/2018</td>
<td>6:00–7:35 p.m.</td>
<td>Speedway &amp; Mountain</td>
<td>WB</td>
<td>167 ft</td>
<td>Three lanes turn</td>
<td>Training</td>
</tr>
<tr>
<td>06/28/2018</td>
<td>5:15 a.m.–2:20 p.m.</td>
<td>Speedway &amp; Mountain</td>
<td>WB</td>
<td>167 ft</td>
<td>Three lanes turn</td>
<td>Training</td>
</tr>
<tr>
<td>07/07/2018</td>
<td>4:15–8:00 p.m.</td>
<td>Speedway &amp; Stone</td>
<td>EB</td>
<td>70 ft</td>
<td>Two lanes</td>
<td>Test</td>
</tr>
</tbody>
</table>

*The distance between the AED and the stop bar (1 ft = 0.305 m).
completed, volume can be estimated through the volume estimation stage on a cycle-by-cycle basis.

**Dynamic Hidden Markov Model**

**Hidden States Determination.** The hidden Markov model (HMM) was developed based on a Markov Chain to analyze the hidden states \( Q \) and observations \( O \). Three more parameters are needed to build the HMM, that is, start probability \( \pi \), transition probability \( A \), and emission probability \( B \).

The time occupancy can be considered as a set of observations because the time occupancy can be a potential indicator for the vehicular volume. However, the signal status could be affecting the vehicular volume. For instance, the time occupancy for a single vehicle passing the AED during the green time would be shorter than that during the red time. Therefore, the combination of the signal status and the vehicular volume is deemed as a set of hidden states. The “on” and “off” signal statuses of the AED are used to represent the signal status of each detection event. All possible SSGs are defined as below:

- Red-to-Green (RG): the AED is turned on during the red time of the previous cycle and is turned off during the green time of the current cycle. RG happens when the queue covers the AED in the previous cycle.
- Green-to-Green (GG): the AED is turned on during the green time and is turned off during the same green time. GG is more likely to happen because vehicles usually travel at a higher speed, resulting in longer headways between two consecutive vehicles. GG is the most commonly observed SSG.
- Green-to-Red (GR): the AED is turned on during the green time and turned off during the red time.
- Red-to-Red (RR): the AED is turned on during the red time and turned off during the red time. RR often happens when there is light traffic.

The HMM requires a finite number of hidden states \( (11) \). Owing to the randomness of vehicle arrivals, the vehicular volumes passing through the AED per detection-event duration would result in a larger number of hidden states. Each hidden state requires an emission probability matrix that is calibrated from the ground-truth data. To avoid this time-consuming calibration process, the vehicular volumes with small percentages for each SSG should be grouped into one hidden state, referred to as the additional hidden state.

The probability density function (PDF) of vehicular volumes can be fitted using the lognormal distribution model

\[
N(v_v|\mu_i, \sigma_i) = \frac{1}{v_v \sigma_i \sqrt{2 \pi}} \exp \left[ -\frac{(\ln v_v - \mu_i)^2}{2\sigma_i^2} \right], q_i > 0 \tag{1}
\]

where \( v_v \) is the vehicular volume of SSG\( i \), \( i \in \{RG, GG, GR, RR\} \); \( \mu_i \) and \( \sigma_i \) the mean and variance of \( \ln v_v \).

Based on Equation 1, the 95th percentile of the cumulative distribution function (CDF) for each SSG can be calculated as the critical value to define the vehicular volumes with small percentages. When the vehicular volumes are lower than the critical value, the vehicular volumes are assigned to the hidden states for each SSG. Otherwise, the vehicular volumes will be grouped in an additional hidden state for each SSG. This grouping decision criterion is

\[
\begin{align*}
&\{ v_v = a_i, \text{if } \Phi_f(\ln(v_v)) > 95\% \\
&v_v = v_v, \text{otherwise}
\end{align*}
\tag{2}
\]

where \( \Phi_f \) is the CDF of the normal distribution of the SSG\( i \) and \( a_i \) is the additional hidden state of the SSG\( i \) for volume estimation.

**Transition Probability.** The first order Markov Chain is widely used in the HMM. One significant assumption of the first order Markov Chain is that the probability of the state at time \( t \), \( Q_t \), equal to the \( m \)th hidden state \( q_m \), only depends on the previous state at time \( t - 1 \). This means it is independent of other previous time states \( (12) \), as

\[
P(Q_t = q_m|Q_{t-1} = q_n, Q_{t-2} = q_l, \ldots, Q_1 = q_k) = P(Q_t = q_m|Q_{t-1} = q_n) \tag{3}
\]

where \( P(Q_t = q_m|Q_{t-1} = q_n) \) is the state transition probability \( a_{mn} \) from state \( q_n \) to \( q_m \).

To make a realistic estimation of vehicular volumes for each cycle, two assumptions need to be made. The first assumption is that vehicles cannot occupy the AED longer than a cycle length without a gap. This assumption is used to prevent the HMM from estimating vehicles continuously for more than one cycle. The second assumption is that GG must follow RG because RG generally happens with long queues being cleared followed by vehicles arriving with longer headways before GR or RR happens. The second assumption is made to train the model using minimum transition probability matrices and to avoid unnecessary calculations.

The DHMM is estimated for each cycle from the beginning of the green time to the end of the red time. For each HMM, both the vehicular volume and the SGG need to transform into different states. However, each SGG could only transform into certain specific SGGs in one cycle. Table 2 shows six possible transformations, referred to as TTSSGs. Within a single cycle,
Table 2. Transformations of Two Signal Status Groups

<table>
<thead>
<tr>
<th>From/To</th>
<th>RG</th>
<th>GG</th>
<th>GR</th>
<th>RR</th>
</tr>
</thead>
<tbody>
<tr>
<td>RG</td>
<td>√</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GG</td>
<td></td>
<td>√</td>
<td></td>
<td>√</td>
</tr>
<tr>
<td>GR</td>
<td></td>
<td></td>
<td>√</td>
<td></td>
</tr>
<tr>
<td>RR</td>
<td></td>
<td></td>
<td></td>
<td>√</td>
</tr>
</tbody>
</table>

Figure 6. Example of one possible HMM.

GG could transform into GG, GR, or RR, and GR can only transform into RR. However, RG can only transform into GG, not GR or RR according to the second assumption. Hence, six transition probability matrices are used in the DHMM. For the initial probability, four initial probability vectors corresponding to four SSGs will need to be calibrated to explain the initial transition between the start of the estimation process and the hidden states at the first timestamp.

The numbers of detection events and the SSGs vary for each cycle. Therefore, different combinations of initial probability vectors and transition probability matrices should be selected for each cycle from the training results. Figure 6 shows one possible DHMM structure within one cycle. At $t_1$, the SSG is RG that, combined with the corresponding vehicular volumes and the additional group of volumes, are defined as the hidden states. $o_1$ represents the observation of time occupancy at $t_1$. Other observations follow the same process to determine their respective hidden states. The initial probability is the vector of RG. The transition probability matrix is determined based on the two consecutive SSGs. Take the first transition in Figure 4, for example. The transition probability is the matrix showing the transition from RG at $t_1$ to GG at $t_2$.

The optimal state transition probability $a_{nm}$ and initial probability $\pi_n$, should be calculated by training the model using the ground-truth data. The Baum–Welch algorithm (13) is commonly used to train the HMM. The underlying idea of the algorithm is that an initial $\hat{a}_{nm}$ is updated by $\tilde{a}_{nm}$ until it reaches convergence. $\tilde{a}_{nm}$ is the expected number of transitions from state $n$ to $m$ divided by the expected number of transitions from state $n$. The original Baum–Welch algorithm cannot handle DHMM with multiple transition probability matrices and initial probability vectors. Therefore, the algorithm is revised to be applied to this paper. During one cycle, GG to GG and RR to RR would occur several times in order. However, the rest of four TTSSGs would either only happen once or not happen during one cycle. Firstly, let $j$ denote one of the TTSSGs, $j \in \{RG, GG, GG, GR, GG, RR, GR, RR, RR, RR\}$. $C$ is the total number of cycles in the training dataset and let $E_j$ denote the set of all cycles including TTSSG. If the $c^h$ cycle in the training dataset has the TTSSG, the cycle can be stored in $E_j$. There is a total of $C_j$ cycles in $E_j$. For each cycle in $E_j$, the TTSSG starts from the timestamp $t_j'$ and ends with the timestamp $T_j'$. $\hat{a}_{nm}(TTSSG_j)$ can be calculated as the summation of probability from state $n$ to $m$ of each cycle and each timestamp including the TTSSG, divided by the summation of the probability from state $n$ to any state $k$ of each cycle and each timestamp including the TTSSG. This expected frequency percentage is

$$\hat{a}_{nm}(TTSSG_j) = \frac{\sum_{c' = 1}^{C_j} \sum_{t = t_j'}^{T_j'} \xi_1(n,m)M_j}{\sum_{c' = 1}^{C_j} \sum_{t = t_j'}^{T_j'} \sum_{k = 1}^{N_j} \xi_1(n,k)}$$ (4)

where $\xi_1(n,m)$ is the probability of being in state $n$ at time $t$ and state $m$ at time $t + 1$, given the time occupancy sequences and previously updated parameters. The details of the calculation of $\xi_1(n,m)$ can be found in Baum (13). $M_j$ is the total of hidden states where TTSSG would transform.

Similarly, initial $\pi_n$ is updated by $\tilde{\pi}_n$ until convergence. Firstly, let $i$ denote one of the SSGs, $i \in \{RG, GG, GR, RR\}$. Let $F_i$ denote the set of all cycles including SSG. If the $c^h$ cycle in the training dataset begins with the SSG, the cycle can be stored in $F_i$. There is a total of $C_i'$ cycles in $F_i$. The expected frequency probability of state $n$ at the beginning, $\hat{\pi}_n$, is calculated using

$$\hat{\pi}_n(\text{SSG}_i) = \sum_{c' = 1}^{C_i'} \gamma_1(n)$$ (5)

where $\gamma_1(n)$ is the probability of state $n$ at the beginning of the $c''$ cycle, given the time occupancy and the parameters updated in previous iteration. A detailed
calculation of $\gamma_t(n)$ can be found in Baum (13). Both $a_{nm}$ and $\pi_n$ are pseudo probabilities for each update, so the probabilities should be normalized each time for the next update.

In the Baum–Welch algorithm, the emission probability is used to calculate $\xi_t(n,m)$ and $\gamma_t(n)$ when updating the initial and transition probabilities. The mixture probability models are applied to fit the distribution of time occupancy for each hidden state to estimate the PDF of the observations. Then, the PDF is utilized as the emission probability $b_{ml} = P(O_t = o_l|Q_t = q_m)$ for the DHMM.

Gaussian and Lognormal Mixture Model for Emission Probability. Since time occupancy is continuous, the Baum–Welch algorithm cannot estimate the emission probability directly. To deal with the continuous observations, mixture possibility models are the common method to fit the distribution of observations and to calculate the PDF. In this paper, both the Gaussian mixture model (GMM) and the lognormal mixture model (LMM) are utilized to fit the time occupancy distribution of each hidden state. The mixture models are

\[
\text{GMM}(O_m|\theta_m) = \sum_{i=1}^{K_m} \pi_v \frac{1}{\sqrt{2\pi} \sigma_v} \exp \left[ -\frac{(x_m - \mu_v)^2}{2\sigma_v^2} \right]
\]

\[
\text{LMM}(O_m|\theta_m) = \sum_{i=1}^{K_m} \pi_v \frac{1}{\sqrt{2\pi} \sigma_v} \exp \left[ -\frac{(\ln x_m - \mu_v)^2}{2\sigma_v^2} \right]
\]

where $O$ is a set of time occupancy of the $m$th hidden state; $\pi_v$ is the weight of the $v$th Gaussian or lognormal distribution; $\sum_v$ is the covariance matrix of the two types of models respectively; $\mu_v$ is the mean of the two types of models respectively, and $\theta_m$ represents all optimal parameters of the GMM and the LMM respectively. $\theta_m$ can be estimated using the expectation maximization (EM) algorithm when the component number $K_m$ is known. The details of EM can be found in Bishop (14).

The goal of the DHMM is to calculate the vehicular volume through selecting a sequence of hidden states. This process of hidden state selection is called the decoding task. One common method to decode the hidden states of the DHMM is the Viterbi algorithm (15). The Viterbi algorithm is a kind of dynamic programming method. At each timestamp, the most probable sequences are calculated by taking the maximum likelihood of all previous state sequences. The best hidden state sequence can be selected based on the maximum likelihood. In our research, the decoding result includes a group of vehicular volume hidden states and the additional hidden state.

The additional hidden state needs further processing to determine the vehicular volume.

Modeling for the Additional Hidden State

Since the vehicular volume is unknown for the additional hidden state, a linear regression model is proposed to estimate the vehicular volume. The general form is

\[
v = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \epsilon
\]

where $\alpha$ is the intercept; $\beta$ is the time occupancy, the distance the detector was occupied by vehicles; $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ are the coefficients of the variables; and $\epsilon$ is the error between estimated results by the model and ground-truth data.

In addition, the distance $d$ between the AED and the stop bar would affect the data collection of the time occupancy when the SSG is RG. This is because the queue is more likely to reach the AED closer to the stop bar, resulting in a longer time occupancy. Long time occupancy would overestimate the vehicular volume using Equation 8 when the SSG is RG. To handle this issue, the weighting factor $\omega$, as defined in Equation 9, was introduced to reduce the impact of different AED’s distances. $vv$ would be multiplied by $\omega$ when the SSG of the additional hidden state is RG.

\[
\omega = \frac{d'}{d}
\]

where $d'$ is the distance between the AED and the stop bar of the applied location, and $d$ is the distance between the AED and the stop bar at the location used to build the linear model.

Implementation

As shown in Figure 3, ground-truth data was collected from Speedway Blvd and Mountain Ave WB to train the models in the proposed method. The ground-truth data was collected from each of the three locations to evaluate the performance of the proposed method.

Model Calibration

In Figure 7, the histograms show the observations of vehicular volume, and the red lines are the fitted results of a lognormal probability model based on Equation 1. The number shown on the top of each arrow is the critical value of each SSG. This critical value is to distinguish the hidden states of vehicular volumes from the additional hidden state.
For GG and RR, the critical value is eight and three vehicles, respectively. Both critical values of GG and RR were calculated based on Equation 2. Therefore, the hidden states of GG are \{GG1, GG2, GG3, ..., GG8, GGa\} and the hidden states of RR are \{RR1, RR2, RR3, RRa\}. For RG and GR, because of the limited training data size, it cannot be used to generate the PDF of time occupancy for each vehicular volume if the critical values are calculated by Equation 2. Considering RG and GR comprise only a small percentage, three vehicles are selected as the critical value of both RG and GR. Therefore, the hidden states of RG are \{RG1, RG2, RG3, RGa\} and the hidden states of GR are \{GR1, GR2, GR3, GRa\}.

The emission probability of each hidden state is calculated using the mixture models in Equations 6 and 7. The optimal number of \(K\) and the most suitable mixture model are determined by maximizing the log likelihoods. Figure 8 shows the PDF of time occupancy for each hidden state using various mixture models. The results could efficiently be used to calculate the emission probability of each hidden state. Please note that the distribution of RGa is almost flat. This might be because a wider range of vehicular volumes are grouped into the additional hidden state.

After the emission probabilities are determined, six transition probability matrices are calculated using the revised Baum–Welch algorithm. Figure 9 visually presents the transition probability for each TRSSG. The transition probability with GG involved is significantly different from the one with RR involved. The transition probability matrices with GG involved (except GG_RR) are more evenly distributed. For RG_GG, the probability of transforming to GGa is a little higher than transforming into other hidden states. Since vehicles are starting to depart from the long queue during RG at a lower speed, the vehicle platoon has a higher probability of passing the AED during the green time, leading to GGa following RG. However, the probability of transforming to GRa is a little lower than transforming into other hidden states under GG_GR. This is mainly because the headways are longer at the beginning of the queuing process, leading to the lower probability of GRa. All of the three transition probability matrices with RR involved have a much higher probability to transform into RR1. This is because RR always happens when the traffic is light. In addition, vehicles travel slowly with shorter headways during the red time. If multiple vehicles arrive, RR tends to be converted into RG. Based on the reasons mentioned above, the probability of transforming to RR1 is higher than transforming into other hidden states.

To decode the additional hidden state into the vehicular volume, the linear model was built using 465 observations (11% of the training dataset), where the vehicular volume is greater than the critical value of each SSG. Table 3 summarizes the calibration results of the linear model in Equation 8. It is expected that the variable RG is not significant. However, the low significance level of
RG does not create a large impact on volume estimation. Moreover, the impact is further mitigated by the weight \( w \) defined in Equation 9.

**Performance Evaluation**

As shown in Table 1, the data collected from three locations are used to evaluate the performance of the proposed method using the mean absolute percentage error (MAPE), as

\[
MAPE = \frac{1}{X} \sum_{x=1}^{X} \frac{|V_x - \hat{V}_x|}{V_x}
\]  

(10)

where \( V_x \) is the observation volume during the \( x^{th} \) 15-min interval; \( \hat{V}_x \) is the estimated volume during the \( x^{th} \) 15-min interval, and \( X \) is the total number of samples.
Figure 10 shows the results of these comparisons. The MAPEs are 14.1%, 10.3%, and 11.6% for the three study locations, indicating reasonable accuracy of the proposed method. Generally, the proposed method can successfully estimate the traffic volume during the daytime without being affected by lane configuration. The MAPE of Speedway Blvd and Mountain Ave WB is slightly higher compared with the other two cases. Figure 10a shows that the higher MAPE is mostly because of the estimated results before 6:30 a.m. It was found that the headlights and shadows of vehicles increase the estimation error. The AED was triggered early because of the headlights of vehicles, resulting in longer time occupancy than the one in normal cases. Therefore, the proposed method overestimated the volume in the early morning as indicated in Figure 10a. Similar issues can be observed for the nighttime case in Figure 10c. Additionally, the error at 4:00 p.m. of Figure 10b is also higher. By checking the recorded videos, one fire truck was observed traveling west during 4:00–4:15 p.m., resulting in long queues occupying the AED over two cycles. With such a long time occupancy, the probability density for each hidden state is close to zero. The fitted results would have erroneous emission probabilities, especially for those of the additional hidden state. Then the volume would be underestimated. Moreover, the internal issue of video-based sensors would cause additional errors for volume estimation. For instance, during 6:30–6:45 p.m. in Figure 10c, the AED was continuously triggered for about 30 s without any vehicles. This internal issue resulted in higher estimated volume.

To further investigate effects of cancellation between the positive and negative volume estimation over a period of time, the proposed method is also applied in the scenarios of 5-min estimation intervals. In the scenarios of the 5-min interval, the MAPEs for Speedway Blvd and Mountain Ave, Speedway Blvd and Campbell Ave, and Speedway Blvd and Stone Ave are 18.3%, 17.1%,

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient</th>
<th>Standard error</th>
<th>t-value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>7.626</td>
<td>0.522</td>
<td>14.61</td>
<td>0.000</td>
</tr>
<tr>
<td>Time occupancy (s)</td>
<td>0.510</td>
<td>0.018</td>
<td>27.81</td>
<td>0.000</td>
</tr>
<tr>
<td>GR (0 = no; 1 = yes)</td>
<td>-5.619</td>
<td>1.031</td>
<td>-5.449</td>
<td>0.000</td>
</tr>
<tr>
<td>RG (0 = no; 1 = yes)</td>
<td>-1.038</td>
<td>0.733</td>
<td>-1.417</td>
<td>0.157</td>
</tr>
<tr>
<td>RR (0 = no; 1 = yes)</td>
<td>-5.487</td>
<td>0.782</td>
<td>-7.016</td>
<td>0.000</td>
</tr>
<tr>
<td>Multiple R-squared</td>
<td>0.706</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Adjusted R-squared</td>
<td>0.703</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Figure 10. Comparison between the ground-truth values and estimated values: (a) Speedway Blvd and Mountain Ave WB, (b) Speedway Blvd and Campbell Ave eastbound (EB), and (c) Speedway Blvd and Stone Ave EB.
and 13.7%, respectively. As expected, volume estimation with a shorter interval generally had slightly larger MAPEs compared with the results with a 15-min interval. Considering most traffic volume applications use a 15-min or a longer interval, the results are satisfactory and the 15-min interval is recommended to be used to estimate the volume.

Conclusion

This study developed an innovative method to estimate the volume at signalized intersections using event-based data collected by video-based sensors. The focus of this approach was to model the vehicle arrivals as a stochastic process. A DHMM was built to estimate the vehicular volume using signal status and detector time occupancy. Moreover, one linear model was utilized to decode the additional hidden state to improve estimation results. Compared with ground-truth volume data, the accuracy of the proposed method has MAPEs of 14.1%, 10.3%, and 10.5% at three different study locations, respectively. Moreover, it was found that the proposed method can accurately estimate volume without being affected by lane configuration.

Since the performance of this proposed method relies heavily on the detection quality of the video sensor, this issue can be addressed by collecting additional data from more scenarios to train the proposed method. Additionally, the proposed method could potentially integrate the data collected from virtual-loop detectors from video sensors or other detectors, for example, loop detectors, to improve the estimation performance. In the future, the proposed method will be tested on more locations and compared with other algorithms and sensor products, for example, radar and LIDAR sensors for further evaluation.

Acknowledgments

The authors would like to thank the City of Tucson for funding and data support. We would also like to thank Dr. Zong Tian, Paul Burton, Diahn Swartz, Bob Hunt, Roger Kestler, and Christopher French for providing valuable advice and technical support in this project. Also, the authors would like to thank Eric Huettner for collecting ground-truth data. Special thanks to Adrian Cottam for his assistance in English proofreading.

Author Contributions

The authors confirm contribution to the paper as follows: study conception and design: XL, Y-JW, Y-CC; data collection: XL; analysis and interpretation of results: XL; draft manuscript preparation: XL, Y-JW. All authors reviewed the results and approved the final version of the manuscript.

References


The Standing Committee on Artificial Intelligence and Advanced Computing Applications (ABJ70) peer-reviewed this paper (19-01828).