# Vehicle Re-Identification using Wireless Magnetic Sensors: Algorithm Revision, Modifications and Performance Analysis

Rene O. Sanchez*, Christopher Flores†, Roberto Horowitz*, Ram Rajagopal‡ and Pravin Varaiya†
*Department of Mechanical Engineering
University of California, Berkeley, CA 94720, US
Emails: r2sanche@me.berkeley.edu and horowitz@me.berkeley.edu
†Sensys Networks, Inc, Berkeley, CA 94710, US
Emails: chrisf@sensysnetworks.com and pvaraiya@sensysnetworks.com
‡Department of Civil and Environmental Engineering
Stanford University, Stanford, CA 94305, US. Email: ram.rajagopal@stanford.edu

***Abstract*—A vehicle re-identification method based on matching vehicle signatures obtained from wireless magnetic sensors was studied on a single lane loop on-ramp. Different modifications were implemented in the algorithm in order to address limitations of the system when vehicles stop/move slowly over the detectors. The original and modified vehicle re-identification algorithm results were compared against ground truth data obtained from video. Based on the ground truth data it was possible to determine the percentage of vehicles that are re-identified and the number of those vehicles that are misidentified as a function of different algorithm parameters. For this analysis, vehicles were divided into two subsets: i) uncongested and ii) congested. The original method mismatched percentage was around or below 15% for the uncongested vehicle subset and between 20% to 60% for the congested one. With the modified method it was possible to improve the matching rate as well as the accuracy of the matching algorithm. For the uncongested subset, the modified method showed a higher vehicle re-identification rate while maintaining the mismatched percentage around or below 8%. The main improvement over the original method was achieved on the congested vehicle subset, since the number of re-identified vehicles was increased over the original method while keeping the mismatched percentage around or below 14%.**

***Keywords: Vehicle Re-Identification; Magnetic Sensors; Signal Matching***

## I. INTRODUCTION

The work presented in this paper was motivated by the results obtained from a queue estimation field test performed at the Hegenberger Rd. loop on-ramp in the Caltrans Bay Area District, described in [1]. Four different queue estimation methods were studied, including one that is based on a vehicle re-identification algorithm, presented in [2].

The queue estimation method based on the vehicle re-identification algorithm currently used by Sensys Networks, Inc. performed better than the other three methods, but it significantly under-performed in estimating queue lengths during congested on-ramp conditions. This was due to the fact that the number of re-identified vehicles was very low when vehicles were stoping and moving at low speeds as they were going over the detectors.

Vehicle re-identification by matching electromagnetic signatures captured from inductive or magnetic sensors appears as one of the most efficient and cost-effective methods to re-identify vehicles without raising privacy concerns associated with vehicle tracking [3]. However, electromagnetic signature methods tend to be unreliable under congested conditions, when stop and go traffic is present. Vehicle re-identification methods based on loop detectors, like the one described in [4], are not only sensitive to speed changes in between vehicle detection stations, but also to changes in velocity as the vehicles go over the detectors. Even when speed normalization is decoupled from the matching algorithm, as in [3], the algorithms still rely on the assumption that vehicles go over the detectors at constant speed. This assumption is unrealistic for implementations in arterial streets, congested freeways, and on-ramps where stop-and-go traffic is unavoidable. The vehicle re-identification method discussed in this paper does not rely on the constant speed assumption and does not require vehicle speed. However, the poor performance of the vehicle matching algorithm in the queue estimation study called for an algorithm revision, improvement and performance analysis.

The paper is organized as follows: the vehicle re-identification method is summarized in Section II. The test site is described in Section III. The ground truth (GT) and vehicle detection system (VDS) data are explained in Section IV. The revision details are presented in Section V. Section VI discusses the modifications. Section VII contains the performance analysis of the original and modified vehicle re-identification method. Conclusions are collected in Section VIII.

## II. VEHICLE RE-IDENTIFICATION METHOD

The vehicle re-identification method summarized in this section is described in [2].

### A. *Vehicle Magnetic Signature*

The magnetic vehicle signature consists of a collection of peak value sequences (local maxima and minima) extracted from the 'raw' magnetic signals measured by an array of sensors. Each sensor has a three-axis magnetometer that measures the $x$, $y$ and $z$ directions of the earth's magnetic field as a vehicle goes over it. Each extracted peak is paired with a local time stamp, which can be used to determine relative separation among peaks generated from the same sensor. Each sensor generates three peak sequences extracted from the $x$, $y$ and $z$ component signals, which constitute a signature slice $X^i = (X^i_x, X^i_y, X^i_z)$. Seven slices constitute the vehicle's signature. Figure 4 (a) shows two slices of a vehicle's signature measured at the entrance and at the exit arrays plotted using the local time stamp component and the peak amplitudes before any processing.

The vehicle re-identification algorithm takes two signatures, $X_i = (X^1, \cdots, X^7)$ and $Y_j = (Y^1, \cdots, Y^7)$ where $(X^q, Y^r)$ are slices, and computes a distance (a measure of dissimilarity) between each pair of slices componentwise. The distance $\delta(X_i, Y_j)$ is defined as the minimum of the distances between all pairs of slices $(X^q, Y^r)$. This method uses a dynamic time warping approach to calculate the distance between signatures.

### B. *Vehicle Re-Identification Algorithm Summary*

The vehicle re-identification is done in two steps:

*1) Signal Processing Step:* In this step, each pair $(X_i, Y_j)$ of entrance and exit vehicle signatures is compared to produce a distance $d(i, j) = \delta(X_i, Y_j) \geq 0$ between them. The smaller $\delta(X_i, Y_j)$ the more likely it is that $X_i, Y_j$ are signatures of the same vehicle. This step reduces the two signature arrays $X = \{X_i, i = 1, \cdots, N\}$ and $Y = \{Y_j, j = 1, \cdots, M\}$ to the $N \times M$ distance matrix $D = \{d(i, j) \mid 1 \leq i \leq N, 1 \leq j \leq M\}$.

*2) Matching Step:* In the second step a *matching function* assigns to each distance matrix $D$ a matching $\mu : \{1, \ldots, N\} \to \{1, \ldots, M, \tau\}$,with the following interpretation: $\mu(i) = j$ means that the upstream vehicle $i$ is declared to match (be the same as) downstream vehicle $j$; $\mu(i) = \tau$ means $i$ is declared not to match any downstream vehicle.

It is assumed that the distance matrix $D$ is characterized by two probability density functions (pdf), $f$ and $g$: $f$ is the pdf of the distance $\delta(X_v, Y_v)$ between the signatures at the entrance and exit sensor arrays of the same randomly selected vehicle $v$, and $g$ is the pdf of the distance $\delta(X_v, Y_w)$ between two different randomly selected vehicles $v \neq w$. The $f$ and $g$ pdfs are assumed Gaussian and their statistics, the mean $\mu$ and the standard deviation $\sigma$, are part of the algorithm parameters that must be determine beforehand.

The Probability of Turn $(\beta)$ is another matching algorithm parameter. In an arterial implementation, $\beta$ can be defined as the percentage of vehicles that went through the upstream array but turned before reaching the downstream array, and it can be determined from field observations or experience. But another consideration may govern the choice of $\beta$. The larger $\beta$ is, the more stringent is the requirement of a match, and the lower is the probability of an incorrect match.

## III. TEST SITE

The Hegenberger on-ramp is a suitable location to analyze the vehicle re-identification method. It is a single lane ramp, which allows for the testing of the algorithm without having to take into account multiple lanes on-ramp dynamics.

The vehicle detection system deployed at the Hegenberger on-ramp for this study was developed by Sensys Networks, Inc. This system consists of an Access Point (AP240-ESG), a repeater (RP240-B), and 14 wireless magnetic sensors (VDS240) installed as shown in Figure 1 (b) and (c). See [5] for details on this vehicle detection system.


Fig. 1. (a) Hegenberger on-ramp entrance and exit. (b) Entrance Vehicle Re-Id Array. (c) Exit Vehicle Re-Id Array.

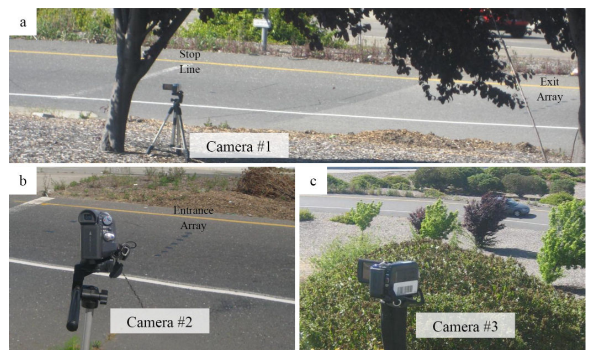


Fig. 2. (a) Camera recording vehicles leaving the on-ramp. (b) Camera recording vehicles entering the on-ramp. (c) Camera to capture any incident occurring at the test site.

## IV. DATA

### A. *Ground Truth Data*

Ground truth data was obtained from videos recorded on May 11, 2010 from 4:07 pm to 5:35 pm.

Three independent cameras were used to obtain the ground truth data (see Figure 2). From the second camera (Figure 2 (b)) it was possible to obtain the time $s_{GT_k}$ when each entering vehicle $k$ crossed the entrance array, where $s_{GT_1} < s_{GT_2} < \cdots s_{GT_{N_{GT}}}$. From the first camera (Figure 2 (a)) it was possible to get the time $t_{GT_l}$ when each exiting vehicle $l$ went through the exit array, where $t_{GT_1} < t_{GT_2} < \cdots t_{GT_{M_{GT}}}$. The GT data consists of two vectors $\{s_{GT_k}, k = 1, \cdots, N_{GT} = 543\}$ and $\{t_{GT_l}, l = 1, \cdots, M_{GT} = 534\}$. The GT matching of upstream to downstream vehicles $k \to l$ was done visually and resulted in 534 matches.

### B. *Vehicle Detection System Data*

Consider the link formed by the entrance and exit arrays shown in Figure 1. During the video recording time interval, detection events indexed $i = 1, \cdots, N$ were registered by the entrance array at times $s_1 < s_2 < \cdots s_N$. Detection events indexed $j = 1, \cdots, M$ were registered by the exit array at times $t_1 < t_2 < \cdots t_M$. The upstream sensor measures a signature $X_i$ each time there is a vehicle detection event $i$ and the corresponding time $s_i$. The downstream sensor measures a signature $Y_j$ each time there is detection event $j$ and the corresponding time $t_j$. For this study, the vehicle detection system data consists of two arrays $\{(s_i, X_i), i = 1, \cdots, N = 522\}$ and $\{(t_j, Y_j), j = 1, \cdots, M = 527\}$. Note that due to the nature of the vehicle detection system, detection errors cannot be avoided and may create multiple signatures of the same vehicle at one location or may result on missing signatures due to undetected vehicles at the entrance and/or exit array [1].

To be able to determine the number of vehicles that are correctly matched by the algorithm and the percentage of those vehicles that are mismatched, mappings of the form $k \to i$ and $k \to l \to j$ were obtained. With them it is possible to determine if a signature $(X_i, Y_j)$ corresponds to the same vehicle or to different vehicles.

### C. Vehicle Subsets

During the video recording period, the ramp presented two traffic modes: uncongested and congested. For this analysis, uncongested conditions correspond to the time interval from 4:07 pm to 5:07 pm, when the on-ramp queue was below on-ramp capacity and vehicles with index $k$ for $1 \leq k \leq 399$ went through the on-ramp. The uncongested vehicle subset is constituted by these vehicles. The congested conditions time interval occurs from 5:07 pm to 5:35 pm, when the queue length was around or beyond on-ramp capacity, vehicles were stopping or going slowly over the entrance array, and vehicles with entering vehicle index $k$ for $400 \leq k \leq 534$ went through the ramp. The congested vehicle subset is composed of these vehicles. Based on the vehicle detection system data, it is possible to achieve a maximum correct matching of 477 vehicles assuming a perfect matching algorithm, where 362 vehicles correspond to the uncongested vehicle subset and 115 to the congested one.

### D. 23 Chosen Vehicles

In order to be able to analyze the vehicle re-identification algorithm in detail, 23 vehicles were chosen from the 543 that entered the ramp, as shown in Table I. The selection criteria was that their entrance and exiting signatures were available and unique and that they were traveling through the middle of the lane when going over the arrays.

## V. VEHICLE RE-ID METHOD REVISION

### A. Signal Processing Step Revision

The plots on Figure 3 are gray scale coding of the distance matrix of the 23 chosen vehicles signatures (left) and the distance matrix of the complete vehicle signatures data set (right) calculated using the original signal processing algorithm. Each square in the plots represents a distance between a $X_i$ and $Y_j$ vehicle signature combination; a darker color indicates shorter distance and a greater chance that $X_i$ and $Y_j$ come from the same vehicle. The gray scale used in Figure 3 and Figure 5 goes from $\min(D)$ to $.75 median(D)$, where $D$ represents the distance matrix being plotted.

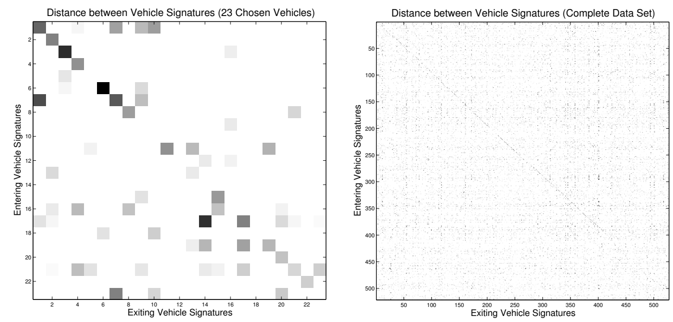


Fig. 3. Original Distance Matrix from entrance to exit array for the (left) 23 Chosen Vehicles (right) Complete Vehicle Data Set

Figure 3 (left) shows multiple dark squares along the diagonal, where diagonal entries correspond to distances between signatures of the same vehicle. The distance value of the diagonal entries are listed in Table I. Note that the diagonal entries corresponding to the congested vehicle subset are not as dark (i.e. small) as the uncongested ones. It seems that phenomena occurring specifically during congestion affect the $f$ and $g$ probability density functions (pdf). This is further corroborated by the plot of the distance matrix for the complete data set, shown on Figure 3 (right). For the uncongested vehicle subset, a diagonal line is present and is consistently darker and visually distinguishable from the off-diagonal $D$ entries. However, dark diagonal entries vanish at the lower right portion of the plot, where the distances between vehicles signatures from the congested vehicle subset are plotted.

The original signal processing algorithm produces different $f$ and $g$ pdfs for different on-ramp modes. The $f$ and $g$ pdfs become similar during on-ramp congested conditions, which reduces the effectiveness of the matching algorithm. The signal processing method should be able to maintain the $f$ and $g$ pdfs invariant to traffic conditions.

| Uncongested Conditions | | | | Congested Conditions | | | |
|---|---|---|---|---|---|---|---|
| Index $k$ | Veh. Type | Org. Dist. | Mod. Dist. | Index $k$ | Veh Type | Org. Dist. | Mod. Dist. |
| 76 | minivan | .25 | .23 | 449 | car | .50 | .41 |
| 135 | truck | .28 | .23 | 459 | SUV | .30 | .35 |
| 140 | truck | .20 | .18 | 471 | pickup | .42 | .24 |
| 187 | pickup | .29 | .24 | 472 | car | .28 | .16 |
| 194 | truck | .71 | .70 | 490 | SUV | .40 | .38 |
| 222 | car | .15 | .18 | 492 | car | .33 | .29 |
| 236 | minivan | .24 | .14 | 516 | car | .34 | .35 |
| 259 | SUV | .30 | .30 | 517 | car | .36 | .25 |
| 282 | pickup | .45 | .32 | 519 | car | .35 | .18 |
| 305 | car | .46 | .28 | 527 | car | .53 | .15 |
| 380 | car | .29 | .25 | | | | |
| 393 | SUV | .48 | .48 | | | | |
| 419 | van | .38 | .29 | | | | |

TABLE I
23 CHOSEN VEHICLES.

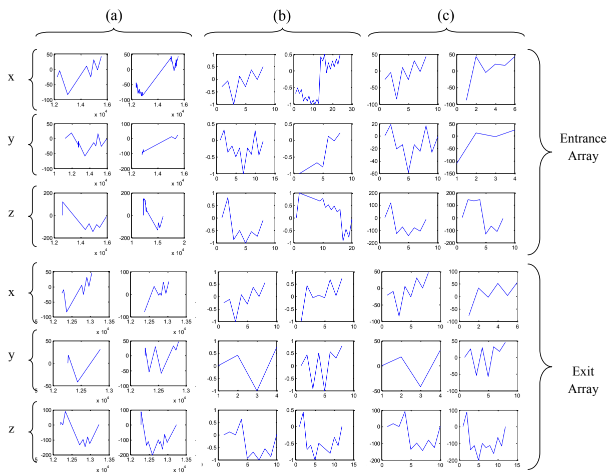


Fig. 4. Vehicle $k = 527$ two sensor signature slice at the entrance and exit arrays (a) Raw peak values (b) Processed for distance calculation (Original Method) (c) Processed for distance calculation (Modified Method)

### B. Vehicle Signatures Revision

Entrance and exit signatures of the 23 chosen vehicles were studied in order to explain the different $f$ and $g$ pdfs observed for the uncongested and congested vehicle subsets.

Figure 4 (b) shows two slices of the entrance and exit signatures of vehicle $k = 527$. Note that the upstream slices show more peaks than the downstream ones. Vehicle 527 went slowly over the entrance array and at free flow speed over the exit array. The extra peak phenomenon was observed in the signatures of vehicles going slowly or stopping over the entrance arrays. During congestion, vehicle signatures tend to have more noisy peaks, most of which are small compared to the dominant peaks of the signatures. A vehicle signature with noisy peaks leads to a larger $\mu_f$ and $\sigma_f$, since the difference on the number of peaks is penalized in the distance calculation independently of their magnitude.

Another finding was related to the importance of the $x$, $y$ and $z$ component of a vehicle signature slice when calculating distances. The original distance method assigned the same weight to the distance obtained between the $x$, $y$, and $z$ components when calculating distances between signature slices. By changing this weights, it was observed that it is possible to increase the dissimilarity between the $f$ and $g$ pdfs during congestion.

Sometimes signature slice components ($x$,$y$ and $z$) of different vehicles look very similar after the peak sequences have been normalized by the maximum absolute value of their elements as a preprocessing step before the distance calculation is performed. The distance between such slices is small and may leads to significant errors in the matching step. It was observed that when the raw amplitude of the peaks is considered in the distance calculation between such signatures, the distance increases for signatures from different vehicles and remains unchanged for signatures from the same vehicle.

| | Default | 23 Veh. | | Iterative | |
|---|---|---|---|---|---|
| | *Both* | *Org* | *Mod* | *Org* | *Mod* |
| $\mu_f$ | .125 | .36 | .27 | .30 | .28 |
| $\sigma_f$ | .058 | .12 | .09 | .09 | .08 |
| $\mu_g$ | .67 | .54 | .60 | .56 | .62 |
| $\sigma_g$ | .23 | .11 | .17 | .13 | .16 |

TABLE II
DIFFERENT $f$ AND $g$ STATISTICS FOR THE ORIGINAL AND MODIFIED SIGNAL PROCESSING METHOD

### C. Matching Step Revision

The matching algorithm was studied using the 23 chosen vehicles. It was concluded the match rate and accuracy are directly related to $\mu_f$, $\sigma_f$, $\mu_g$, $\sigma_g$ and $\beta$. Furthermore, it was observed that one of the reasons that explain the low matching rate during the queue estimation study, even during uncongested on-ramp conditions, was directly related to the $f$ and $g$ parameters used at the Hegenberger on-ramp. The $f$ and $g$ statistics used at the ramp where assumed to be equal to the ones satisfactorily used at many arterial installation sites (see Table II, second column). This assumption was incorrect, since $f$ and $g$ statistics based on the 23 chosen vehicles distance matrix (see Table II), show that the default $f$ and $g$ statistics do not model accurately the signature distances at the on-ramp.

It is important to calculate $f$ and $g$ parameters for each test site, since they are site dependent and influence the matching rate. These parameters can be obtained using an iterative method as the one suggested in [2], which does not require GT. Note that the iteratively obtained $f$ and $g$ parameter extracted from the complete vehicle data set are very similar to the ones extracted from the 23 chosen vehicles for both the modified and the original signal processing method, and very different to the default values.

## VI. VEHICLE RE-ID METHOD MODIFICATIONS

### A. Signal Processing Algorithm Modification

The following modifications were implemented in the original signal processing algorithm:

i) There were adjustments in the way vectors of different sizes are compared using dynamic time warping. This included modifying the way in which extra peaks are penalized when vectors being compared are of different sizes. ii) A peak processing step was implemented before the distance calculation in order to remove noisy peaks resulting from vehicles traveling slowly or stopping on top of the arrays. This step uses the local time stamp component available for each signature peak described in Section II. See Figure 4 for a comparison between the preprocessed signature slices for the original and the modified method. iii) Different weights were assigned to the $x$, $y$ and $z$ components of the distance between two vehicle slices. The $x$ component was assigned the larger weight and the $y$ component the smaller one. iv) The distance calculation is performed without normalizing the peak sequences. Once a distance is obtained between the components of two signature slices, a normalization step is performed.

### B. Signal Processing Improvements

The plot on Figure 5 (left) is a gray scale coding of the distance matrix of the 23 chosen vehicles signatures using the modified signal processing algorithm. Note that in contrast to Figure 3 (left), dark squares are present along the entire diagonal, thus suggesting that $f$ and $g$ statistics are somehow invariant to on-ramp traffic conditions. In Figure 5 (right), where the distance matrix of the complete data set calculated using the modified distance method is plotted, it is possible to see a darker diagonal band at the bottom right corner of the plot. This diagonal band corresponds to distances generated from the same vehicle belonging to the congested vehicle subset. This distinction between diagonal and off-diagonal entries was not present in the original distance matrix shown in Figure 3 (right), which helps explain why the matching rate was specially low during congestion.

From Table I it is possible to see that the modified distance among signatures from the same vehicle, $\delta(X_k,Y_k)$, are similar for the uncongested and congested vehicle subsets. Figure 6 shows for each of the 23 chosen vehicles, $k$, $D_k = \{d(k,l) = \delta(X_k,Y_l) \mid 1 \leq l \leq 23\}$. For each $D_k$, the data points given by $\delta(X_k,Y_l)$ for $l \neq k$ are plotted as black dots while the $\delta(X_k,Y_k)$ distance is plotted as a blue circle. The top plot contains the distances obtained using the original method. This plot shows that for vehicles from the congested vehicle set, $\delta(X_k,Y_k)$ is generally not the smallest entry of $D_k$, which is undesired and affects the performance of the matching algorithm. With the modified signal processing method, as displayed in Figure 6 (bottom), the $\delta(X_k,Y_k)$ data point, represented by a blue circle, is generally the lowest element of $D_k$ for vehicles belonging to both uncongested and congested vehicle subsets. This improves the matching algorithm accuracy.

Finally, after comparing the $f$ and $g$ pdfs for the original and modified method, listed in Table II, it is clear that the modified distance method $f$ and $g$ statistics will benefit the matching algorithm performance due to the larger difference between $\mu_f$ and $\mu_g$ and the smaller values of $\sigma_f$ and $\sigma_g$.

### C. Matching Algorithm Modification

The matching algorithm was not modified. However, it was observed that the $f$ and $g$ statistics as well as $\beta$ play an important role in the number of matched vehicles and the percentage of mismatched vehicles produced by the algorithm.

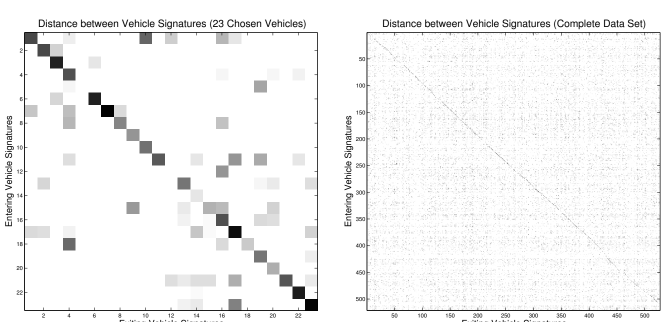


Fig. 5. Modified Distance Matrix from entrance to exit array for the (left) 23 Chosen Vehicles (right) Complete Vehicle Data Set

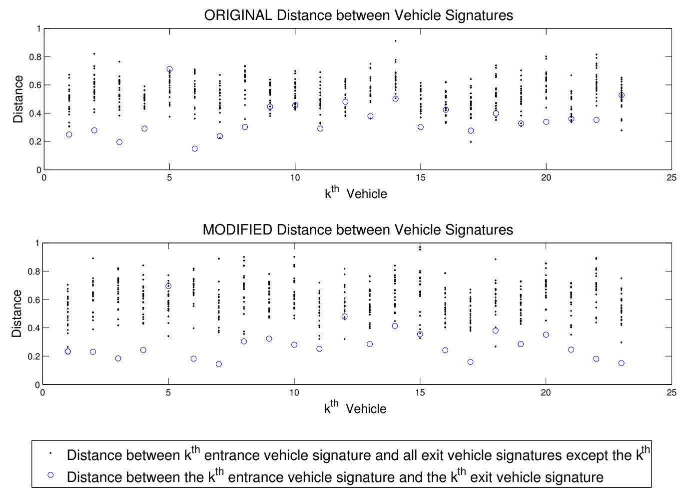


Fig. 6. Comparison Between the Original Distance Method and the Modified Distance Method

## VII. VEHICLE RE-ID RESULTS

### A. Default vs Iterative $\mu_f$, $\sigma_f$, $\mu_g$, $\sigma_g$ Results

In this section, two different sets of $\mu_f$, $\sigma_f$, $\mu_g$, $\sigma_g$ were used to study the effect of $f$ and $g$ pdfs variations on the matching algorithm performance for the original and the modified signal processing methods. The default and the iteratively obtained set of parameters, listed in Table II, were used.

Table III shows the results obtained with the matching algorithm for the complete vehicle data set, the uncongested vehicle subset and the congested vehicle subset. The modified method has a higher number of re-identified vehicles in comparison to the original method for both set of parameters. The number of matches over the original method results increased by 56% for the default case and by 13 % for the iterative one. The increase in vehicle matches did not result in an increase of the percentage of incorrectly matched vehicles. For the original method 23 % of vehicles were incorrectly matched when using the default parameters and 16 % when using the iterative parameters. For the modified method, from 166 matched vehicles obtained using the default parameters, 7% were incorrectly matched, while 7 % out of 368 vehicles were

| | Default | | Iterative | |
|---|---|---|---|---|
| | *Org* | *Mod* | *Org* | *Mod* |
| Matched(total) | 106 | 166 | 325 | 368 |
| Incorrectly M. | 24 | 11 | 53 | 24 |
| Matched(uncong) | 90 | 139 | 257 | 279 |
| Incorrectly M. | 14 | 8 | 32 | 14 |
| Matched(cong) | 16 | 27 | 68 | 89 |
| Incorrectly M. | 10 | 3 | 21 | 10 |

TABLE III
MATCHING RESULTS USING THE DEFAULT AND ITERATIVE $f$ AND $g$ STATISTICS USING THE MODIFIED AND ORIGINAL METHOD

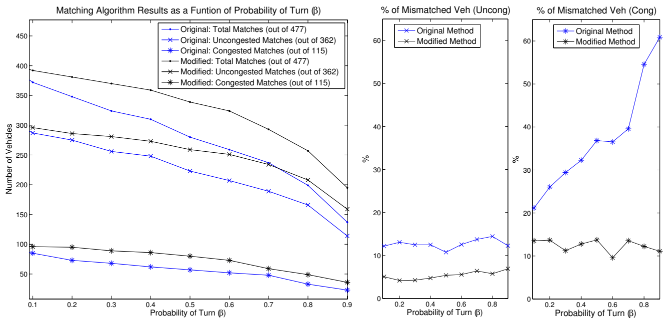


Fig. 7. (left) Matched Veh. as a function of $\beta$ (middle) % of Uncongested Mismatched Veh. as a function of $\beta$ (right) % of Congested Mismatched Veh. as a function of $\beta$

misidentified using the iterative values. The modified distance method seems to increase the matching rate in comparison with the original distance method while keeping the percentage of incorrect matches low and constant for significant $\mu_f$, $\sigma_f$, $\mu_g$, $\sigma_g$ variations. Note that the use of f and g statistics that properly model the distance between signatures (e.g. iteratively obtained parameters), makes a significant difference in the number of matched vehicle but does not impact the accuracy of the matches.

When the results of the matching algorithm are analyzed by vehicle subset, further differences were encountered. First, the percentage of mismatched vehicles is larger for the congested vehicle subset in comparison to the uncongested one, for the original and the modified signal processing methods. The percentage of incorrectly matched vehicles changes from 15% to 62.5% when using default parameter values and the original method. This value changes from 12% mismatched percentage to 31% when the iterative parameters are used instead. The modified distance method also shows differences in the accuracy obtained for the uncongested and congested vehicle subset results. The percentage of incorrect matches changes from 6% to 11% when using the default parameters. This number changes to 5% incorrect matches for the uncongested vehicles subset and 11 % for the congested subset when using the iterative parameters. The decrease in performance of the matching algorithm during congestion is observed for both the original and the modified signal processing method. Note that the original method accuracy is highly dependent on the values of $\mu_f$, $\sigma_f$, $\mu_g$, $\sigma_g$ while the accuracy of the modified method seems to remain unchanged.

### B. Iterative $\mu_f$, $\sigma_f$, $\mu_g$, $\sigma_g$ with varying $\beta$ Results

The matching algorithm results presented in this section were obtained using the iteratively obtained $\mu_f$, $\sigma_f$, $\mu_g$, $\sigma_g$ parameters and using both the original and the modified vehicle re-identification method as $\beta$ was varied.

From Figure 7 (left) it is observed that the modified method has higher matching rate for all $\beta$ values considered for this analysis for the complete vehicle data set, the uncongested vehicle subset and the congested vehicle subset. Figure 7 (middle) shows the percentage of incorrectly matched vehicles for the uncongested vehicle subset as function of $\beta$. Note that both methods have a mismatch percentage that remains somehow constant as $\beta$ is varied. However, an advantage of the modified distance method is that the percentage of incorrect matches is around or below 8% , in comparison to the 15% obtained with the original method. Figure 7 (right) shows the percentage of incorrectly matched vehicles for the congested vehicle subset. The percentage of incorrect matches are larger for all $\beta$ for the original and modified methods in comparison to the uncongested results. Observe that while the original method mismatch percentage increases from 20% at low $\beta$ values to 60% for large ones, the modified distance method percentage of incorrectly matched vehicles remains around or below 14% for all $\beta$ values.

The original algorithm matching rate and accuracy is affected by $\beta$ variations. The modified distance method matching rate is affected by changes in $\beta$, but the accuracy of the matches seems to remain constant.

## VIII. CONCLUSION

The vehicle re-identification method based on matching vehicle magnetic signatures obtained with wireless magnetic sensors was studied on a single lane on-ramp. Based on this study, different modifications were implemented in the algorithm in order to address limitations of the system when vehicles travel slowly or stop while going over the detectors. The original and modified vehicle re-identification algorithm results were compared against ground truth data obtained from video. Based on the ground truth data it was possible to determine the percentage of vehicles that were re-identified and the number of vehicles that were misidentified as a function of different algorithm parameters. The modified distance method resulted in an increase in the number of re-identified vehicles over the original system while keeping an overall incorrectly matched vehicle percentage below 10% when matching algorithm parameters like $\mu_f$, $\sigma_f$, $\mu_g$, $\sigma_g$, and $\beta$ were varied. The best performance was observe during uncongested on-ramp conditions, with a percentage of mismatched vehicles around or below 8%, while during congested on-ramp conditions this number increased to 14%.

## REFERENCES


[1] R. O. Sanchez, R. Horowitz, and P. Varaiya, "Analysis Of queue estimation methods using wireless magnetic sensors," accepted for publication at the *Trans. Res. Rec., Journal of the Transportation Research Board*, 2011.

[2] K. Kwong, R. Kavaler, R. Rajagopal, and P. Varaiya. "Arterial travel time estimation based on vehicle re-identification using wireless magnetic sensors," in *Trans. Res. Part C: Emerging Technol.*, vol. 17, no. 6, pp. 586-606, 2009.

[3] M. Ndoye, V. F. Totten, J. V. Krogmeier, and D. M. Bullock, "Sensing and signal processing for vehicle reidentification and travel time estimation," in *IEEE Trans. Intell. Transp. Syst.*, vol. 12, no. 1, pp. 119-131, March 2011.

[4] C. Sun, S. G. Ritchie, K. Tsai, and R. Jayakrishnan, "Use of vehicle signature analysis and lexicographic optimization for vehicle reidentification on freeways," *Transp. Res. Part C: Emerging Technol.*, vol. 7, no. 4, pp. 167185, Aug. 1999.

[5] A. Haoui, R. Kavaler, and P. Varaiya, "Wireless magnetic sensors for traffic surveillance," in *Transp. Res. Part C: Emerging Technol.*, vol. 16, no. 3, pp. 294-306, 2008.