Geographic Information System Estimation Using Inter-Vehicle Communication

Keith Yu Kit Leung,	MASc Candidate	Email: kykleung@lair.uwaterloo.ca
Thanh-Son Dao,	PhD Candidate	Email: tsdao@engmail.uwaterloo.ca
Christopher Clark,	Assistant Professor	Email: cclark@mecheng1.uwaterloo.ca
Jan Huissoon,	Professor	Email: jph@uwaterloo.ca

Lab for Autonomous and Intelligent Robotics
Department of Mechanical Engineering
University of Waterloo
200 University Avenue West
Waterloo, ON, Canada, N2L 3G1

I. Introduction

Many automotive manufacturers currently offer vehicles equipped with a GPS receiver coupled to a geographic information system (GIS) that serves as a digital road map. A vehicle can use this database of roads and highways together with GPS position readings to localize itself. In the ideal case, the GPS position reading for a vehicle will fall on a certain road within the GIS database. However, GPS readings include bias and error caused by various factors such as atmospheric conditions, the location of GPS satellites, multi-path, etc... This error is generally magnified with lower cost receivers, and due to this bias and error, the reported location of a vehicle may not even fall on a roadway. Recent research efforts have improved this map matching process to provide greater accuracy in vehicle position reported by a GIS through methods such as taking the velocity vector of a vehicle into account in the map matching process. Advanced techniques include the use of fuzzy logic in the map matching routine to estimate vehicle position [1][2]. Nevertheless, the localization system as a whole depends on the availability of a GIS database of the local area with respect to the vehicle. The objective of this paper is to present a method for vehicles to collaboratively estimate the position and orientation of roadways by using global positioning system (GPS) readings from low cost receivers and exchange information via inter-vehicle communication (IVC). The proposed method of estimating road position and orientation will serve as a complementary system to existing GIS databases. The method will allow information to be compressed as road sections will be modeled as Bezier splines. Furthermore, it may be possible in the future for roadways not yet available within a GIS database to be appended using the information generated from the proposed method. At the moment, this method is implemented in two dimensions so the elevation of the road is not being estimated. Simulation results of GIS estimation for a highway located in Waterloo, Ontario, Canada using GPS data will be used to support this. The highway is modeled in a microscopic traffic simulator designed with the capability of handling inter-vehicle communication.

II. Estimation Method

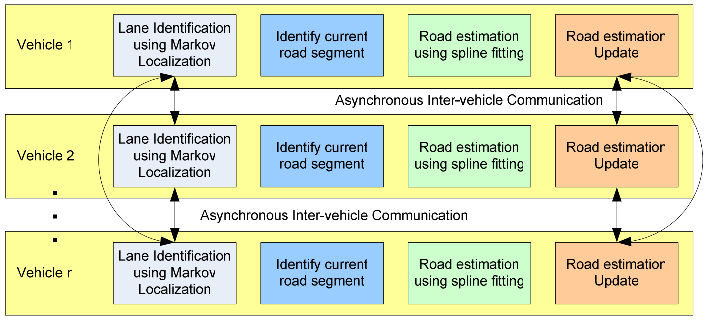

Fig 1: Summary of the road estimation process

The proposed GIS estimation method not only depends on individual vehicles to estimate that position of a road, but it also relies on the availability of inter-vehicle communication to collaboratively combine estimations from multiple vehicles. As summarized in figure 1 above, the GIS estimation method consists of the following steps: 1) lane identification using *Markov localization*, 2) road segment identification, 3) road segment estimation by individual vehicles, and 4) collaborative estimation.

Individual vehicles are first required to localize themselves and determine their lateral position, or lane. This step involves the use of inter-vehicle communication so that a vehicle is aware of the position of its neighbouring vehicles. Lane position is used to determine the offset to the centerline of the road that is being estimated. All vehicles also need to determine the local road section they are currently driving on. To ensure consistency in road segment identification, an inertial reference frame is used to define a square grid of width w that vehicles will use to associate a road segment with as shown in figure 2. The inertial reference frame can be derived from a global coordinate system such as the Earth centered Earth fixed (ECEF) frame in which geodetic coordinates provided by the GPS receivers are given. An identification number is assigned to each grid cell as a mean of reference. Figure 3 shows how road segments are determined using the grid, the process of which will be explained in detail subsequently. When a vehicle has traveled a road segment, it will use GPS coordinates which it has been recording to calculate a curve which best fits its trajectory. Through the use of inter-vehicle communication again, a vehicle will receive a road segment estimation that was generated collaboratively by vehicles that have previously traveled through. The individual vehicle will combine their traveled trajectory with the collaborative estimation to update it and communicate the latest estimate of the road segment to other vehicles.

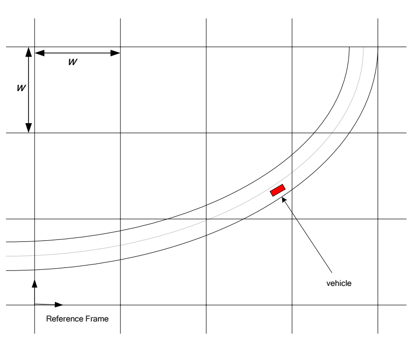

Fig 2: Descretizing the local area into grid cells

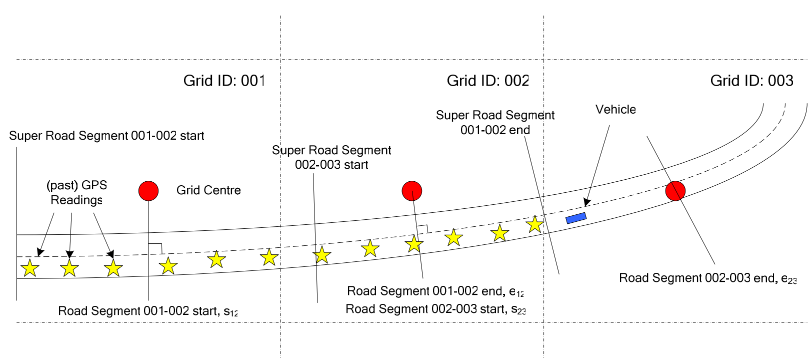

Figure 3: Determining road sections using the defined grid

1) *Markov Localization for Lane Estimation using Inter-vehicle Communication*

To begin the road estimation process, a vehicle is must identify the lane it is traveling in as this indicates the position offset to the centre of a road. Ultimately, an estimation of a road section will be defined by a spline that represents the centre of the roadway. A road width of 3.5 m is used for testing of the proposed road estimation algorithm as it corresponds to the lane width of a local highway in Waterloo, Ontario, Canada. Part of the future work of the lane identification process will be estimating the width of a lane.

Traditionally in the field of robotics, Markov localization is used to track the belief state of a vehicle by using a discretized probability density function that indicates the likelihood of being in all possible states [3]. This method often demands the number of descretized states to be extremely large, making computation impractical. However, for the case of the predicting the occupied lane for a vehicle on a roadway, the number of possible states is drastically reduced and bounded by the number of lanes on a road, making Markov localization practical.

Similar to other methods of localization, the two main steps involved in Markov localization are prediction and correction. In the prediction step, the probability of a vehicle being estimated in a particular lane can be calculated using prior knowledge of the vehicle's lane and the probability of a lane changing action. For vehicle *i* traveling in lane *j*, the confidence in the prior estimated lane position can be expressed mathematically as

$$P_{t-1}(v_i = l_j) \tag{1}$$

Given the lateral displacement Δd, the probability of a lane changing action (which includes the case of not changing lanes) can be approximate with a probability mass function. Δd is a function of previously perceived vehicle heading θ, vehicle position *x, y*, and the current vehicle position obtained using GPS data *x', y'*. For a vehicle *i*, the probability of a motion that carries it from lane j to k can be written as

$$P_{t-1}(v_i = l_k \mid v_i = l_j) = f(\theta, x, y, x', y') \tag{2}$$

From this, conditional probability allows the chance for a vehicle to be in a certain lane (*k*) to be quantified. For a two lane roadway ($j \in \{1,2\}$), this probability can be expressed as.

$$P_t(v_i = l_k) = \sum_{j=1}^{2} P_{t-1}(v_i = l_k \mid v_i = l_j) P_{t-1}(v_i = l_j) \tag{3}$$

The correction step involves using the latest GPS reading to make adjustments to the previous state estimation. For a pair of vehicles, the probability of measuring a certain lateral distance z between them (calculated using newly obtained GPS measurements), given the estimated position the vehicles can be expressed as a probability density function

$$P(z \mid v_1 = l_a, v_2 = l_b) \tag{4}$$

The function above is modeled from specifications of the GPS receiver and field test data corresponding to the receiver's performance.

Determining the lateral position difference z requires the use of inter-vehicle communication to share GPS readings. The probability of the pair of vehicles being in some specific lane configurations can be corrected using Bayes' Rule.

$$P_t(v_1 = l_a, v_2 = l_b \mid z) = \frac{P_t(z \mid v_1 = l_a, v_2 = l_b)\,\hat{P}_t(v_1 = l_a, v_2 = l_b)}{P_t(z)} \tag{5}$$

where $\hat{P}_t\left(v_1 = l_a, v_2 = l_b\right) = P_t\left(v_1 = l_a\right) P_t\left(v_2 = l_b\right)$.

The denominator in the above equation is there to ensure that the probability of all possible lane positions for all vehicles sum to one. This Markov-based algorithm can be extended for use with multiple vehicles and has been proven to work using simulations and real GPS data collected by multiple vehicles on a highway.

2) Road Segment Identification

Estimation of road is performed in a piecewise manner. Every vehicle needs to determine the road segment it is traveling on so that information can be collaborated with other vehicles later.

With reference to figure 3, a road segment S_{ij} contains a starting point s_{ij} and an ending point e_{ij} which serve as boundaries of an estimation. The subscripts in the nomenclature refer to the identification numbers of the grids in which the road estimation spans across. The starting or ending point at a given segment is defined as the closest point of the road estimate to the centre of a grid cell. From the figure, it shows that the starting and ending points for a particular road sector can not reside in the same grid sector. Also, the start and end points for a particular road section should only be in adjacent grid sectors. Continuity demands that the starting point or one road section should correspond to the ending point of another road section.

A vehicle can use its GPS readings to determine when it has entered or exited a road segment the distance to its current grid centre will have reached a minimum. In addition to a road section, a 'super' road section is defined as an extension to a road section where the start and end points are extended to overlap the adjacent road sections. Road position and orientation estimation is actually performed on the super road section as this will provide better continuity in the transition between road section estimates. Handling of intersecting roadways is part of the future work planned for the proposed estimation algorithm.

3) *Road Segment Estimation for Individual Vehicles*

A vehicle that has traveled through a super road section will have recorded positions along its trajectory from its GPS receiver. These data points can then be used for fitting a curve that will represent the position and orientation of a road segment. A generalized Bézier curve, or B-spline is fit in a least square sense to a set of GPS coordinates belonging to a super road section. In accordance with [4], a B-spline curve is defined for a collection of $n+1$ control points $\{Q_i\}_{i=0}^{n}$ by

$$\mathbf{X}(t) = \sum_{i=0}^{n} N_{i,d}(t)\mathbf{Q}_i \tag{6}$$

The control points can be any dimension so long as they are the same for each point. The degree of the curve is d and must satisfy $1 \le d \le n$. The functions $N_{i,d}(t)$ are the B-spline basis function, which are defined recursively and require selection of a sequence of scalars t_i for $0 \le i \le n+d+1$. The sequence is nondecreasing; that is, $t_i \le t_{i+1}$. Each t_i is referred to as a knot and the total sequence of knots composes the knot vector. The basis function that starts the recursive definition is

$$N_{i,0}(t) = \begin{cases} 1, & t_i \le t \le t_{i+1} \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

for $0 \le i \le n+d$. The recursion can be expressed in the following form.

$$N_{i,j}(t) = \frac{t-t_i}{t_{i+j}-t_i} N_{i,j-1}(t) + \frac{t_{i+j+1}-t}{t_{i+j+1}-t_{i+1}} N_{i+1,j-1}(t) \qquad \text{for } 1 \le j \le n+d \text{ and } 0 \le i \le n+d-j \tag{8}$$

The support of a function is the smallest closed interval on which the function has at least one nonzero value. The support of $N_{i,0}(t)$ is clearly $[t_i, t_{i+1}]$. In general, the support of $N_{i,j}(t)$ is $[t_i, t_{i+j+1}]$. This fact means that locally the curve is influenced by only a small number of control points.

The main classification of the knot vector is that it is either open or periodic. If open, the knots are either uniform or nonuniform. Periodic knot vectors have uniformly spaced knots. Uniform knots are

$$t_i = \begin{cases} 0, & 0 \le i \le d \\ \dfrac{i-d}{n+1-d}, & d+1 \le i \le n \\ 1, & n+1 \le i \le n+d+1 \end{cases} \tag{9}$$

Periodic knots are

$$t_i = \frac{i-d}{n+1-d}, \quad n+1 \le i \le n+d+1 \tag{10}$$

The method of least squares is used to fit a spline to a set of data points $\{(s_k, \mathbf{P}_k)\}_{k=0}^{m}$, where s_k are times and $\mathbf{P}_k$ are position data. The sample times are assumed to be increasing: $s_0 < s_1 < ... < s_m$. A B-spline curve that fits the data is parameterized by $t \in [0,1]$, so the sample times need to be mapped to the parameter domain by $t_k = (s_k - s_0)/(s_m - s_0)$.

The control points $\mathbf{Q}_i$ of the spline are unknown quantities that need to be determined. The control points are considered to be column vectors, and collection of control points may be arranged into a single column vector

$$\hat{Q} = [\boldsymbol{Q_0} \quad \boldsymbol{Q_1} \quad \dots \quad \boldsymbol{Q_n}]^T \tag{11}$$

Similarly, the samples $\mathbf{P}_k$ are considered to be column vectors, and the collection written as a single column vector

$$\hat{P} = [\boldsymbol{P_0} \quad \boldsymbol{P_1} \quad \dots \quad \boldsymbol{P_n}]^T \tag{12}$$

For a specified set of control points, the least-squares error function between the B-spline curve and the sample points is the scalar-valued function

$$E(\hat{Q}) = \frac{1}{2}\sum_{k=0}^{m}\left|\sum_{j=0}^{n} N_{j,d}(t_k)\mathbf{Q}_j - \mathbf{P}_k\right|^2 . \tag{13}$$

The quantity $\sum_{j=0}^{n} N_{j,d}(t_k)\mathbf{Q}_j$ is the points on the spline curve at the scaled sample time t_k. The term within the summation on the right-hand side of equation measures the squared distance between the sample point and its corresponding curve point. The error function measures the total accumulation of squared distances. The method of least squares should define control points that minimize the error by equating the derivatives $\frac{\partial E}{\partial \mathbf{Q}_i}$ to zero. For the proposed road estimation method, five control points are used to define a spline in each road section.

4) *Collaboration Road Segment Estimation*

After a vehicle has made its own estimation of a road segment, it will combine its estimate with the collaborative estimate made by vehicles that have previous traveled the road segment. The updated collaborative estimation will then be passed onto other vehicles through inter-vehicle communicataion. If there is no history of previous estimates for a road section, a vehicle will use its own estimation and inform other vehicles that its spline fit is currently the best estimate of the road.

The combination of a new estimate from a vehicle and the collaborative estimate passed on by previous vehicles is determined by a weighting scheme that depends on the number of contributions made to the collaborative estimate. An equal number of coordinates are sampled from the individual and the collaborative estimates or splines, which will be denoted by the vectors $\boldsymbol{P}_t$ and $\boldsymbol{P}_{t-1}$ respectively. These coordinates will be combined to form a new vector $\boldsymbol{P}$ using the weighting scheme

$$\boldsymbol{P} = \lambda \boldsymbol{P}_t + (1-\lambda)\boldsymbol{P}_{t-1} \tag{14}$$

Such that,

$$\lambda = \max\left(\frac{1}{n+1}, \lambda_{\min}\right)$$

The weight λ is dependent on the number contributions that have been made to the collaborative road estimate. However, it is important to limit the number of contributions to (or in other words the history of the collaborative estimation) because the representation of the road should change as GPS error and bias evolves. Further work is required to determine an optimal contribution limit λ_{min}.

The B-spline fitting routine presented in the road segment estimation step is now performed on the coordinates in vector $\boldsymbol{P}$. Information corresponding to this new updated estimate is then be communicated to other vehicles.

III. Simulation Platform

The simulator used for testing the developed road estimation concepts is a microscopic traffic simulator that is capable of handling vehicle communication. A microscopic traffic simulator is one that models individual vehicles as particles governed by a vehicle model. The simulator used for testing the road estimation concept is specifically developed to assist in inter-vehicle communication related research. The underlying simulator engine and program environment and interface is based on the commercial microscopic traffic simulator package Vissim. An external driver model dynamic link library (dll) file contains functions which serve as an interface between the simulator engine and the components designed for handling inter-vehicle communication. Vissim itself already includes driver models that govern the behaviour of each vehicle which can be overrode through the external driver model. Additionally, Vissim includes the infrastructure for generating a traffic network for simulation.

A satellite image of highway 85 in Waterloo, Ontario, Canada as shown in figure 4 is used as a template for creating the traffic network that is used for testing the road estimation algorithm. The sections labeled in the figure correspond to the figures that will be presented in the next section. The dimension of a square grid used for segment identification has been set to 100m for simulation of the proposed estimation method but fine tuning of this value is expected in the future for improved performance when road density increases.

GPS readings are generated for each vehicle based on the coordinate of the vehicle in the simulation. Error is modeled with a Gaussian distribution using a standard deviation of 3m. A slow random bias drift is also present that can cause readings to deviate from the true position by as much as 5m. GPS readings are shared amongst vehicles within communication range as determined by the simulator. Results from the simulator are logged to file.

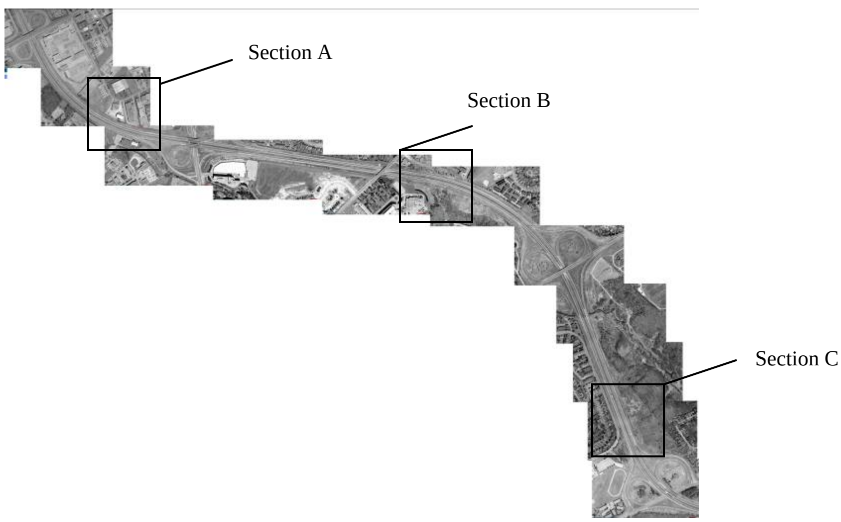

Fig 4. Satellite imaged used for constructing the simulation traffic network and sections specifically examined

IV. Simulation Results

Results demonstrate the feasibility of the proposed algorithm. The highway shown in figure 4 was set to have a traffic flow rate of 2000 vehicles per hour in the simulator. Since the simulated highway spans over several kilometers, the road estimations for selected sections will be shown to enable a more magnified scale. Figures 5 through 7 show the collaborative road estimations corresponding to section A of figure 4 for different number of contributions. The solid lines in these figures represent the true road location, while the dotted lines are the spline fit estimations constructed from the control nodes that are shown as solid dots. Figure 8 through 10 shows the change in average error of the collaborative road estimation as the number of vehicle contributions increases for section A, B and C of figure 4 respectively. The results consistently show the decrease in error of the collaborative road estimation as the number of vehicle contributions increases.

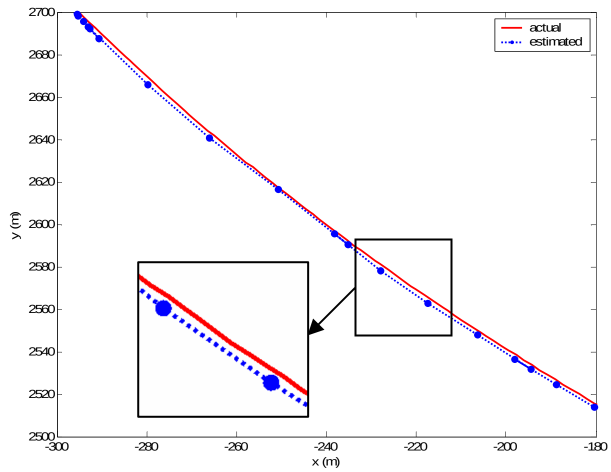

Fig 5. Collaborative estimation with 1 contribution, section A

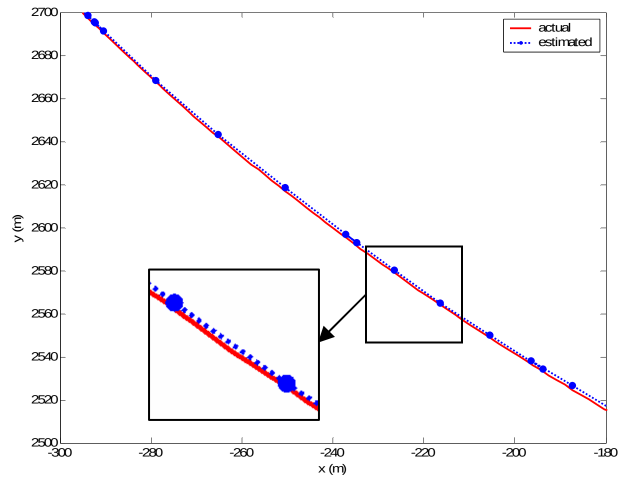

Fig 6. Collaborative estimation with 2 contributions, section A

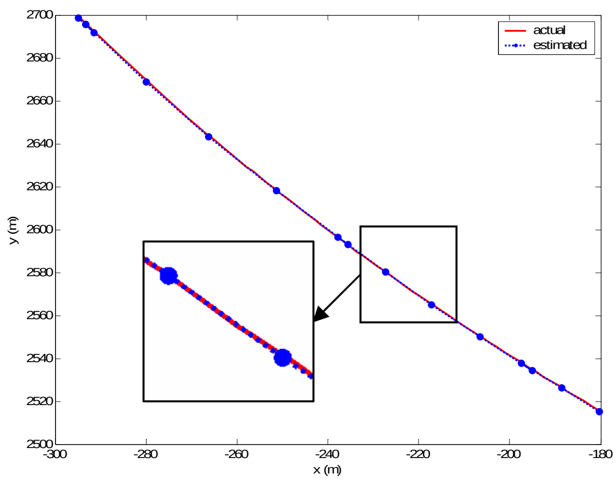

Fig 7. Collaborative estimation with 10 contributions, section A

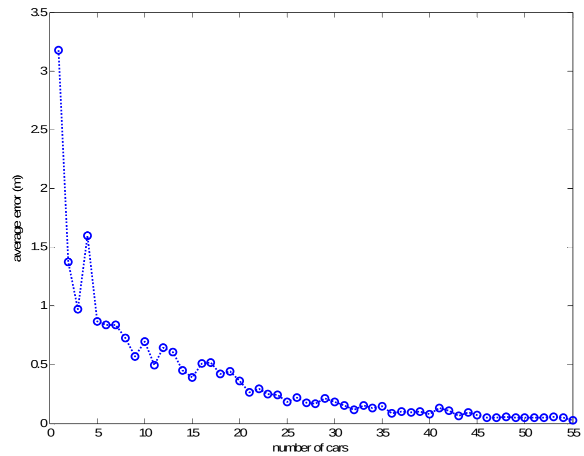

Fig 8. Average error as the number of contributions increases, section A

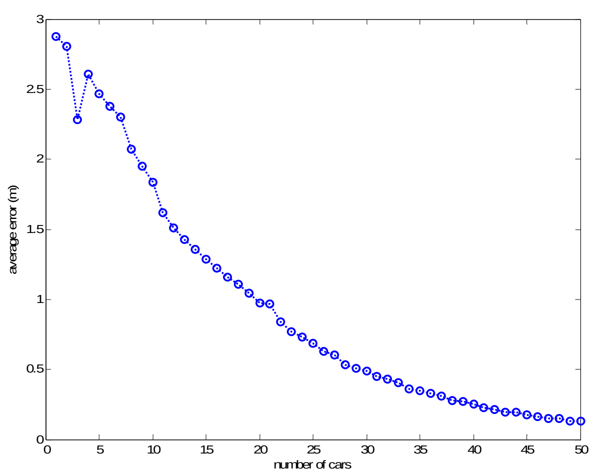

Fig 9. Average error as the number of contributions increases, section B

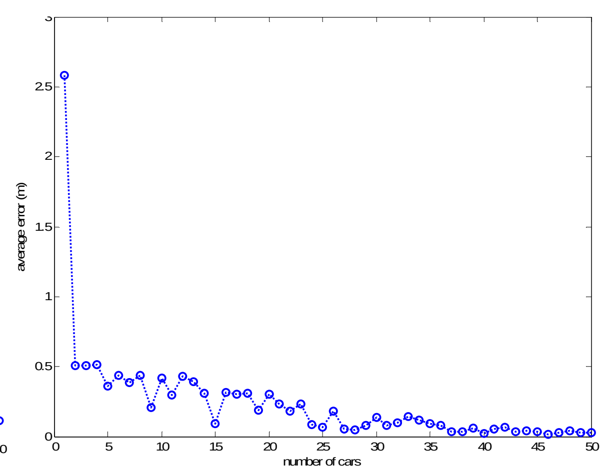

Fig 10. Average error as the number of contributions increases, section C

V. Conclusions

A method of road estimation using low cost GPS receivers and inter-vehicle communication was presented in this paper. The method was programmed into a microscopic traffic simulator for testing. Simulation results show that the proposed algorithm works well in estimating the position of a highway and estimation error always decreases with increasing contributions to the collaborative estimation. Using characteristics of a low cost GPS for the simulation that has a standard deviation of 3m and a bias drift of 5m, collaborative estimation is able to reduce the average error to a much smaller distance. The ability to achieve these results with reading characteristics of a low cost GPS receiver implies that it may be possible to implement the proposed method in a large scale in the future. At the moment additional development is required to cope with more complex traffic networks. The dependency of the road estimation algorithm on inter-vehicle communication gives an additional incentive for continual development of the technology and provides another benefit of collaborative driving in addition to safety improvement of the roads.

VI. Future Work

Future work will explore the effectiveness of the proposed road estimation algorithms in more complex traffic networks. Currently, the algorithm has been shown to work well for a freeway. However, modifications are expected for the proposed method to work on traffic networks where roads intersect each other or where roads are closely spaced. Optimization of the grid sizes for road segmentation will also be investigated.

The simulator is able to generate GPS data using a model based upon data collected from a GPS receiver, but field testing of the proposed algorithm is expected in the future to give stronger evidence that the estimation method is feasible. Since it is impractical and costly to gather a fleet of vehicles and equip each with a GPS receiver, only a couple of vehicles will be used. The vehicles will travel on a specific section of a highway numerous times to collect GPS data, which will then be divided appropriate offline to represent data from multiple vehicles. The road estimation will be applied on these data sets.

VII. Acknowledgement

The research presented in this paper is a part of the Dynamic Collaborative Driving project of AUTO21, an automotive research initiative funded by the government of Canada.

VIII. References

[1] Syed, S. and M.E. Cannon, Fuzzy Logic Based-Map Matching Algorithm for Vehicle Navigation System in Urban Canyons, Proceedings of the ION National Technical Meeting, Long Beach, 2004.

[2] Basnayake, C., O. Mezentsev, G. Lachapelle, M.E. Cannon, A Portable Vehicular Navigation System Using High Sensitivity GPS Augmented with Inertial Sensors and Map-matching. Paper 04CONG-98, CD-ROM Proceedings of 2004 Society of Automotive Engineers, Detroit, March 2004.

[3] R. Siegwart, I. R. Nourbaksh, Introduction to Autonomous Mobile Robots, Bradford Book, England, 2004.

[4] G. Wahba,. Spline Models for Observational Data, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 59, Philadelphia, 1990.