Title
Anonymous vehicle reidentification using heterogeneous detection systems

Permalink
https://escholarship.org/uc/item/0ps814kz

Authors
Oh, Cheol
Jeng, Shin-Ting
Ritchie, Stephen G.

Publication Date
2007-09-01

Peer reviewed
Anonymous Vehicle Reidentification Using Heterogeneous Detection Systems

Cheol Oh, Member, IEEE, Stephen G. Ritchie, and Shin-Ting Jeng

Abstract—An innovative feature of this paper is the demonstration of 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 detector cards) at both locations. In this paper, an opportunity arose for the first time to collect a downstream data set 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 and homogeneous detection systems could be supported. This initial paper 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.

Index Terms—Genetic algorithm (GA), lexicographic optimization, travel time estimation, vehicle feature, vehicle reidentification.

I. 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 various advanced transportation management and information systems to alleviate traffic congestion and its associated impacts.

A promising approach to obtain travel times is by tracking vehicles 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. The use of a global positioning system and in-vehicle tag-based automatic vehicle identification technologies has been successfully demonstrated to obtain accurate travel times. However, privacy issues still remain with such systems, and 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 reidentification systems using point-based detections 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]–[5], laser-based detection systems [6], [7] providing vehicle length, and video-based vehicle signature generation [8] using video image processing technology. As a further application of analyzing vehicle features, Sun et al. [9] proposed a method utilizing both inductive signatures and vehicle color information to derive travel times. They collected inductive signatures of vehicles that traversed two detector stations spanning a section of an arterial and the corresponding video of these signatures.

Previous studies performed by the authors [2]–[5] have proven that accurate travel times can be obtained from inductive signature-based vehicle reidentification using the 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 paper, an opportunity arose for the first time to collect a downstream data set from a temporary installation of a prototype innovative inductive loop sensor known as a “blade” sensor in conjunction with conventional inductive loops upstream. 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, for example, monitoring traffic conditions and supporting dynamic traffic management in work zones. However, it should be noted that the life cycle of blade sensors needs to be thoroughly investigated in the case of temporal installation, because various environmental conditions.
The following section introduces the blade sensor that is able to produce unique vehicle signatures. Data collection and vehicle feature extraction for the blade signature are presented in Section III. Section IV describes an algorithm for vehicle reidentification with the heterogeneous detection system used in this paper. An analysis of travel times using the outputs of the algorithm is then presented. Finally, conclusions including comments and findings are provided.

II. 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 [10]. A physical configuration of 6 ft × 6 ft (1.8 m × 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 ft × 6 ft (1.8 m × 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 contained within a single solid “sensor blade” that is then embedded in a 3/16-in-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 analog-to-digital converter.

A continuous stream of signed integers is generated by the blade sensor, which can be monitored by a dedicated onboard microprocessor. The resulting measurement data produce the vehicle’s inductive signature. Fig. 1 shows the temporary surface installation of blade sensors as deployed in this paper as well as 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 ft 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, number of axles, and wheel-based vehicle length. The prominent peaks shown in Fig. 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.
Fig. 2. Data collection layout for blade vehicle signature.

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.

III. 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.

A. Data Collection

In this paper, 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, CA, 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 40-GB 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. Fig. 2 shows the blade signature data collection layout.

One hour of data collected from 11:40 A.M. to 12:40 P.M. constituted the available data set for both conventional loop and blade loop data. In addition, 140 blade vehicle signatures collected in the rightmost lane of the downstream detector station and 858 upstream conventional loop signatures constituted the valid signature data set. We were not able to collect sufficient data due to hardware and power supply problems. The vehicle reidentification algorithm was developed and tested based on visual inspection identifying an upstream vehicle on a monitor and then searching to match 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.

B. Vehicle Feature Extraction From Blade Sensor Signatures

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. 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 [2], [4].

Each piece of information embedded in the blade signature is extracted and converted into individual feature vectors representing characteristics of individual vehicles, which are then used for the algorithm inputs. Moreover, vehicle lane information, vehicle speed, and vehicle arrival time are recorded for individual vehicles and are also considered as feature vectors. The raw vehicle signature is normalized by two procedures in order to eliminate the location variations and the effects of traffic conditions. First, the magnitude of vehicle signature, i.e., $y$-axis, is divided by its maximum magnitude (Max_mag). Second, the $x$-axis representing the time while the vehicle is present on the loop detectors is multiplied by the vehicle speed. The typical normalized blade signature obtained, as shown in Fig. 3, can subsequently be divided into two components: the upper region above a threshold value for vehicle body information analysis and the lower region for wheel information.
analysis. In vehicle body information analysis, measures such as skewness, kurtosis, degree of symmetry (DOS), shape parameter (SP), equally spaced interpolations, area, and standard deviation are used to determine the unique and intrinsic characteristics of each vehicle body. On the other hand, dimensional information is obtained from the wheel information analysis, which provides the total vehicle length, vehicle wheel base, and vehicle width. Fig. 4 presents both conventional inductive
loop vehicle signatures and blade loop vehicle signatures for different types of vehicles.

C. Vehicle Feature Analyses

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

As shown in Fig. 5, the overlapping areas \( \Phi_i \) 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. We selected salient features empirically based on comparing a priori densities of the features with the assumption that the features were normally distributed. Fig. 6 shows examples of the comparison of the probability density functions for vehicle features.

It was found that seven features were salient features based on the comparison of a priori densities. Those features are lane, vehicle length, maximum magnitude of inductance change, standard deviation for whole vehicle signature, SP for the whole vehicle signature, DOS for the body part of the signature, and standard deviation for the body part of the signature.

IV. 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 because each detector system has unique characteristics for representing vehicle images, which results from the different levels of detection sensitivity. In order to develop a robust vehicle reidentification algorithm that can be successfully used with the 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 to enhance the matching capability.

Prior to applying lexicographic optimization for vehicle reidentification, input features should be adjusted since the downstream and upstream vehicle features are from different detection systems. Adjustment factors \((k_i, l_i)\) were employed to adjust the feature differences between conventional inductive loop signatures and blade signatures. Adjustment factors were estimated by the ordinary least squares method. Therefore, the distance measure of vehicle feature \(i\) between an upstream loop vehicle feature \((vf_{\text{up}}^f)\) and a downstream blade vehicle feature \((vf_{\text{bl}}^f)\) is described by

\[
d_i(vf_{\text{up}}^f, vf_{\text{bl}}^f) = \sum_{n=1}^{q} |vf_{\text{up}}^f(n) - (k_i \times vf_{\text{bl}}^f(n) + l_i) |
\]

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

In this application, adjustment factors for three vehicle features, including vehicle length, maximum magnitude, and SP, were estimated by a simple linear regression using the ordinary least square method. The estimated adjustment factors are presented in Table I. The vehicle length that presented the highest value of \(R^2\) was identified as the most similar vehicle features between two detector technologies.

The lexicographic method is a sequential approach to solve the multiobjective 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. The lexicographic optimization approach has the following benefits [11].

- 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 [11]. The goal programs that can be used for search space reduction are described as follows:

- goal for “time window”: \(f_1(x) = t(x) = \text{z}_1\) such that \((z_1 \in [L_1, U_1]), x \in S, S^1 = \{x \in S : f_1(x) = \text{z}_1\}\)
- goal for “lane”: \(f_2(x) = d_l(x) = \text{z}_2\) such that \((z_2 < T_1), x \in S^1, S^2 = \{x \in S^1 : f_2(x) = \text{z}_2\}\)
- goal for “maximum magnitude”: \(f_3(x) = |d_m(x)| = \text{z}_3\) such that \((z_3 < T_m), x \in S^2, S^3 = \{x \in S^2 : f_3(x) = \text{z}_3\}\)
- goal for “length”: \(f_4(x) = |d_l(x)| = \text{z}_4\) such that \((z_4 < T_{vl}), x \in S^3, S^4 = \{x \in S^3 : f_4(x) = \text{z}_4\}\)

...
• STD: Standard Deviation, SP: Shape Parameter, DOS: Degree of Symmetry
• PC: Passenger Car, PU: Pickup Truck, VAN: Van, SUV: Sport Utility Vehicle

Fig. 6. Vehicle feature distribution analyses.

<table>
<thead>
<tr>
<th>Table I</th>
<th>Estimated Adjustment Factors</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Length</td>
</tr>
<tr>
<td>K</td>
<td>0.7503</td>
</tr>
<tr>
<td>L</td>
<td>1.872</td>
</tr>
<tr>
<td>R²</td>
<td>0.8207</td>
</tr>
</tbody>
</table>

where
- $x$ is the feature vector;
- $f$ is the objective function;
- $z$ is the objective value;
- $t(x)$ is the travel time between upstream and downstream vehicle arrival times for individual vehicle;
- $L_t-U_t$ is the lower and upper bound for feasible travel time;
- $S$ is the feasible set of vehicle pairs;
- $T$ is the threshold value for feature vectors;
- $d$ is the vehicle feature distance;
- $l$ is the lane;
- $m$ is the maximum magnitude;
- $vl$ is the vehicle length.

This process can generally continue until all objectives are considered, although this paper 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

$$\min f_5 = p_a |d_a(x)| + p_b |d_b(x)| + p_c |d_c(x)| + p_d |d_d(x)| \cdots s.t. x \in S^4$$

where $p$ is the set of coefficients associated with the feature vector differences.

In this application, three vehicle features, including vehicle length, maximum magnitude, and SP, were used for the fifth-level lexicographic objective.

V. Parameter Calibration With GA

In order to obtain an optimal 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 GA 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 [12]. The algorithms are not only robust but also simple and do not require the assumption of knowledge of the search space. A more detailed description of the GA can be found in the literature [13].

GA was applied to solve the maximization problem for the vehicle reidentification system. The problem in this paper was to maximize the correct matching rate (CMR). The fitness function to be optimized by the GA is the vehicle reidentification algorithm. A set of coefficients for feature vector differences \( \bar{P} \) that were used in computing discriminant scores were prepared by the GA optimizer. The output of the fitness function is the CMR. The maximization of CMR is defined as

\[
\max \sum \text{CMR} \{ \text{REID}(\Sigma) \}
\]  

(2)

where CMR is the correct matching rate, REID is the vehicle reidentification algorithm, and \( \sum \) are the parameters to be optimized (coefficients for discriminant function).

The steps of the GA performed in this paper can be summarized as follows:

1) Initialization.
2) Retrieval of fitness (CMR) from the vehicle reidentification algorithm.
3) Selection process.
4) Crossover and mutation.
5) Repeat Steps 2)–4).

Table II shows the framework for obtaining the optimal set of parameters by GA for the vehicle reidentification algorithm.

<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 : Matching 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>

VI. RESULTS

Performance measures for vehicle reidentification algorithm evaluation included the total matching rate (TMR), CMR, mismatching rate (MMR), and matching reliability rate (MRR). TMR is the percentage of 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 the algorithm matches incorrectly. MRR is the ratio of CMR to TMR and the proportion of matched vehicles that are correctly matched. Table I summarizes the vehicle reidentification performance. As shown in Table II, the CMR of the training data set was 41.43%, while the CMR of the testing data set was 50.00%.
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 the signal control, highly variable travel times result, and the effect of the aggregation period on travel time accuracy needs to be investigated. Fig. 8 shows the relationship among time window sizes, aggregation periods, and travel time accuracies. In this paper, nine aggregation periods were ranging from 2 to 10 min. In order to evaluate the travel time accuracy, the mean absolute percentage error (MAPE) was calculated as

$$\text{MAPE} = \frac{1}{N} \sum_{n=1}^{N} \left\{ \frac{|\text{Time}_{\text{obs}},n - \text{Time}_{\text{est}},n|}{\text{Time}_{\text{obs}},n} \times 100 \right\}$$  \hspace{1cm} (3)

where

- $\text{Time}_{\text{obs}},n$ observed travel time at time step $n$ (ground truth);
- $\text{Time}_{\text{est}},n$ estimated travel time at time step $n$ (reidentification algorithm);
- $N$ total number of time steps.

As shown in Fig. 8, it is obvious that longer aggregation intervals yield smaller errors than those of shorter intervals. The best time window size to produce the highest travel time accuracy for most aggregation intervals was identified as 112 s. Less than 10% MAPE was achieved for 5 min and longer aggregation periods. The shorter aggregation periods such as 2, 3, and 4 min were also able to produce less than 15% MAPEs when a 112-s time window was applied to derive travel times. Fig. 9 shows comparisons of the estimated travel times obtained by the vehicle reidentification algorithm with the true travel times. It should be noted that the 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 accuracy. In addition, shorter aggregation intervals have bigger travel time variations than those of 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.

VII. CONCLUSION

This paper explored the vehicle reidentification problem based on vehicle signatures collected from different types of detection technologies, including conventional square inductive loops and newly developed blade inductive loop sensors. A lexicographic optimization algorithm together with a GA was introduced to solve the vehicle reidentification problem. Goal programming approaches for search space reduction in the vehicle reidentification algorithm improved both 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-min 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 reidentification approaches for real-time traffic performance measurement, management, and control, it would be most beneficial and practical if heterogeneous and homogeneous detection systems could be supported. This initial paper 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.

ACKNOWLEDGMENT

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,
Fig. 9. Comparison of travel times. (a) 2-min. aggregation. (b) 3-min. aggregation. (c) 4-min. aggregation. (d) 5-min. aggregation.

REFERENCES


Cheol Oh (M’04) received the Ph.D. degree in civil engineering—transportation systems engineering from the University of California, Irvine.

He is currently an Assistant Professor with the Department of Transportation Systems Engineering, Hanyang University at Ansan, Kyunggi-do, Korea. His research interests include traffic surveillance systems, traffic data filtering and processing, travel time estimation and prediction, and data fusion.
Stephen G. Ritchie received the Ph.D. degree in civil engineering—transportation from Cornell University, Ithaca, NY.

He is currently a Professor with the University of California, Irvine. He studies transportation systems engineering and is specifically interested in advanced traffic management and control systems. He is primarily focused on the development and application of new information technologies to support more efficient transportation systems.

Shin-Ting (Cindy) Jeng received the M.S. degree in civil engineering with a major in transportation engineering from the National University of Singapore, Singapore. She is currently working toward the Ph.D. degree in transportation system engineering at the University of California, Irvine.

Her research interests are in traffic surveillance systems, traffic simulation, and incident management.