

PROBABILISTIC ESTIMATION OF LANES USING VEHICLE GPS TRAJECTORIES

E.R.I.A.C. Manel

(108006 M)



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
Degree of Master of Philosophy
www.195.mrt.ac.lk

Department of Electronic and Telecommunication Engineering

University of Moratuwa

Sri Lanka

November 2013

PROBABILISTIC ESTIMATION OF LANES USING VEHICLE GPS TRAJECTORIES

Epa Ranasinghe Imia Appuhamilage Champika Manel

(108006 M)

Dissertation submitted in partial fulfillment of the requirements for the degree of



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Department of Electronic and Telecommunication Engineering

University of Moratuwa

Sri Lanka

November 2013

Declaration

I, declare that this is my own work and this dissertation does not incorporate without acknowledgement any material previously submitted for a Degree or Diploma in any other University or institute of higher learning and to the best of my knowledge and belief it does not contain any material previously published or written by another person except where acknowledgement is made in the text.

Also, I hereby grant to University of Moratuwa the non-exclusive right to reproduce and distribute my dissertation, in whole or in part in print, electronic or other medium. I retain the right to use this content in whole or part in future work (such as articles or books)

Signature:

Date:

E.R.I.A.C. Manel



The above candidate has carried out research for the MPhil dissertation under our supervision.

University of Moratuwa, Sri Lanka.

Electronic Theses & Dissertations

www.lib.mrt.ac.lk

Signature of supervisor 1:

Date:

Professor S.A.D. Dias

Signature of supervisor 2

Date:

Dr. A. S. Perera

Abstract

This dissertation presents a novel process for generating lane-level information for road maps using a collection of trajectories from vehicles travelling on the road. The information thus generated will aid a range of Intelligent Transportation System (ITS) applications.

Recently, there has been a surge of interest in research in the arena of Intelligent Transportation Systems (ITS). These systems are expected to ensure driver/passenger safety, assist the driver, support green concepts and to improve the overall efficiency and the performance of transportation systems. However, ITS require more information than what current road maps provide. Generating and refining road maps having such level of detail using the existing methods such as surveying and digitization are time consuming, costly and incompatible with the real-time and the dynamic nature of the road network. Therefore, finding new ways of generating this additional information is of high importance in making ITS a reality.

The new method we propose generates lane-level information such as lane centerlines, boundaries and lane-width using vehicle trajectory data. This is achieved by modeling the Probability Density Function (PDF) of trajectories across the road using the non-parametric Kernel Density Estimation (KDE). Unlike the existing methods that use Differential GPS (DGPS) data or improved GPS data, the proposed method uses ordinary GPS data obtained from vehicles moving along the road. It does not require any information regarding the road parameters and is completely automatic. Furthermore, it is completely independent of the lane/road width and does not use stringent assumptions on lane parallelism and constant lane width. In particular, it estimates the locations of lane centers, locations of lane boundaries and lane width. The proposed method for calculating the lane centers was proven to be successful in different road geometries such as straight sections, curved sections and sections with lane splits and merges. The method proposed for calculating lane boundaries produced good results when there are no gaps in between lanes. The lane width calculated using the proposed method is compatible with the recorded standard lane width of the chosen road.

Keywords: GPS, Lane level maps, Kernel density estimation, Kernel bandwidth



To My Parents
University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

Acknowledgement

I extend my sincere appreciation, first to the University of California, Berkley for providing the Mobile Century Dataset that served as a key component from the inception of the project. I equally thank the Staff Development Centre, University of Moratuwa for funding for the two years of my MPhil. Professor S.A.S Kulathilake, director, Staff Development Centre deserves a big thank you for his kind support extended to me with the funding application. I am also grateful to Dr. J.C Balasooria, former Head, Department of Information Technology, University of Moratuwa and Dr. Lochandaka Ranatunga, current Head, Department of Information Technology, University of Moratuwa for granting me two years of leave for my studies.

I would like to thank, in a special way, Professor Dileeka Dias and Dr. Shehan Perera, my supervisors, for their advice, guidance and enthusiasm and for supporting and standing by me throughout my first steps in research. I equally thank Professor J.M.S.J Bandara and Dr. Ajith Pasqual, the progress review committee members for directing and providing valuable input throughout this research. I am grateful to Dr. K.C.B. Wewagedara, postgraduate coordinator of the Department of Electronic and Telecommunication Engineering, for supporting and guiding through the administrative work related to this research. I do thank Nimalika Fernando for all her advices and support extended during the initial phase of the project. Nimalika, the discussions with you accelerated this research and were a relief during difficult times. Thank you very much Indika Karunaratne, for all the valuable comments for the draft version of the thesis. Finally, a big thank you for you, Anjana Uduwaragoda, my husband for all the support and encouragement extended throughout these two years.

TABLE OF CONTENTS

Declaration	i
Abstract	ii
Dedication	iii
Acknowledgements	iv
Table of Contents	v
List of Figures	ix
List of Tables	xii
List of Abbreviations	
List of Appendices	xiv
1. Introduction	1
1.1 Background	1
1.2 Problem Statement	4
1.3 Aim and the Research Questions	4
1.4 Organization of the Thesis	5
2. Related Work	7
2.1 Introduction	7
2.2 Promise of Future Intelligent Transportation Systems	7
2.3 ITS and Digital Maps	13
2.4 GPS Based Map Generation – Existing Methods	17
2.5 GPS Based Map Generation – Summary of Features in Existing Methods	36
2.6 GPS Based Map Generation – Evaluation Techniques Used by Existing Methods	37
2.7 Directions for Future Research	41
2.8 Summary	43

3.	Probability Density Estimation	44
3.1	Introduction	44
3.2	Probability Density Estimation	43
3.3	Parametric Methods	45
3.4	Non-parametric Methods	45
3.4.1	Histogram	46
3.4.2	Naïve Estimator	46
3.4.3	Kernel Density Estimator	47
3.5	Summary	54
4.	Research Methodology	55
4.1	Introduction	55
4.2	Stage I : Trajectory Outlier Removal	56
4.2.1	Stage 1.1 - Phase I of NECTOD	59
4.2.2	Stage 1.2 - Phase II of NECTOD	61
4.3	Stage II: Choosing a Suitable Kernel Function and a Bandwidth calculation Method for the KDE	64
4.4	Stage III: Road Map Enhancement -The Proposed Approach for Lane Level Road Data Extraction from GPS Trajectories	65
4.4.1	Finding the number and locations of lane centers and calculating lane width	66
4.4.2	Finding the lane boundaries	71
4.5	The Overall Process	71
4.6	Summary	73

5.	Experimental Evaluation and Analysis of Results	74
5.1	Introduction	74
5.2	Experimental Data & Data Conversion	75
5.2.1	Experimental data	75
5.2.2	Data conversion	77
5.3	Stage I: Experimental Evaluation of the Proposed Method for Network Constrained Trajectory Outlier Removal	78
5.3.1	Stage 1.1 - Experimental evaluation of Phase I of NECTOD	79
5.3.2	Stage 1.2 - Experimental evaluation of Phase II of NECTOD	81
5.3.3	Statistical evaluation	82
5.3.4	Selection of parameters for NECTOD	85
5.4	Stage II: Choosing a Suitable Kernel Function and a Bandwidth Calculation Method	86
5.4.1	Choice of a kernel function	86
5.4.2	Choice of a method for calculating the optimal bandwidth	87
5.5	Stage III: Evaluation of the Proposed Methods for Calculation of Lane Centers and Boundaries	95
5.5.1	Experimental design	95
5.5.2	Determination of lane centerlines and lane width	96
5.5.3	Determination of lane boundaries	99
5.5.4	Overall evaluation of results	100
5.6	Discussion	104

6. Conclusion	106
6.1 Introduction	106
6.2 Contributions	106
6.3 Challenges	109
6.4 Directions for Future Research	109
References	111
Appendix A - Kernel Bandwidth Selection Experiment - Road Section-wise Results	125
Appendix B – Code Snippets	127



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

LIST OF FIGURES

Figure 1.1	Examples of anticipated Intelligent Transportation Systems (ITS)	1
Figure 1.2	Structure of the Thesis	6
Figure 2.1	Benefits of ITS	9
Figure 2.2	Interlocking subareas in digital mapping for ITS	15
Figure 2.3	A framework for GPS data driven road map generation, update and refinement	17
Figure 2.4	Different objectives of existing GPS based map making research	18
Figure 2.5	The method proposed by Rogers et al.	22
Figure 2.6	The method proposed by Edelkamp et al. [64] and Schroedl et al	24
Figure 2.7	The method proposed by Davies et al.	25
Figure 2.8	The method proposed by Guo et al.	26
Figure 2.9	The method proposed by Worrall and Nebot	26
Figure 2.10	The method proposed by Chen and Cheng	27
Figure 2.11	The method proposed by Shi et al	28
Figure 2.12	The method proposed by Gao and Krumm	29
Figure 2.13	The method proposed by Chen and Krumm	30
Figure 2.14	The method proposed by Zhang et al.	31
Figure 2.15	The method proposed by Agamennoni et al.	32
Figure 2.16	The method proposed by Karagiorgou and Pfoser	32
Figure 2.17	The method proposed by Liu et al	33
Figure 2.18	The method proposed by Biagioni and Eriksson	34
Figure 2.19	The method proposed by Knoop et al	35
Figure 3.1	Shapes of some Kernel Functions	49
Figure 3.2	KDE is obtained by summing the Kernels	49
Figure 3.3	Effect of Kernel function on the final density estimate	50
Figure 3.4	For a given dataset, a family of density estimates could be generated using different values for the bandwidth	50
Figure 3.5	Larger bandwidths over-smoothes the density estimate whereas smaller bandwidths under-smoothes the density estimate	51

Figure 4.1	The stages of the research methodology	55
Figure 4.2	Phases in NECTOD	58
Figure 4.3	Process flow diagram of Phase I of NECTOD	60
Figure.4.4	Three distance components of TRAOD	61
Figure 4.5	Process flow diagram of Phase II of NECTOD	63
Figure 4.6	Trajectory density at lane centers is greater than the trajectory density at lane boundaries	66
Figure 4.7	Trajectories intersect with the perpendicular line drawn through the centerline	67
Figure 4.8	The kernel density estimate of Xc for a selected road segment	68
Figure 4.9	Process flow of finding lane cluster centers	70
Figure 4.10	Process flow diagram of the overall process	72
Figure 5.1	The stages of the experimental evaluation process	74
Figure 5.2	Nokia N95 Phone	76
Figure 5.3	Nimitz freeway (I-880)	76
Figure 5.4	Geographical zones in the SPCS	77
Figure 5.5	The sample of selected GPS points containing outliers	78
Figure 5.6	Comparison of output generated by Phase I for Dataset I	80
Figure 5.7	Output of Phase II of Dataset I	82
Figure 5.8	Histograms showing the segment distribution across the road	83
Figure 5.9	Comparison of results of NECTOD and TRAOD	84
Figure 5.10	Effect of the kernel function on the kernel density estimate	87
Figure 5.11	Kernel Density Estimates obtained for a given road segment using different values for the bandwidth	88
Figure 5.12	Design of the experiment	90
Figure 5.13	Comparison of different bandwidth calculation methods when N is large (N=1019), for a road segment with four lanes (from Section 1 of the road) – density estimates in one graph	91
Figure 5.14	Comparison of different bandwidth calculation methods when N is large (N=1019), for a road segment with four lanes (from Section 1 of the road) – density estimates in different graphs	92

Figure 5.15	Comparison of different bandwidth calculation methods when N is small (N=144), for a road segment with four lanes (from Section 1 of the road) – density estimates in one graph	93
Figure 5.16	Comparison of different bandwidth calculation methods when N is small (N=144), for a road segment with four lanes (from Section 1 of the road) – density estimates in different graphs	94
Figure 5.17	Road parts selected for the experiment	96
Figure 5.18	Lane centerlines generated for the selected road parts	97
Figure 5.19	KDE graphs for some selected centerline segments of road part 3	98
Figure 5.20	Lane boundaries generated for the selected road parts	99
Figure 5.21	The difference between actual and calculated lane boundaries	100
Figure 5.22	Error in calculation of number of lanes vs. the number of traces	102
Figure 5.23	Straight section - Generated lane centerlines overlaid on Google Earth	102
Figure 5.24	Curved section - Generated lane centerlines overlaid on Google Earth.	103
Figure 5.25	Section with the exit lane- Generated lane centerlines overlaid on Google Earth.	103
Figure 5.26	A portion of the generated centerline map overlaid on Google Earth – a closer view	104



University of Moratuwa, Sri Lanka.
 Electronic Theses & Dissertations
www.lib.mrt.ac.lk

LIST OF TABLES

Table 2.1	Major research initiatives related to mapping for ITS	16
Table 2.2	Classes of ITS based map generation and respective research references	19
Table 2.3	Strengths and the limitations of the existing approaches of GPS based map generation approaches	20
Table 2.4	Summary of the general features of existing GPS-based map generation and refinement methods	37
Table 2.5	Ground truth and the evaluation methods used by existing map generation/refinement approaches	39
Table 3.1	Kernel Functions	48
Table 4.1	Methodology used to determine the parameters for kernel density estimation	65
Table 5.1	Trajectory datasets used for the experiment	79
Table 5.2	Summary statistics of phase I	81
Table 5.3	Summary statistics for Phase II	81
Table 5.4	Comparison of the number of trajectory segments before and after each phase	85
Table 5.5	Features of the road sections selected for the experiment	89
Table 5.6	Results of the experiment (for all sections) conducted to find the best bandwidth selector	93

LIST OF ABBREVIATIONS

Abbreviation	Description
ADAS	Advanced Driver Assistance Systems
BCV	Biased Cross Validation
BOT	Kernel bandwidth calculation method by Botev et al. [97]
CALM	Communications Access for Land Mobiles
DGPS	Differential Global Positioning System
D ² ITS	Data Driven Intelligent Transportation Systems
DVB-SH	Digital Video Broadcasting - Satellite services to Handhelds
GMM	Gaussian Mixture Model
GPS	Global Positioning System
ISO-GDF	ISO- Geographic Data Format
ITS	Intelligent Transportation Systems
KDE	Kernel Density Estimation
LDM	Local Dynamic Maps
LSA	Least Squares Approximation
LCSV	Least Squares Cross-Validation
NDS	Navigation Data Standard
NECTOD	Network Constrained Trajectory Outlier Detection
OSM	Open Street Maps
PDA	Personal Digital Assistant
PDF	Probability Density Function
PSF	Physical Storage Format
PPP-GPS	Precise Point Positioning – Global Positioning System
RT	Rule of Thumb for Gaussian
SJ	Sheather & Jones Plug-in Method
SPCS	State Plane Coordinate System
TRAOD	Trajectory Outlier Detection
UCV	Unbiased Cross Validation
WAVE	Wireless Access in Vehicular Environments



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk

LIST OF APPENDICES

Appendix A	125
Appendix B	127



University of Moratuwa, Sri Lanka.
Electronic Theses & Dissertations
www.lib.mrt.ac.lk