Vehicle Reidentification using Heterogeneous Detection Systems

UCI-ITS-TS-WP-04-3

Cheol Oh
Stephen G. Ritchie
Shin-Ting Jeng

Department of Civil and Environmental Engineering and
Institute of Transportation Studies
University of California, Irvine
cheolo@uci.edu, sritchie@uci.edu, tjeng@uci.edu

January 2004
Institute of Transportation Studies
University of California, Irvine
Irvine, CA 92697-3600, U.S.A.
http://www.lts.uci.edu
Vehicle Reidentification using Heterogeneous Detection Systems

Cheol Oh
Department of Civil and Environmental Engineering and Institute of Transportation Studies
University of California, Irvine, CA 92697-3600, U.S.A.
Voice) 949-824-5623 Fax) 949-824-8385
choo@uci.edu

Stephen G. Ritchie
Department of Civil and Environmental Engineering and Institute of Transportation Studies
University of California, Irvine, CA 92697-3600, U.S.A.
Voice) 949-824-4214 Fax) 949-824-8385
srichie@uci.edu

Shin-Ting Jeng
Department of Civil and Environmental Engineering and Institute of Transportation Studies
University of California, Irvine, CA 92697-3600, U.S.A.
Voice) 949-824-5623 Fax) 949-824-8385
sjeng@uci.edu

Word Count: 4,075 + 10*250 = 6,575

Vehicle reidentification using heterogeneous detection systems

Cheol Oh, Stephen G. Ritchie, and Shin-Ting Jeng
Institute of Transportation Studies and Dept. of Civil and Environmental Engineering
University of California, Irvine, CA 92697, USA

ABSTRACT
This paper investigates the feasibility of real-time vehicle reidentification algorithm development at a signalized intersection where different traffic detection technologies were employed at upstream and downstream locations. Previous research by the authors on vehicle reidentification has utilized the same traffic sensors (e.g. conventional square inductive loops) and detectors (e.g. high speed scanning detectors cards) at both locations. In this study, an opportunity arose for the first time to collect a downstream dataset from a temporary installation of a prototype innovative inductive loop sensor, known as a “blade” sensor, in conjunction with conventional inductive loops upstream. At both locations advanced high speed scanning detector cards were used. Although the number of vehicles for which data could be collected was small, encouraging results were obtained for vehicle reidentification performance in this system of mixed traffic detection technologies. In future large-scale applications of vehicle reidentification approaches for real-time traffic performance measurement, management and control, it would be most beneficial and practical if heterogeneous as well as homogeneous detection systems could be supported. This initial study yielded many useful insights about this important issue, and demonstrated on a small scale the feasibility of vehicle reidentification in a system with heterogeneous detection technologies.
1. Introduction

Travel time has been identified as a particularly important traffic parameter for evaluating the performance of dynamic traffic systems by transportation researchers and engineers. It is also important because it is an input to advanced transportation management and information systems (ATMIS) to alleviate traffic congestion and its associated impacts.

A promising approach to obtain travel times is tracking vehicle to identify their locations and arrival times, so that travel times can be readily collected. A variety of sensor technologies have been developed and tested for tracing individual vehicles on transportation networks. Use of global positioning system (GPS) and in-vehicle tag-based automatic vehicle identification (AVI) technologies have been successfully used for obtaining accurate travel times. However, privacy issues still remain with such systems, and a limited market penetration does not yet allow us to measure wide-area transportation performance. As a result, it would be advantageous if individual vehicles could be traced without any privacy concerns on wide-area transportation networks.

In order to meet the aforementioned requirement, there has recently been substantial interest in implementing vehicle identification systems that anonymously trace vehicles in a network. Examples include license plate matching (1), use of existing loop detectors with high speed scanning detector cards to generate inductive signatures (2-8), laser-based detection systems (9) providing vehicle length, and video-based vehicle signature generation (10) using video image processing technology.

Previous studies performed by authors (2-8) have proven that accurate travel times can be obtained from inductive signature-based vehicle reidentification using new detector card technology. Because inductive loops are still the dominant surveillance system in the U.S. and many other countries, use of such loops for vehicle reidentification is potentially quite cost-effective.

This paper investigates the feasibility of real-time vehicle reidentification algorithm development at a signalized intersection where different traffic detection technologies were employed at upstream and downstream locations. Previous research by the authors on vehicle reidentification has utilized the same traffic sensors (e.g. conventional square inductive loops) and detectors (e.g. high speed scanning detectors cards) at both locations. In this study, an opportunity arose for the first time to collect a downstream dataset from a temporary installation of a prototype innovative inductive loop sensor, known as a "blade" sensor, in conjunction with conventional inductive loops upstream. At both locations advanced high speed scanning detector cards were used.

The following section of the paper introduces the blade sensor that is able to produce unique vehicle signatures. Data collection and vehicle feature extraction for blade signature is presented in the third section. The next section describes an algorithm for vehicle reidentification with the heterogeneous detection system used in this study. An analysis of travel times using the outputs of the algorithm is then presented. Finally, conclusions including comments and findings are provided.
2. Blade sensor

Traditional applications of inductive loop sensors have focused on counting vehicles or detecting the presence of vehicles. For such purposes, the ideal loop should approximate the vehicle's periphery (11). A physical configuration of 6 x 6 (1.8 x 1.8 m) is a commonly used size for inductive loops that measure counts and presence. More recently, inductive loops have been utilized for outputting inductive signatures for vehicle reidentification purposes. The standard 6 x 6 (1.8 x 1.8 m) loop configuration is not ideal for this purpose since the square geometry results in the integration of the inductive signature over the traversal distance. Therefore, if this smoothing effect, which can remove distinctive features from the inductive signature, can be eliminated it may make vehicle reidentification more effective. The blade sensor addresses the loop configuration problem and incorporates additional improvements to the inductive loop detection system through use of a high-speed scanning detector card.

The blade is a new remote vehicle sensor technology. The physical embodiment of this concept uses two matched oscillating LRC circuits whose induction coils are orientated contained within a single, solid 'sensor blade' that is then embedded in a 3/16 inch wide pavement slot (for a permanent installation). The sensing coil is oriented toward the surface of the pavement and the reference coil is oriented toward the base of slot. Because the sensing coil is positioned nearer overpassing vehicles, it responds more strongly to this stimulus than the reference coil. Data collection is initiated by simultaneously charging both circuits to a threshold voltage using an impulse function and then allowing them to rapidly decay to a base line asymptote. This differential signal is amplified and digitized using an A/D converter.

A continuous stream of signed integers is generated by the blade sensor, which can be monitored by a dedicated on-board microprocessor. The resulting measurement data produce the vehicle’s inductive signature. Figure 1 shows the temporary surface installation of blade sensors as deployed in this study and an example of a blade vehicle signature.

In its present configuration, the blade sensor collects data from two parallel sensor blades separated by a distance of 6 feet and oriented at an angle of 20° to the direction of the traffic flow. This orientation allows for a significant amount of valuable data to be generated including speed, the number of axles, and wheel based-vehicle length. The prominent peaks shown in Figure 1-(d) represent the wheels passing over the sensors. A clearer view of the composite metallic profile of the vehicle, which allows us to differentiate the vehicle wheel part from the vehicle body part, can also be seen.

The temporary surface mounted version of the blade sensor is an out-of-pavement installation that does not require pavement cutting. This version is particularly useful for short-term studies.
3. Vehicle signature analysis and feature extraction/selection

Vehicle signature analysis for vehicle reidentification can be generally separated into two components: feature extraction and classification. The first component seeks to extract salient and parsimonious features from raw detector output, while the second component classifies or matches the vehicles using feature vectors.

3-1. Data collection

In this study, blade sensors were installed next to existing conventional square inductive loop stations upstream and downstream on westbound Irvine Center Drive at the intersection of Alton Parkway and Irvine Center Drive in Irvine, California on January 21, 2003. Vehicle inductive signatures were generated from each type of loop sensor using high speed scanning detector cards.

Each of the detector cards being used to collect the blade signatures had a 40GB hard-drive. The signatures were recorded to the local hard-drives. A laptop computer was used to start the data collection, set the time, etc., and to download the signatures from the cards. Figure 2 shows the blade signature data collection layout.
One-hour of data collected from 11:40 am to 12:40 pm constituted the available data set for both conventional loop and blade loop data. In addition, 140 blade vehicle signatures collected in the right-most lane of the downstream detector station and upstream conventional loop signatures constituted the valid signature data set that could be used for feature analysis and algorithm development for vehicle reidentification. The vehicle reidentification algorithm was developed and tested based on the different detector systems: the conventional inductive loops upstream and blade sensors downstream. Therefore, a vehicle reidentification algorithm was developed with the first 70 vehicle pairs, and the other 70 vehicle pairs were used for algorithm testing. Table 1 presents the through movement vehicles including vehicle types and volumes collected in the right-most lane at the downstream station.

Candecorders were also installed at each station for the purpose of video ground truthing. The video ground truthing was performed based on visual inspection identifying an upstream vehicle on a monitor, and then searching for matching the corresponding vehicle downstream on another monitor. True travel times were obtained by comparing the time stamps of each vehicle at both upstream and downstream stations.
Table 1 Blade vehicle classification data for downstream lane 3 (through movement)

<table>
<thead>
<tr>
<th>Vehicle type</th>
<th># vehicles</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Motorcycle</td>
<td>1</td>
<td>0.69</td>
</tr>
<tr>
<td>Bus</td>
<td>1</td>
<td>0.69</td>
</tr>
<tr>
<td>Passenger car</td>
<td>71</td>
<td>51.724</td>
</tr>
<tr>
<td>Pickup</td>
<td>17</td>
<td>11.724</td>
</tr>
<tr>
<td>SUV</td>
<td>33</td>
<td>23.448</td>
</tr>
<tr>
<td>Trailer</td>
<td>2</td>
<td>1.379</td>
</tr>
<tr>
<td>Truck</td>
<td>2</td>
<td>1.379</td>
</tr>
<tr>
<td>Van</td>
<td>13</td>
<td>8.966</td>
</tr>
</tbody>
</table>

3-2 Vehicle feature extraction from Blade sensor signatures

Vehicle feature extraction is one of the major tasks for accomplishing vehicle reidentification because it seeks to extract salient components of vehicle images that would sufficiently differentiate vehicles. As mentioned in the previous section, blade loops are more sensitive than existing inductive loops, and are capable of capturing vehicle wheel locations in a signature. Use of vehicle wheel information is expected to improve the performance of vehicle reidentification. In this paper, we focus on developing a new method for vehicle signature feature extraction for blade sensors. Detailed information on feature extraction from conventional loops can be found elsewhere (3,8).

Figure 3 shows the feature extraction scheme for vehicles signatures produced by blade loops. Because a blade vehicle signature consists of two vehicle parts, namely, the wheel part and the vehicle body part, each part of a vehicle signature provides different vehicle features as shown in Figure 3. Figure 4 presents both conventional inductive loop vehicle signatures and blade loop vehicle signatures for different types of vehicles.
3.3 Vehicle feature analyses

This section focuses on the selection of vehicle features that will be used for vehicle reidentification. In this study, four vehicle types including passenger car, pickup truck, sport utility vehicle, and van were analyzed.

The feature selection seeks to select the salient features extracted from the vehicle signature that would sufficiently differentiate vehicle types. To select salient features, we used Bayes decision theory, which minimizes the probability of classification error for feature selection. As shown in Figure 5, the overlapping areas, $\Phi_j$, for the probability density functions for each vehicle type represent the probability that could be misclassified. Therefore, vehicle features showing the minimum overlapping area can be regarded as salient features that are capable of classifying vehicle types more effectively, and can be used for vehicle reidentification.

It was found that seven features were salient features based on Bayes decision theory. Those features are lane, vehicle length, maximum magnitude of inductance change, standard deviation for whole vehicle signature, shape parameter for whole vehicle signature, degree of symmetry for the body part of the signature, and standard deviation for the body part of the signature. Figure 6 shows the examples of the comparison of the probability density functions for vehicle features obtained from different vehicle types.
Figure 4 Conventional inductive loop vehicle signatures vs. blade loop vehicle signatures

Figure 5 Misclassification probabilities for hypothetical vehicle classification regions

where,  
\( f(x) \) : probability density function  
\( \Phi \) : misclassification area  
\( VT \) : vehicle type  
\( R \) : classification region
4. Genetically enhanced lexicographic optimization algorithm for vehicle reidentification

The vehicle reidentification problem with heterogeneous detection systems is much more challenging compared to the case of using homogeneous detection systems. It is because each detector system has unique characteristics for representing vehicle images, resulting from the different levels of a detection sensitivity. In order to develop a robust vehicle reidentification algorithm that can be successfully used with heterogeneous detector system, both a mapping procedure for input features and a genetic algorithm (GA) were incorporated into a lexicographic optimization based vehicle reidentification algorithm for enhancing the matching capability.

The lexicographic method is a sequential approach to solve the multi-objective optimization problem. The vehicle reidentification problem was formulated as a lexicographical optimization problem consisting of two main components. The first component has several layers to reduce the search space by eliminating upstream vehicle signatures that are unlikely to match a given downstream vehicle signature. The second
component computes discriminant scores to determine vehicle matching, which involves a multiple criteria decision-making process. The discriminant function of the second component has feature vectors as independent variables. More detailed algorithmic descriptions can be found in Sun et al. (4). The lexicographic optimization approach has the following benefits (12):

• multiple objectives can be addressed with different levels of priority
• sequential reduction of the feasible set from level to level enhances the computational efficiency
• sensitivity analysis can be conducted separately for each level

Search space reduction consists of four levels of optimization procedures with goal programs. The fundamental idea of goal programming is to establish a specific numeric goal for each objective and then search for a solution to minimize the weighted sum of deviations of objective functions from respective goals (12). The goal programs that can be used for search space reduction are described as follows.

\[
\text{goal for 'time window': } (g_f(x) = f(x) = z_1) \quad \text{such that } (z_1 \in [L_0, U_0],) \quad x \in S, \quad S^1 = \{x \in S : f(x) = z_1\}
\]

\[
\text{goal for 'lane': } (g_f(x) = \|d(x) - z_2\|) \quad \text{such that } (z_2 < T_1), \quad x \in S^1, \quad S^2 = \{x \in S^1 : f_f(x) = z_2\}
\]

\[
\text{goal for 'maximum magnitude': } (g_f(x) = \|d(x) - z_3\|) \quad \text{such that } (z_3 < T_2), \quad x \in S^2, \quad S^3 = \{x \in S^2 : f_f(x) = z_3\}
\]

\[
\text{goal for 'length': } (g_f(x) = \|d(x) - z_4\|) \quad \text{such that } (z_4 < T_3), \quad x \in S^3, \quad S^4 = \{x \in S^3 : f_f(x) = z_4\}
\]

where,

\[ x : \text{feature vector} \]

\[ f : \text{objective function} \]

\[ z : \text{objective value} \]

\[ t(x) : \text{travel time between upstream and downstream vehicle arrival times for individual vehicle} \]

\[ L_0, U_0 : \text{lower and upper bound for feasible travel time} \]

\[ S : \text{feasible set of vehicle pairs} \]

\[ T : \text{threshold value for feature vectors} \]

\[ d : \text{vehicle feature distance} \]

\[ l : \text{lane} \]

\[ m : \text{maximum magnitude} \]

\[ v : \text{vehicle length} \]

This process can generally continue until all objectives considered, although this study used four objectives. These first four optimization levels reduce the search space of similar vehicle signature pairs.

The fifth level lexicographic optimization objective can be described as follows:

\[
\min \quad f_5 = p_1|d(x) + p_2|d(x)|^2 + p_3|d(x) + p_4|d(x)| - \quad \text{s.t. } x \in S^4
\]

where,

\[ p : \text{set of coefficients associated with the feature vector differences} \]
Prior to applying lexicographic optimization for vehicle reidentification, input features should be adjusted since downstream vehicle features and upstream vehicle features are from different detection systems. Adjustment factors \((k_i, d_i)\) were employed for adjusting the feature differences between conventional inductive loop signatures and blade signatures. Therefore, the distance measure of vehicle feature \(i\) between an upstream loop vehicle feature \((v_{\text{loop}}^n)\) and downstream blade vehicle feature \((v_{\text{blade}}^n)\) is described by

\[
d_i(v_{\text{loop}}^n, v_{\text{blade}}^n) = \sum_{m=1}^{q} |v_{\text{loop}}^n(m) - k_i \cdot v_{\text{blade}}^n(m) + d_i|,
\]

where \(n\) denotes the \(n\)th element of the feature vector and \(q\) is the vector dimension.

In order to obtain an optimised set of parameters capable of maximizing vehicle reidentification performance, GA was applied. GA is an algorithm that searches the solution space of a function by emulating the mechanism of natural selection, that is, the survival of the fittest strategy. Optimization is performed on a set of strings, where each string is composed of a sequence of characters. Given an initial population of strings, a genetic algorithm produces a new population of strings according to a set of genetic rules. This constitutes one generation. The rules are devised so that the new generation tends to have strings that are superior to those in the previous generation. Successive generations of strings are produced, each of which tends to produce a superior population (13). The algorithms are not only robust but also simple, and do not require the assumption of knowledge of the search space. More detailed description of GA can be found in the literature (14).

GA was applied to solve the maximization problem for the vehicle reidentification system. The problem in this study was to maximize the Correct Matching Rate (CMR). The fitness function to be optimized by GA is the vehicle reidentification algorithm. A set of coefficients for feature vector differences \((P)\) that were used in computing discriminant scores were prepared by the GA optimizer. Output of the fitness function is the CMR. The maximization of CMR is defined as follows:

\[
\max \text{CMR}
\]

where, 

\[
\text{CMR} = \text{REID}(\Sigma)
\]

\[
\text{CMR} = \text{Correct Matching Rate}
\]

\[
\text{REID} = \text{vehicle reidentification algorithm}
\]

\[
\Sigma = \text{parameters to be optimized: coefficients for discriminant function}
\]

The steps of the GA performed in this study can be summarized as follows.

Step 1: Initialization
Step 2: Retrieval of fitness (CMR) from vehicle reidentification algorithm
Step 3: Selection process
Step 4: Crossover and Mutation
Step 5: Repeat Step 2-4

Figure 7 shows the framework for obtaining the optimal set of parameters by GA for the vehicle reidentification algorithm.

Figure 7 Framework for genetically enhanced lexicographic optimization algorithm for vehicle reidentification
5. Results

Performance measures for the vehicle reidentification algorithm evaluation included the total matching rate (TMR), the correct matching rate (CMR), the mismatching rate (MMR), and the matching reliability rate (MRR) were used. TMR is the total number of matched vehicles declared by the algorithm. CMR is the percentage of individual vehicles that the algorithm is able to match correctly. On the other hand, MMR is the percentage of individual vehicles that algorithm matches incorrectly. MRR is the ratio of CMR to TMR, and proportion of matched vehicles that are correctly matched. Table 2 summarizes the vehicle reidentification performance. As shown in Table 2, the CMR of the training data set was 41.43 %, while the CMR of the testing data set was 50.00%.

<table>
<thead>
<tr>
<th>Data</th>
<th>TMR: total matching rate</th>
<th>CMR: correct matching rate</th>
<th>MMR: mismatching rate (TMR-CMR)</th>
<th>MRR: reliability Rate (CMR/TMR)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>97.14 %</td>
<td>41.43 %</td>
<td>55.71 %</td>
<td>42.65 %</td>
</tr>
<tr>
<td>Testing</td>
<td>97.14 %</td>
<td>50.00 %</td>
<td>41.43 %</td>
<td>51.47 %</td>
</tr>
</tbody>
</table>

Sensitivity analysis on the effect of the time window (the first goal program in the vehicle reidentification algorithm) was performed in terms of travel times between the upstream and downstream stations. When a large time window is applied, the algorithm includes many upstream candidate vehicles resulting in increasing the matching possibility of the corresponding vehicle. The computational burden and mismatching possibility then increase simultaneously. On the other hand, the algorithm can find the corresponding vehicle efficiently with a small time window, but the corresponding vehicle could be excluded from the set of candidate vehicles. In addition, since arterial traffic flow is interrupted by signal control highly variable travel times result and the effect of the aggregation period on travel time accuracy needs to be investigated. Figure 8 shows the relationship among time window sizes, aggregation periods, and travel time accuracies. In this study, 9 aggregation periods were ranging from 2-minute to 10-minutes. In order to evaluate travel time accuracy, the mean absolute percentage error (MAPE) was calculated.

\[
MAPE = \frac{1}{N} \sum_{n=1}^{N} \left| \frac{TT_{obs,n} - TT_{est,n}}{TT_{obs,n}} \right| \times 100
\]

where,

- \(TT_{obs,n}\): Observed travel time at time step n (Ground truth)
- \(TT_{est,n}\): Estimated travel time at time step n (Reidentification algorithm)
- \(N\): Total number of time step
As shown in Figure 8, it is obvious that longer aggregation intervals yield smaller errors than those of shorter intervals. 112-second was identified as the best time window size to produce the highest travel time accuracy for most aggregation intervals. Less than 10% MAPE were achieved for 5-minute and longer aggregation periods. The shorter aggregation periods such as 2, 3, and 4-minutes were also able to produce less than 15% MAPEs when a 112-second time window was applied to derive travel times. Figure 9 shows comparisons of the estimated travel times obtained by the vehicle reidentification algorithm with the true travel times. It should be noted that results are quite encouraging despite the small size of the data set used.

The size of aggregation interval is an important issue for designing real-time traffic management and information strategies. As shown in the evaluation results, different aggregation intervals produce different levels of accuracies. In addition, shorter aggregation intervals have bigger travel time variations than those of the longer intervals. Therefore, the use of rolling averages of travel times on the time horizon would be a possible way to reduce the travel time variations. Identifying optimal travel time aggregation intervals for generating useful traffic information accounting for the real-time performance of transportation systems is an important issue in the field of traffic surveillance and information systems.
Figure 9 Comparison of travel times

6. Conclusion

This study explored the vehicle reidentification problem based on vehicle signatures collected from the different types of detection technologies, including conventional square inductive loops and newly developed blade inductive loop sensors.

A lexicographic optimization algorithm together with a genetic algorithm was introduced to solve the vehicle reidentification problem. Goal programming approaches for search space reduction in the vehicle reidentification algorithm both improved the algorithm matching performance and the computational burden. The algorithm performed well. For example, less than 10% travel time error was achieved with a 5-minute travel time aggregation period.

Although the number of vehicles for which data could be collected was small, encouraging results were obtained for vehicle reidentification performance in this system of mixed traffic detection technologies. In future large-scale applications of vehicle...
identification approaches for real-time traffic performance measurement, management and control, it would be most beneficial and practical if heterogeneous as well as homogeneous detection systems could be supported. This initial study yielded many useful insights about this important issue, and demonstrated on a small scale the feasibility of vehicle reidentification in a system with heterogeneous detection technologies.

Acknowledgement

This work was performed as part of the California Partners for Advanced Highways and Transit (PATH) program of the University of California, in cooperation with the Business, Transportation and Housing Agency, California Department of Transportation, and FHWA, U.S. Department of Transportation. The content of this paper reflect the view of the authors, who are responsible for the facts and the accuracy of the data presented. The contents do not necessarily reflect the official views or policies of the state of California. This paper does not constitute a standard, specification, or regulation. The authors gratefully acknowledge the collaboration and assistance of Steven Hilliard, Inductive Signature Technologies, Inc., in conducting this research, especially in providing the prototype blade sensors and portable data collection equipment used in this study.

References


