



**HAL**  
open science

# Integrity Analysis of Data Sources in Multimodal Localization System

Arjun Balakrishnan

► **To cite this version:**

Arjun Balakrishnan. Integrity Analysis of Data Sources in Multimodal Localization System. Robotics [cs.RO]. Université Paris-Saclay, 2020. English. NNT : 2020UPASG060 . tel-03245708

**HAL Id: tel-03245708**

**<https://theses.hal.science/tel-03245708v1>**

Submitted on 2 Jun 2021

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Integrity Analysis of Data Sources in Multimodal Localization System

**Thèse de doctorat de l'université Paris-Saclay**

École doctorale n° 580, Sciences et Technologies de  
l'Information et de la Communication (STIC)  
Spécialité de doctorat: Robotique  
Unité de recherche: Université Paris-Saclay, ENS Paris-Saclay,  
CNRS, SATIE, 91190, Gif-sur-Yvette, France  
Réfèrent: : Faculté des Sciences d'Orsay

**Thèse présentée et soutenue en visioconférence totale,  
le 14/12/2020, par**

**Arjun BALAKRISHNAN**

## Composition du Jury:

<b>Hugues MOUNIER</b> Professeur des Universités, Université Paris-Saclay	Président
<b>Véronique CHERFAOUI</b> Professeur des Universités, Université de Technologie de Compiègne, Sorbonne Universités	Rapporteuse & Examinatrice
<b>David BETAÏLLE</b> Directeur de Recherche (HDR), COSYS/SII, IFSTTAR, Univer- sité Gustave Eiffel	Rapporteur & Examineur
<b>Cindy CAPPELLE</b> Maître de Conférences, Université de Technologie de Belfort- Montbéliard	Examinatrice
<b>Nicole EL-ZOGHBY</b> Ingénieur de Recherche, Groupe Renault	Examinatrice
<b>Patrice AKNIN</b> Directeur de Recherche, Institut de Recherche Technologique SystemX	Examineur
<b>Roger REYNAUD</b> Professeur des Universités, Université Paris-Saclay	Directeur de thèse
<b>Sergio RODRIGUEZ</b> Maître de Conférences, Université Paris-Saclay	Co-encadrant



## **Acknowledgments:**

This PhD thesis was carried out at the SATIE Laboratory in the Paris-Saclay University from October 2017 to December 2020. It was financed with a grant from the French National Education and Research Ministry. I would like to express my gratitude to everyone who has contributed to this research directly or indirectly, without whom I would not have been able to complete this research. I wish to extend my sincere thanks to my supervisors Professor Roger Reynaud and Sergio Rodriguez, for their valuable advices, their continuous encouragements and their personal involvement. Their impeccable knowledge in the subject matter and careful guidance not only helped me tremendously in this journey, but also provided me with precious insights for my scientific and personal life. A special thanks to Flavien Delgehier, for providing all the required technical support during experiments conducted in this work. I am grateful to all members and researchers of SATIE and DIGITEO laboratories, who had taken part in the discussions and evaluations of this work, at its various stages. Finally, I thank my family and friends for their encouragements and the great level of understanding shown throughout the years of this work.



---

## **Abstract**

Intelligent vehicles are a key component in humanity's vision for safer, efficient, and accessible transportation systems across the world. Due to the multitude of data sources and processes associated with Intelligent vehicles, the reliability of the total system is greatly dependent on the possibility of errors or poor performances observed in its components. In our work, we focus on the critical task of localization of intelligent vehicles and address the challenges in monitoring the integrity of data sources used in localization. The primary contribution of our research is the proposition of a novel protocol for integrity by combining integrity concepts from information systems with the existing integrity concepts in the field of Intelligent Transport Systems (ITS). An integrity monitoring framework based on the theorized integrity protocol that can handle multimodal localization problems is formalized. As the first step, a proof of concept for this framework is developed based on cross-consistency estimation of data sources using polynomial models. Based on the observations from the first step, a 'Feature Grid' data representation is proposed in the second step and a generalized prototype for the framework is implemented. The framework is tested in highways as well as complex urban scenarios to demonstrate that the proposed framework is capable of providing continuous integrity estimates of multimodal data sources used in intelligent vehicle localization.



---

## Résumé

Les véhicules intelligents sont un élément clé pour des systèmes de transport plus sûrs, efficaces et accessibles à travers le monde. En raison de la multitude de sources de données et de processus associés aux véhicules intelligents, la fiabilité de l'ensemble du système dépend fortement de la possibilité d'erreurs ou de mauvaises performances observées dans ses composants. Dans notre travail, nous nous intéressons à la tâche critique de localisation des véhicules intelligents et relevons les défis de la surveillance de l'intégrité des sources de données utilisées dans la localisation. La contribution clé de notre recherche est la proposition d'un nouveau protocole d'intégrité en combinant les concepts d'intégrité des systèmes d'information et les concepts d'intégrité existants dans les Systèmes de Transport Intelligents (STI). Un cadre de surveillance de l'intégrité basé sur le protocole d'intégrité proposé qui peut gérer les problèmes de localisation multimodale est développé. Dans la première étape, une preuve de concept pour ce cadre est développée sur la base d'une estimation de cohérence croisée des sources de données à l'aide de modèles polynomiaux. Sur la base des observations de la première étape, une représentation des données «Feature Grid» est proposée dans la deuxième étape et un prototype généralisé pour le cadre est mis en œuvre. Le cadre est testé sur les autoroutes ainsi que dans des scénarios urbains complexes pour démontrer que le cadre proposé est capable de fournir des estimations d'intégrité continue des sources de données multimodales utilisées dans la localisation intelligente des véhicules.



---

## Synthèse en Français

Les véhicules intelligents promettent un monde avec des transports efficaces et sûrs. Les progrès des systèmes de navigation par satellite (GNSS), des capteurs toujours plus performants et des méthodes de perception basées sur l'intelligence artificielle (IA) ont propulsé la recherche sur les véhicules intelligents plus près de l'objectif des systèmes de transport intelligents (STI) entièrement fonctionnels. Cependant, en raison de la multitude de sources de données et de processus associés aux STI, la fiabilité de l'ensemble du système dépend fortement de la possibilité d'erreurs ou de mauvaises performances observées dans ses composants. Dans cette thèse, nous nous intéressons à la tâche critique de localisation des véhicules intelligents et abordons les défis de la surveillance de l'intégrité des sources de données utilisées pour la localisation. Bien que plusieurs travaux aient fourni des méthodes de surveillance de l'intégrité pour certaines sources de données notamment les systèmes GNSS et les cartes dans le contexte de certaines applications ou environnements d'exploitation, il existe encore de la place pour une approche généralisée et évolutive dédiée à la surveillance de l'intégrité des sources de données multimodales utilisées pour la localisation de véhicules dans divers scénarios.

Le premier chapitre de cette thèse est dédié à l'étude des différentes interprétations du concept d'intégrité dans le domaine scientifique et de leur application actuelle à la localisation des véhicules intelligents. Les limites des approches actuelles sont identifiées et un nouveau protocole d'intégrité est défini en combinant les concepts d'intégrité des systèmes d'information et les concepts existants dans les STI. Dans le deuxième chapitre, le problème de la localisation multimodale est décrit en catégorisant et en analysant différents types d'approches de localisation en fonction des sources de données

utilisées. Cette recherche s'est également concentré sur la modélisation des sources de données utilisées dans différentes approches de localisation et sur l'analyse complète des avantages et des inconvénients de ces modèles. Le troisième chapitre propose un cadre de surveillance de l'intégrité utilisant le protocole d'intégrité théorisé qui peut gérer les problèmes de localisation multimodale. Des arguments pour la pertinence du cadre proposé et le choix des méthodes d'estimation de l'intégrité sont également présentés dans ce chapitre. Le quatrième chapitre décrit la modélisation géométrique des sources de données et la technique d'estimation de l'intégrité basée sur la cohérence croisée développée pour la validation du cadre de surveillance de l'intégrité proposé. Une seule entité (forme de la route) est détectée à partir de toutes les sources de données et modélisée à l'aide de polynômes quadratiques. Les résultats obtenus à partir de cette méthodologie servent de preuve de concept pour le cadre proposé et montrent qu'il est capable de surveiller l'intégrité des données du GNSS, de la carte et de la vision dans des scénarios autoroutiers et semi-urbains. Dans le cinquième chapitre, un modèle de représentation des données amélioré et une technique d'estimation d'intégrité basée sur la cohérence sont proposés. La représentation des données à l'aide des «Feature Grids» proposées incorpore de multiples caractéristiques spatiales de l'environnement, telles que les structures routières, les marquages routiers et la géométrie de l'infrastructure environnante. Les résultats des expériences utilisant des ensembles de données accessibles au public ainsi que des acquisitions de données réelles sont analysés pour montrer l'efficacité du cadre proposé. Les expériences ont montré que le cadre proposé est capable de fournir des estimations d'intégrité continues de différents types de sources de données dans des scénarios routiers ainsi que dans des scénarios urbains complexes. De plus, les paramètres d'intégrité de localisation classiques sont également calculés à l'aide de ce cadre et comparés aux exigences d'intégrité de localisation standard disponibles dans le domaine des véhicules intelligents. Enfin, cette thèse se conclut en mettant en évidence les principaux enseignements de cette recherche, les potentiels et les améliorations possibles de l'approche destination d'intégrité proposé.

---

# Contents

Acknowledgments . . . . .	i
Abstract . . . . .	iii
Résumé . . . . .	v
Synthèse en Français . . . . .	viii
List of Tables . . . . .	xi
List of Figures . . . . .	xiv
Acronyms . . . . .	xvii
<b>1 What Do We Know?</b>	<b>1</b>
1.1 Introduction . . . . .	2
1.2 Defining Integrity . . . . .	2
1.3 Integrity in Intelligent Vehicles . . . . .	8
1.4 Conclusion . . . . .	19
<b>2 What Is The Problem?</b>	<b>23</b>
2.1 Introduction . . . . .	24
2.2 Different Types of Localization . . . . .	25
2.3 Odometry . . . . .	29
2.4 SLAM . . . . .	37
2.5 Map-Based Localization . . . . .	41
2.6 Modeling Multimodal Data . . . . .	44
2.7 Conclusion . . . . .	50
<b>3 The Proposed Solution</b>	<b>51</b>
3.1 Introduction . . . . .	52
3.2 Why Integrity of Sources? . . . . .	53

---

3.3	Modeling of Data Sources . . . . .	55
3.4	Integrity Assessment Technique . . . . .	57
3.5	Integrity Markers . . . . .	59
3.6	Conclusion . . . . .	60
<b>4</b>	<b>The Proof Of Concept</b>	<b>61</b>
4.1	Introduction . . . . .	62
4.2	Data Handling . . . . .	62
4.3	Integrity Assessment . . . . .	71
4.4	Experiments and Discussion . . . . .	74
4.5	Conclusions . . . . .	83
<b>5</b>	<b>The Generalized Prototype</b>	<b>85</b>
5.1	Introduction . . . . .	86
5.2	Methodology . . . . .	89
5.3	Experiments and Discussions . . . . .	105
5.4	Conclusion . . . . .	114
<b>6</b>	<b>Conclusions and Perspectives</b>	<b>117</b>

---

# List of Tables

1.1	Quantification of Alert Limit requirements in highways and urban scenarios	19
2.1	Different representations used for map data . . . . .	49
3.1	Modality and properties of data sources . . . . .	55
4.1	Performance evaluation on Dataset 1 . . . . .	82
5.1	Comparison of results obtained using FG-based method and polynomial-based method . . . . .	110
5.2	Meta-Analysis of availability of data integrity monitoring process . . . . .	115



---

# List of Figures

1.1	Summary of Integrity concepts in Information systems . . . . .	6
1.2	The Stanford-ESA Diagram . . . . .	10
1.3	Integrity events based on PE, PLs, ALs in ground vehicle navigation . .	11
1.4	FDI principle for localization using multiple measurements . . . . .	12
1.5	Principle of classic RAIM implementation . . . . .	13
1.6	Challenges for GNSS positioning in urban environments . . . . .	15
2.1	GNSS trilateration process . . . . .	26
2.2	Classification of self-localization concepts . . . . .	30
2.3	Different GNSS trajectory models . . . . .	46
3.1	Localization module in Intelligent Vehicles . . . . .	53
3.2	Framework for integrity monitoring of sources . . . . .	58
4.1	Functional block diagram of the methodology proposed . . . . .	63
4.2	Vision data processing pipeline . . . . .	64
4.3	Lane markings detection and representation . . . . .	65
4.4	Different frames associated with the analysis . . . . .	66
4.5	Point-to-Curve map matching algorithm . . . . .	68
4.6	Representation of data from different sources in ego-frame during typical scenario . . . . .	69
4.7	Error evolution for Dataset 1. . . . .	75
4.8	Different scenarios from dataset 1 . . . . .	75
4.9	Error evolution for Dataset 2 . . . . .	77
4.10	Different scenarios from dataset 2 . . . . .	77
4.11	Error evolution for Dataset 3 . . . . .	78

---

4.12	Different scenarios from dataset 3 . . . . .	78
4.13	Error evolution for Dataset 4 . . . . .	79
4.14	Different scenarios from dataset 4 . . . . .	79
4.15	Performance evaluation on Dataset 1 . . . . .	80
5.1	Integrity issues in map sources . . . . .	87
5.2	Framework for integrity assessment of multi-modal data sources . . . . .	90
5.3	Example of features used in our work, detected from a complex scenario . . . . .	91
5.4	Steps for detecting features from vision . . . . .	93
5.5	Detection criteria for LiDAR data . . . . .	94
5.6	3D ROI to grid representation for integrity monitoring of sources . . . . .	96
5.7	Example of modeling data from different sources using Feature Grid representation . . . . .	97
5.8	Matching operation between FGs of two sources . . . . .	98
5.9	Transformation optimization process using sequential particle filters . . . . .	102
5.10	Illustration of Protection Levels for localization of ground vehicles . . . . .	104
5.11	Comparison Results of dataset 2011_09_26_drive_0029 . . . . .	107
5.12	Specific scenarios from dataset 2011_09_26_drive_0029 . . . . .	108
5.13	Comparison results of dataset 2011_09_26_drive_0028 . . . . .	109
5.14	Examples of complex scenarios . . . . .	111
5.15	Integrity risk allocation . . . . .	112
5.16	Protection Levels (PL) comparison . . . . .	114

---

# Acronyms

---

AL	Alert Limit
BEV	Bird-Eye-View
BE	Boulder Error
CDF	Cumulative distribution function
DR	Dead Reckoning
DTAM	Dense tracking and mapping
DBSCAN	Density-based spatial clustering technique
DPM	Dirichlet Process Mixture
ENU	East-North-UP
ECA	Error Characterization Approaches
EGNOS	European Geostationary Navigation Overlay Service
EKF	Extended Kalman Filter
FD	Fault Detection
FDE	Fault Detection and Exclusion
FDI	Fault Detection and Isolation
FP	Fault/Feasibility Predictor
FG	Feature Grid
FMCW	Frequency Modulated-Continuous Wave
GMM	Gaussian Mixed Model
GIS	Geographic Information System
GDOP	Geometric Dilution Of Precision
GNSS	Global navigation satellite System
GPS	Global positioning System
GoF	Goodness-of-Fit

HMI	Hazardously Misleading Information
HMM	Hidden Markov Model
HD	High Definition
HPE	Horizontal Position Error
HPL	Horizontal Protection Level
IR	Integrity Risk
IW	Integrity Weight
ITS	Intelligent Transport Systems
ISO	International Organization for Standardization
IDW	Inverse Distance Weighting
IPM	Inverse Perspective Transform
ICP	Iterative Closest Point
LO	Laser Odometry
LatPL	Lateral Protection Level
LCA	Liability Critical Applications
LOAM	LiDAR Odometry and Mapping
LiDAR	Light Detection and Ranging
LOS	Line-Of-Sight
LBS	Location-Based Services
LonPL	Longitudinal Protection Level
LUT	Look-Up-Table
ML	Machine Learning
MRA	Measurement Rejection Approaches
MEO	Medium Earth Orbit
MI	Misleading Information
MDS	Motion Data Sources
NLOS	non-Line-Of-Sight
OSM	OpenStreetMap
PF	Particle filter
PE	Position Error
PL	Protection Level

QPM	Quadratic Polynomial Model
RO	Radar Odometry
RADAR	Radio Detection and Ranging
RAIM	Receiver Autonomous Integrity Monitoring
ROI	Region Of Interest
SLAM	Simultaneous Localization and Mapping
SDS	Spatial Data Sources
SFM	Structure From Motion
SA	System Available
SU	System Unavailable
TLS	Target Level Safety
TTA	Time to Alert
UKF	Unscented Kalman Filter
VPL	Vertical Protection Level
VIO	Visual Inertial Odometry
VO	Visual odometry
VGI	Volunteered Geographic Information
WAAS	Wide Area Augmentation System
WGS	World Geodetic System



# Chapter 1

## What Do We Know?

### Contents

---

<b>1.1</b>	<b>Introduction</b>	<b>2</b>
<b>1.2</b>	<b>Defining Integrity</b>	<b>2</b>
<b>1.3</b>	<b>Integrity in Intelligent Vehicles</b>	<b>8</b>
1.3.1	Components of Localization Integrity	9
1.3.2	Classical Approaches to Localization Integrity Monitoring	11
1.3.3	Integrity in Urban Localization	15
1.3.4	Quantification of Localization Integrity Components	18
<b>1.4</b>	<b>Conclusion</b>	<b>19</b>

---

## 1.1 Introduction

In this chapter, we outline the current state of the research in the field of integrity monitoring of localization. Defining integrity in the context of autonomous vehicles is a complicated task due to the multitude of subsystems and processes associated with them. This chapter is organized thematically. In the first section, discussions are started from the most pragmatic definitions of integrity and consistently narrowed down to the definitions of integrity in scientific branches related to the work presented in this thesis. Successively, we examine the possible candidates of integrity definitions that can be applied to our work and their components. Based on the analysis of the applicability of these potential candidate definitions, we formalize the definitions, components and implications of the integrity concept used in this work. In the following sections, an attempt to substantiate the need for this approach is presented and the techniques used for realizing these specific components of integrity in the context of autonomous vehicles are studied and described. Real world applications of integrity monitoring in autonomous vehicles are studied. Finally, the area of integrity monitoring of data sources is isolated and examined in detail.

## 1.2 Defining Integrity

Majority of words that are used in scientific literature are formalized versions of their counterparts from linguistic glossaries. The term 'integrity' is no different. According to *Merriam-Webster's collegiate dictionary*, integrity is defined around three core concepts: **Incorruptibility** - firm adherence to a code of especially moral or artistic values, **Soundness** - an impaired condition and **Completeness** - the quality or state of being complete or undivided. While these definitions paint a pretty good picture of a human with integrity, applying them on a system to describe its integrity is not a fruitful exercise. To that extent, one can even argue that the concept of integrity is specific to the types of systems and needs to be redefined according to the type. For example, the definition of integrity in 'structural integrity' of an architectural system is vastly different from the same in 'biological integrity' of a natural habitat. But upon closer inspection, one can also see that both structural integrity and biological integrity are different manifestations

of the same concept to describe the incorruptibility, soundness and completeness of the respective systems. This elicits the question: what is integrity in intelligent vehicles and autonomous navigation?

Since intelligent vehicles are complex systems, it is logical to start by looking at what is system integrity in the field of applied sciences. The correct performance of an intelligent vehicle is ensured when all of its systems achieve desired tasks (driving, parking etc.) while complying with the safety standards designed to avoid harm to passengers, surroundings or itself. From this point, it is worth noting that we use the vocabulary and definitions provided by International Organization for Standardization (ISO) for road vehicles engineering in ISO (2011) in this work. According to ISO (2011), safety is defined as the absence of unacceptable risk and risk is the combination of the probability of occurrence of harm and the severity of that harm. In the case of intelligent vehicles, harmful situations are the result of trust to suit different sub-systems such as localization, perception, mapping etc. In an attempt to derive the integrity framework for land vehicles from aviation integrity standards, Reid et al. (2019) describe integrity as the frequency of occurrences of a true error outside of the estimate of maximum possible error. Such occurrences feed hazardous information to the vehicle's sub-systems leading to system failure. Hence, they quantify integrity using the probability of failures per hour of operation.

While the approach taken by the works mentioned above revolves around the processes and systems associated with the intelligent vehicle, one specific process demands a slightly different approach; data handling. Integrity monitoring of data at the input level of all systems poses few challenges while applying the integrity concepts as described above. Input data is not a part of the information supplied by a system, but is a part of information supplied to the system. Irrespective of whether a system can be used for intended operation or not, input data can have integrity of its own. For example, in a system where multiple data sources are used to carry out a process, each data source can have different integrity levels independent to the ability of the process to mitigate their impacts while providing output. A good starting point for the attempt to define integrity of data sources is the vocabulary used in the field of computer science, particularly information systems. In the discussion on defining information integrity concepts presented in Boritz (2005), integrity is defined as an unimpaired or unmarred condition,

hence providing the entire correspondence of a representation with an original condition. When it comes to information integrity, it is a measure of representational faithfulness of the information to the condition or subject that is being represented by the information (Boritz 2005). The core attributes of information integrity are identified as accuracy, completeness, timeliness and validity in Boritz (2005), Boritz (2004), and CobiT (2002) and Boritz (2005) also lists the enablers of information integrity as security, availability, understandability, consistency, predictability, verifiability and credibility. While core attributes of information integrity refer to the minimum criteria that must be satisfied while judging representational faithfulness of information, enablers are the properties or factors of information that help realize those core attributes. Boritz (2005) also argues that perfect representational faithfulness is unachievable and hence it should be viewed as not an absolute quality but a degree of achievement. This helps us distinguish between the core attributes of integrity and the enablers of integrity - secondary attributes of information that lead to one or more core attributes of integrity. Since these concepts do not necessarily correspond to their linguistic definitions and most of them are frequently used in this work, we make some effort to summarize their relevant definitions here.

## **Core Attributes of Information Integrity**

**Accuracy** Quantification of correspondence of the information with real world object or event with some degree of precision (Boritz 2004). Practically accuracy is the same as correctness and it is linked to neutrality in representation of object or event.

**Completeness** The ratio of available information amount and the total information of the real world (Naumann and Freytag 2005). Completeness is closely tied to accuracy as the degree of achievable completeness sets the upper bound for achievable accuracy (Boritz 2005). In other words, if the information available has more coverage of the real world, the accuracy can also be higher.

**Timeliness** The availability of information at a time suitable for its use (Bailey and Pearson 1983). The information is current and timely and within preset definitions of the duration of time in an information period (Flowerday and Solms 2007). The limit of timeliness is defined according to the context of applications and timeliness can affect accuracy and completeness.

**Validity** Measure of whether the information represents real conditions, rules or relationships in the way or domain it was intended to represent (Boritz 2005). Validity of information ensures representational faithfulness beyond just correspondence with the physical world.

## Enablers of Information Integrity

**Security** Protection of information against unintentional and intentional malicious acts such as unauthorized creation, modification, or destruction, as well as inadvertent errors that could compromise its integrity (Boritz 2005; Flowerday and von Solms 2005).

**Availability** For information to be deemed accessible or available, users need to be able to work with the information in a way that meets their needs (Wang and Strong 1996). Information that is not available at the time of intended use will impact accuracy, completeness and timeliness of information, thus affecting user's activities or decision making (Boritz 2005).

**Understandability** Understandability of information refers to the quality of fulfilling the application-specific granularity/aggregation requirement to enable information integrity (Boritz 2005; Flowerday and Solms 2007).

**Consistency** A property that multiple recordings of the values for any of the attributes be the same or closely similar across time and space (Wang, Reddy, and Kon 1995). In Boritz (2005), consistency is defined as the stability of measurements over time or space when compared against each other according to the standards defined by particular representation rules. Uncertainties in the measurement environment can perturb information systems to cause adverse effects on their stability and consistency, hence their comparability.

**Dependability** The dependability of information is facilitated by consistency in how information is delivered to the system, the predictability of information processing and the predictability of the events that the systems are designed to process information about (CobiT 2002; Boritz 2004).

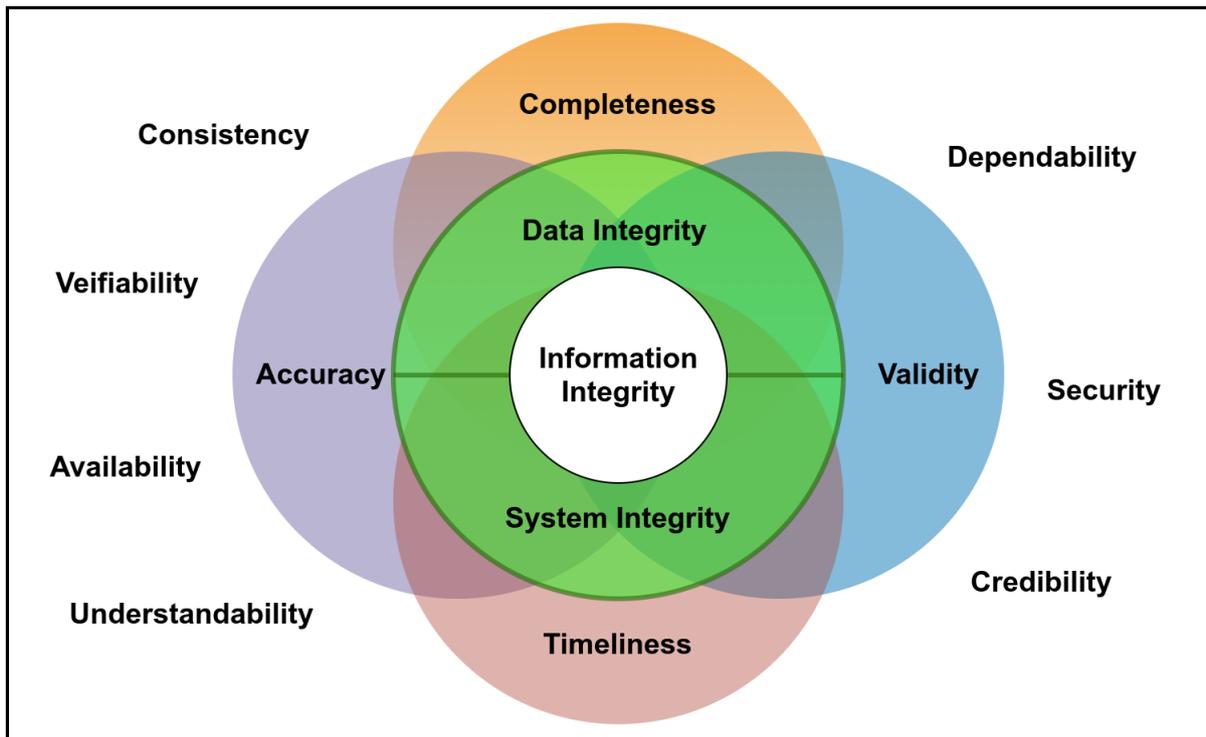


Figure 1.1: Summary of Integrity concepts in Information systems

**Verifiability** The ability of replicating a particular information using the same processes and tolerance limits for attributes of information integrity used to create the original information (Boritz 2005; Flowerday and von Solms 2005).

**Credibility** A level of certitude that the information corresponds to a real situation or object or has been obtained using a proper acquisition method. Credibility is often tied to the nature of the procedure used to evaluate verifiability by gathering evidence about its representational faithfulness (Boritz 2005; Flowerday and Solms 2007).

With the familiarity achieved on the integrity-related terminology used in information systems, we can start formalizing the integrity concept framework required for this work. Since we had begun this discussion by contrasting system integrity and data source integrity, we need to address the connection between information integrity and data integrity. According to the definitions provided by Braga and Logan (2018), data refers to pure and simple facts or values without any particular structure or organization and information is structured, organized data composed by syntax and semantics, that is, with reference to a code and to a given meaning. This leads to the inference that information

integrity subsumes data integrity.

Boritz (2004) and Flowerday and von Solms (2005) not only confirm this inference, but also establishes the connection between all the above mentioned integrity concepts and information integrity. They present information integrity as the combination of system integrity and data integrity. Here system integrity refers to data processing integrity of the system and it is achieved if its outputs fully and fairly reflect its inputs, and its processes are complete, timely, authorized and accurate. In other words, the upper-bound of information integrity is set by the integrity of the system processing the data (Boritz 2004). Information integrity can be worse, for example, if the data processed by a system lacks integrity at the time the system receives it, then the data will continue to lack integrity when it is transformed into information despite the system having integrity (Flowerday and von Solms 2005). The concepts discussed so far are shown in Fig. 1.1.

In many applications where integrity monitoring is crucial, the distinction between information integrity, data integrity and system (processing) integrity are quite blurred. As noted in Bovee, Srivastava, and Mak (2003), data can be viewed as facts or pieces of information hence even any organized combination of data is equivalent to information itself. This is particularly true in the case of sensor data. For example, an image from a camera is organized data and for a machine or human, the individual pixel values are of less use but the organization of those pixel values are useful to represent the patterns and relations of the real world. Hence attributes and enablers of data integrity, in most cases, will be the same as that of information integrity (Boritz 2005).

In other cases, they will be subsets of the attributes and enablers of information integrity and the choice of these subsets are specific to application (Wang, Reddy, and Kon 1995; Bovee, Srivastava, and Mak 2003). Following this school of thought, this work also treats data and information as practically synonymous. The discussion on defining the attributes and enablers of data integrity relevant for the application presented in this work is further developed in Chapter 3 as a precursor to the methodology developed. In the following section, we will look into the state-of-the-art of integrity monitoring approaches for both data sources and system processes in the field of intelligent vehicles.

### 1.3 Integrity in Intelligent Vehicles

According to *Merriam-Webster's collegiate dictionary*, Intelligence, in its absolute sense, is the ability to learn, understand or to deal with new or trying situations . Even though automotive industry has made great strides towards incorporating intelligence into vehicles with the help of artificial intelligence and robotics, this goal is far-fetched as of now. Current 'intelligent vehicles' can realize certain well-defined tasks reliably and assist human intelligence (the driver) to perform its tasks more efficiently and accurately. Hence, a realistic definition for 'intelligent vehicle' is given by Eskandarian (2012) as a vehicle that performs certain aspects of driving either autonomously or assists the driver to perform his/her driving functions more effectively, all resulting in enhanced safety, efficiency, and environmental impact. In this context, it is worth outlining the term 'autonomous driving' as it is a key component of intelligent vehicles. Autonomous driving is a perception-response process carried out by an intelligent vehicle to understand the environment and perform maneuvering functions to achieve safe driving (Eskandarian 2012). Intelligent vehicles are comprised of several such processes at various degree of automation. An intelligent vehicle with highest degree of automation in its driving process can plan the most feasible route from origin to desired destination, control its trajectory along the planned route and ensure navigational safety without any human intervention.

Transposing of integrity concepts from aviation on to intelligent vehicles is a relatively new frontier (Worner et al. 2016). In the case of aviation, localization is the key feature in operation and processes like route planning and navigational safety are traditionally carried out with the help of humans. This resulted in extensive integrity considerations in accurate localization data and process, while integrity of other components are rarely studied. Naturally, the adoption of aviation integrity concepts to develop integrity frameworks for intelligent vehicles favored integrity monitoring of road vehicle localization and its components. Except a few recent advancements in integrity monitoring of new-age data sources (LIDAR, High Definition Maps), majority of the literature in this area are based on integrity framework and concepts developed from localization integrity from aviation industry. Hence, it is important to understand this well-established state-of-the-art integrity concept.

### 1.3.1 Components of Localization Integrity

The core motivation of the integrity framework used in aviation and in turn in intelligent vehicles, is to quantify the requirements on localization safety. In ICAO (1973), integrity is considered as a measure of the trust that can be placed in the correctness of the information supplied by the total system. Integrity includes the ability of a system to provide timely and valid warnings to the user (alerts) when the system must not be used for the intended operation. Based on this definition, four attributes of localization integrity are defined in ICAO (1973) and Speidel et al. (2013) and widely used in majority of the high-integrity localization methods.

**Protection Level (PL)** is the estimated upper bound of true position error (PE) provided by the localization system. Different PLs can be defined according to the dimensionality of localization in a particular application. In aviation, Horizontal Protection Level (HPL) - the PL of positioning error in horizontal plane - and Vertical Protection Level (VPL) - the PL of height estimation - are defined. But in the case of ground vehicle localization, VPL is not addressed as the vertical localization (altitude estimation) of ground vehicles is less important. On the other hand, HPL is subdivided into longitudinal and lateral PLs (lonPL, latPL) as positioning errors of the vehicle 'along the road' and 'across the road' are important in ground vehicle navigation. **Alert Limit (AL)** is the allowable upper bound of PL, beyond which alert (warning) should be made to the user or the system. The time taken between surpassing the AL and issuing an alert is also a parameter of localization integrity, **Time to Alert (TTA)**. Finally, **Integrity Risk (IR)** is the probability of providing a localization estimate that is out of AL without warning the user within TTA and it is typically represented as number of such occurrences per-hour. Among these attributes, AL is defined by the application whereas PL is calculated by the system (or user).

Based on these four attributes, certain integrity events are outlined in the GNSS integrity survey provided in Zhu et al. (2018). The relation between PE and PL with respect to AL is used to define these integrity events and illustrated in Standard-ESA diagram in Tossaint et al. (2007) as shown in Fig. 1.2. When PL is less than AL, the positioning system can be considered available and denoted by the state **System Available (SA)**. If this condition is not met, positioning should be deemed faulty and

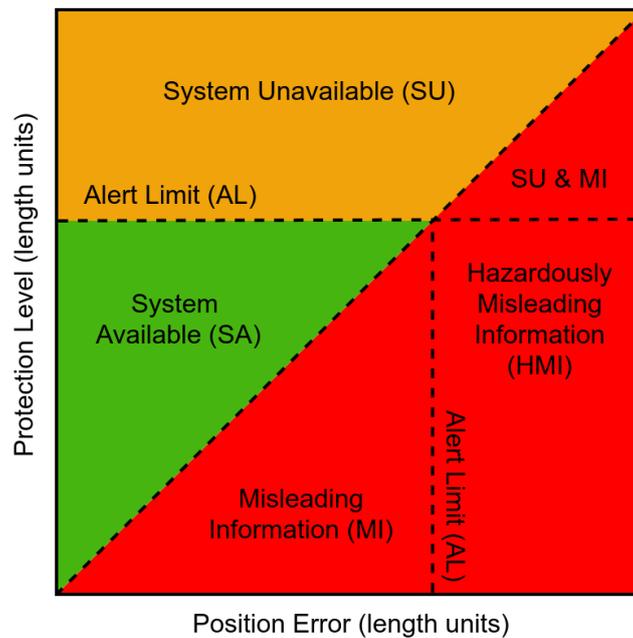


Figure 1.2: The Stanford-ESA Diagram (Tossaint et al. 2007)

belongs to the state **System Unavailable** (SU). According to Tossaint et al. (2007), **Misleading Information** (MI) is an integrity event that happens when, position error is higher than the PL but lesser than AL so that warning is not issued and system is treated as available. **Hazardously Misleading Information** (HMI) occurs when system is considered available but position error is higher than PL as well as AL. Although well specified and used in aviation, specifications of HMI for ground vehicle localization has not been defined yet (Zhu et al. 2018). If any of the integrity event persisted without an alert within TTA, that event is also regarded as an **Integrity Failure** (Zhu et al. 2018). However, the fact that the true position error is often unknown in practice, makes Standard-ESA diagram rather a conceptual tool and less applicable in real situations. One of the attempt to calculate these specifications for a GNSS positioning system by determining PL and PE for all possible satellite combinations, multiple times during representative tests, is presented in Sanz Subirana et al. (2008). In this work, vertical localization integrity attributes are modeled (VPL, vertical PE and AL) regarding an aerial platform.

When extended to localization of ground vehicles, latPL and lonPL can be used to represent integrity events as shown in Fig. 1.3. Fig. 1.3 shows two instances of a localization system with an application-specific lateral and longitudinal ALs. Each time, estimated positions are used to calculate the PLs. Due to the difference in the deviation of

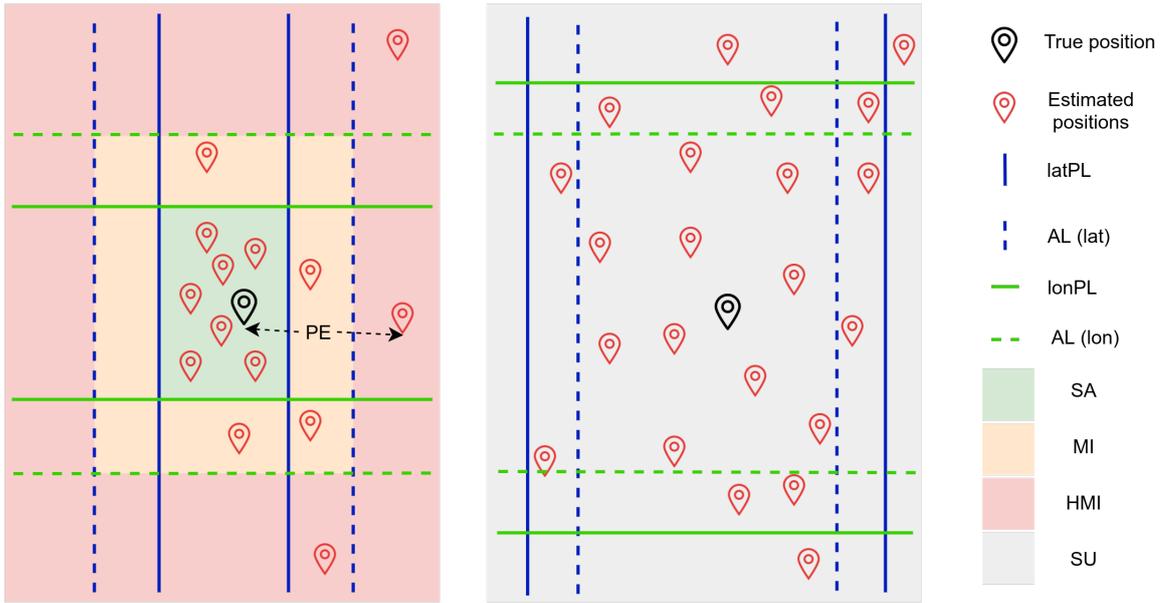


Figure 1.3: Integrity events based on PE, PLs, ALs in ground vehicle navigation

estimated locations from the true position, the PLs are calculated differently. When the PLs are higher than the ALs, the navigation system will be unavailable. But when PLs are lesser than ALs, any position estimate which is beyond any of the ALs is considered as HMI. Since area of HMI represents the state when position estimates are highly faulty and the system was unable to raise an alert, no counter measure can be taken. To mitigate such HMI events, several works such as Hwang et al. (2009), Worner et al. (2016) etc. have proposed Fault Detection and Isolation (FDI) and Bounder Error (BE) based approaches for alert generation.

### 1.3.2 Classical Approaches to Localization Integrity Monitoring

Fault detection and Isolation (FDI) is a concept derived from control theory that can be used to supervise a process in order to detect abnormalities and identify their causes. In this context, a fault is defined as an unacceptable deviation in one or more properties of a variable (Isermann 1997). If the cause is not important, FDI reduces to just Fault Detection (FD). FDI is achieved by comparing the same information provided by multiple sources. For example, by comparing the positioning estimation provided by the wheel encoders with the estimation provided by GNSS pseudorange measurements, we can identify the faults occurred in wheel speed measurements caused by wheel slippages. Redundancy of physical sensors are one way to achieve FDI with ease of implementation. However,

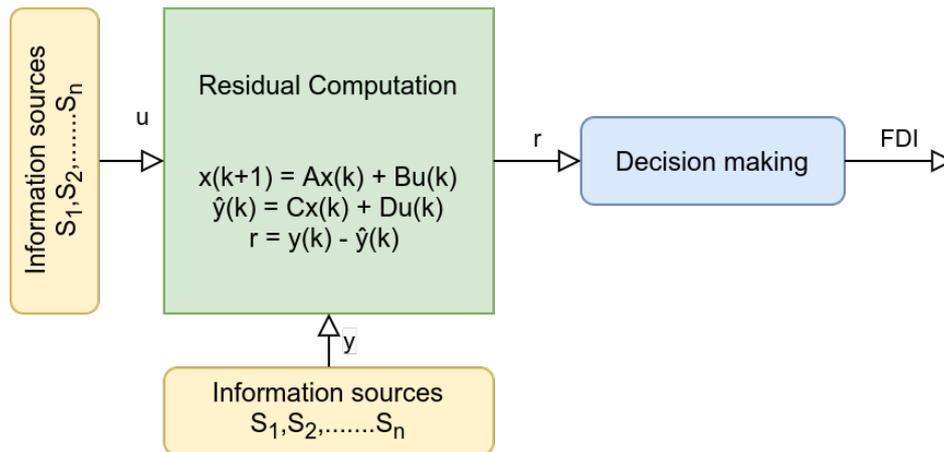


Figure 1.4: FDI principle for localization using multiple measurements (Worner et al. 2016)

since this approach is expensive, advanced methods employ mathematical models for the system and its relationship between sensor outputs to realize an approximate theoretical redundancy (Hwang et al. 2009).

In Worner et al. (2016), the basic principle of such methods are explained and Fig. 1.4 shows the block diagram for this. Though inputs  $u$  and measurements  $y$  are adopted from control theory notations, independent information sources used in localization process can be treated as input or measurements according to computational convenience (Worner et al. 2016). The key step in residual generation is the transformation of inputs  $u$  to the domain of measurements  $y$  using a system model (eg. acceleration measurement from inertial measurement unit is transformed to the velocity domain of an odometry system). After the transformation, residuals can be generated by comparing inputs and measurements directly. If it is a vector of residuals, the decision making block evaluates their magnitude and direction and compare them to detect fault. The output of residual generation can also be several structured residuals, each is sensitive to a specific probable faults at the source (Worner et al. 2016). In this case, along with detection of fault, fault isolation is also possible as residuals are correlated with the cause of faults. Additionally, Fault Detection and Exclusion (FDE) is presented in the literature, where once the fault is detected in an information source, the sensor or subsystems that caused it will be removed before providing a localization estimate.

In this context, the precursor to fault detection, isolation and exclusion approaches in localization need to be mentioned. Receiver Autonomous Integrity Monitoring (RAIM) is a

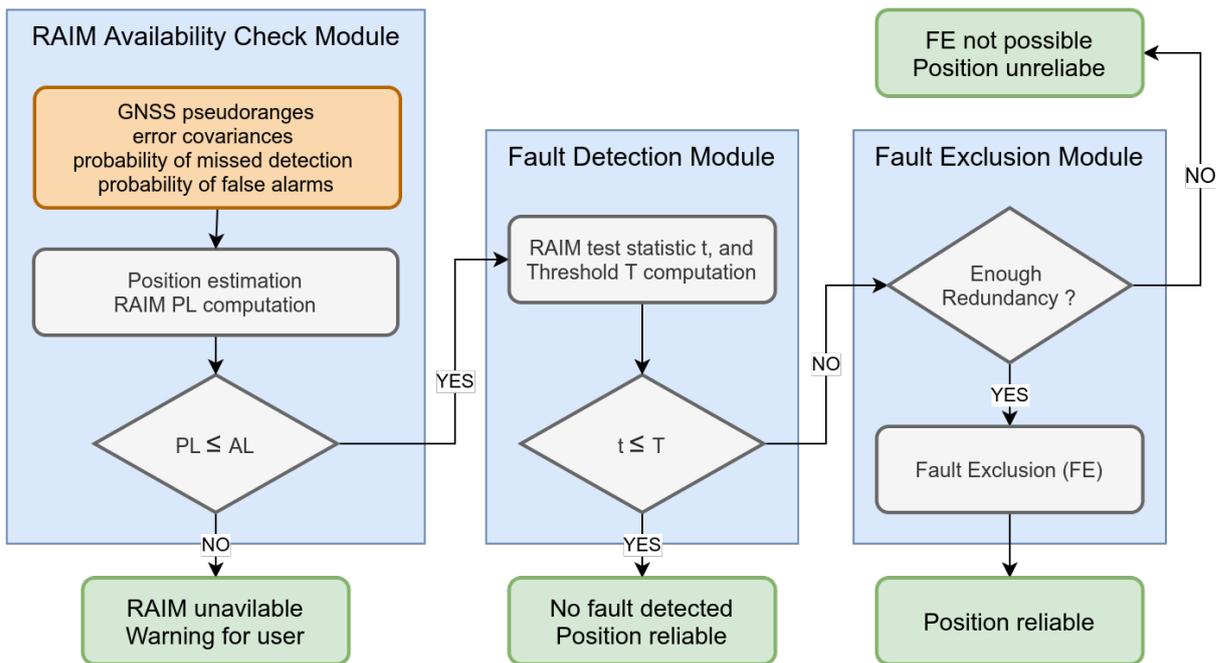


Figure 1.5: Principle of classic RAIM implementation (Zhu et al. 2018)

self-evaluation integrity monitoring framework for GNSS receivers, which was developed in the second half of 1980s for aviation field (Lee et al. 1986; Brown 1992). RAIM algorithms can provide correct positioning (FDE) or reliable alerts about faulty satellites (FDI). RAIM works by performing a consistency check on the redundant information from range measurements from a sufficiently large number of satellites (more than the minimum requirement for positioning). With more and more GNSS system deployed, the efficiency of RAIM algorithms have increased recently due to the abundance of satellites observable by a GNSS receiver at any location on earth (Eskandarian 2012). Based on the part of information used for consistency check there are three types of RAIM: range comparison based, parity-space based, least square residuals based (Worner et al. 2016; Zhu et al. 2018). In practice, they all achieve the same level of fault detections despite the conceptual differences, hence called snapshot algorithms of RAIM (Brown 1992). The classic implementation of RAIM algorithm is shown in Fig. 1.5.

In the first module, pseudorange measurements from different satellites combinations (subsets of available satellites) are combined to estimate positions corresponding to each of them. If there are five or more satellites available, PL is computed using RAIM techniques. Classic PL computation method is based on the relation between three factors (Kaplan and Hegarty 2005):

- error in GNSS solution - calculated as  $\sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2 + \sigma_t^2}$ , where  $\sigma_x^2$ ,  $\sigma_y^2$ ,  $\sigma_z^2$ ,  $\sigma_t^2$  are the variances of position solution error in three dimensions and time.
- pseudorange error factor - User Equivalent Range Error  $\sigma_{URE}$ , is the quality of range measurements of the GNSS receiver and it is estimated based on errors in orbit determination, synchronization, atmospheric transmission and receiver noises.
- geometry factor - Geometric Dilution Of Precision  $GDOP$ , is based on user-satellite geometry (satellites in view of the user) and used to amplify the standard deviations of pseudorange measurement errors to obtain reliable solution

These factors are related to each other through the equation (Kaplan and Hegarty 2005),

$$\sigma_{pos} = \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2 + \sigma_t^2} = \sigma_{URE} \times GDOP \quad (1.1)$$

Where  $\sigma_{pos}$  is the position confidence. However, position confidence is generally expressed separately as confidence of position in each dimension - vertical, horizontal, lateral, longitudinal etc. PLs are computed based on these separate position confidences as (Zhu et al. 2018),

$$XPL = k_X \cdot \sigma_X \quad (1.2)$$

where  $X$  is the dimension,  $k_x$  is an inflation factor determined using missed detection probability and  $\sigma_x$  is the position confidence in  $X$ .

If RAIM algorithm can compute PL and the PL is below AL, fault detection module calculates a test statistic based on pseudorange measurement residuals corresponding to each satellite combinations. When the test statistic is below a threshold value derived based on the required probability of false alarms, the position estimates are considered reliable. In the case of contrary, a fault is detected and a Fault Exclusion module is employed. If there is enough redundancy (ie., measurements from enough number of satellites), this module removes constellations that include satellites providing faulty information and recalculates a reliable position. If not, a non-excludable fault is detected and user is warned about an unreliable position estimate. However, in practice, the RAIM availability check module can also use a PL which is predicted using the satellite/user geometry, the nominal error characteristic (error variance) as well as the

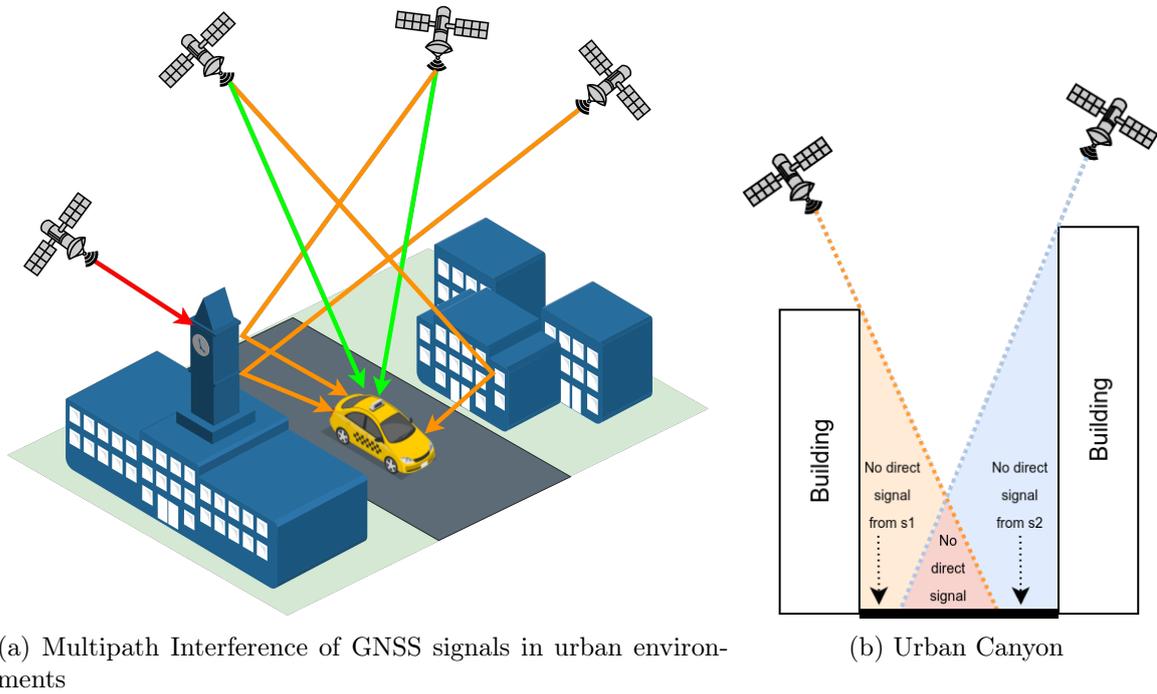


Figure 1.6: Challenges for GNSS positioning in urban environments

integrity probability requirements instead of employing current measurements (Zhu et al. 2018). In order for the RAIM to proceed to the FD module, PL should be always below AL. Once FDE is achieved, the actual uncertainty level can be estimated with the help of the satellite and receiver geometry, the residuals, the integrity probability requirements and error variance (Zhu et al. 2018).

### 1.3.3 Integrity in Urban Localization

Modern urban settlements provide a lot of challenges to navigation of intelligent vehicles. Multi-lane intersections, presence of high-rise buildings, pavements and pedestrians, high-density of vehicles etc. present in the urban environments introduce far more complex scenarios than those faced during highway navigation. As far as localization is concerned, urban environments presents several usage situations, where GNSS systems suffers to produce reliable localization estimates. The motivation behind several sensor fusion based localization systems, as explained in Chapter 2, is the poor performance of GNSS positioning in complex scenarios. In this section, we study the reasons for poor integrity of GNSS localization and current limitations of classical snapshot algorithms of RAIM to provide integrity monitoring in urban environments.

The term 'urban canyons' refers to a structural phenomena where a narrow ground level section (usually road sections) is situated between several high-rise structures on both sides as shown in Fig. 1.6b. These buildings can act as obstacles that block line-of-sight (LOS) coverage of GNSS satellites on some portion of the road. In Fig. 1.6b, the shaded regions show such areas and depending on the satellite configuration at an urban canyon these areas can intersect to result in an area where no LOS signal from any satellite is available. Since the blockage of signal happens from the satellites on either sides of the road than the ones along the road, the position solution accuracy will be greatly reduced in lateral direction of the vehicle movement (Groves et al. 2013). This affects the GDOP, which is a parameter that depends on the user-satellite geometry used in the calculation of protection levels. Hence, presence of urban canyons directly interferes with the classical GNSS integrity monitoring methods.

The next problem is the presence of flat and reflective surfaces present in the urban environment. Fig. 1.6a shows an example of a satellite configuration where signals from one or more satellites reflect on building surfaces and reach the GNSS receiver, along with LOS signals from satellites. Receiving signals from same satellite through different paths is called multipath reception. This can cause interference with LOS signals, making the time calculations inaccurate and in turn, the localization. Integrity monitoring methods which use multiple pseudorange measurements to examine the consistency between them (classical RAIM and its snapshots) can be unreliable in such situations. Multipath effects will be further worsened in non-LOS (NLOS) situations, where there are no LOS signals reaching the receiver but only the reflected signals. Due to the delay in signal reception due to reflection, a positive localization error will be observed (typically tens of meters, but can go up to one kilometer in the presence of distant skyscrapers (Groves et al. 2013)) and RAIM based methods will be unable to provide reliable integrity estimation because of large PLs. In any case, the assumptions of redundancy and single detectable fault at a time, used in RAIM algorithms does not hold in urban environments (NLOS and poor satellite visibility). In the survey provided by Zhu et al. (2018), it is noted that the RAIM availability (based on criteria inspired from aviation) decreases considerably in urban environments compared to rural and HPL computed in such methods is too big to be usable in urban applications.

Evidently, the error models used in pseudorange domain and position domain to

develop aviation-inspired integrity analysis need to be changed to formulate integrity analysis of GNSS in the context of urban environments. Although several works have proposed methods to mitigate multipath and NLOS errors to provide better localizations in urban environment, integrity monitoring in the presence of such errors mostly remains theoretical (Zhu et al. 2018). Regarding the pseudorange domain, DeCleene (2000) replaces the assumption of zero-mean Normal error distribution of pseudorange measurements with an error distribution that is symmetric and whose cumulative distribution function (CDF) is bounded. Zhu et al. (2018) note that the overbounding based error modeling can be a promising method for urban situations due to its ability to deal with bias due to NLOS and multipath effects, though it has been rarely addressed in current literature. To accommodate non-zero mean characteristics of urban NLOS signals, Gaussian Mixed Model (GMM) is proposed by Viandier et al. (2008) where the model is the combination of several weighted Gaussian distribution. But since the proposed GMM is dependent on the previous measurements, sudden variations in signal receptions cannot be modeled reliably. Viandier et al. (2010) address this issue by employing a Dirichlet Process Mixture (DPM) model which has a sequential estimation solution instead of dependency on previous measurements. However, DPM models are complex to implement and expensive to compute (Zhu et al. 2018). In the position domain, Ahmad, Sahmoudi, and Macabiau (2014) estimated the CDF of Horizontal PE (HPE) using empirical data and proposed few models that are good fits for specific situations. Rayleigh distribution was suitable for modeling HPE when there is open sky and Pareto distribution is found to be a good model for urban scenarios.

There are two main families of RAIM algorithms in the literature that are applied to urban scenarios. Measurement Rejection Approaches (MRA) remove faulty pseudorange measurements and ensure that localization solution and PLs are estimated using only valid measurements, thanks to Fault Detection and Exclusion (FDE) techniques. However, MRA algorithms generally perform poorly with low solution availability in urban environments compared to applications in civil aviation, since a high percentage of measurements will be excluded due to NLOS multipath (Cosmen-Schortmann et al. 2008). On the other hand, Error Characterization Approaches (ECA) characterize the measurement errors and compute protection levels based on the characterization, without the requirement FDE techniques. NLOS errors are considered in this approach leading to

higher protection levels. Although not enough to meet tighter civil aviation requirements, the high availability of solution makes ECA approaches useful in localization integrity monitoring for Liability Critical Applications (LCA) in urban environments (Cosmen-Schortmann et al. 2008).

### 1.3.4 Quantification of Localization Integrity Components

Up to this point, the integrity framework used in the localization of intelligent vehicles and its various computation methods and use cases are discussed in this chapter. However, most of these works focus on the integrity estimation processes under specific conditions or applications. One important component of integrity, accuracy, has been quantified in several works, but with respect to specific applications or scenarios. For example, in Basnayake et al. (2010), required accuracy of localization system for vehicle-to-everything (V2X) communication is estimated for different applications as-

- which-road:  $< 5.0$  m
- which-lane:  $< 1.5$  m
- where-in-lane  $< 1.0$  m

In a similar attempt, based on the meta-analysis of several processes in intelligent vehicles, Stephenson (2016) estimates requirements for the localization of current and future intelligent vehicles as-

- collision avoidance:  $< 2.0$  m
- fully-autonomous navigation:  $< 0.05$  m
- lane-departure warning:  $< 0.2$  m
- active vehicle control:  $< 0.1$  m

When it comes to the question of integrity requirements while designing or testing a localization solution, quantified values for integrity parameters are necessary. To this extent, only few works can be found in the literature, which provides definitive localization requirements independent of scenarios and applications.

Vehicle Type	Urban Roads			Highways		
	LatAL (m)	LonAL (m)	VAL (m)	LatAL (m)	LonAL (m)	VAL (m)
Mid-size	0.48	0.48	1.47	0.85	1.50	1.47
Full-size	0.42	0.42	1.47	0.80	1.50	1.47
Standard Pickup truck	0.38	0.38	1.47	0.76	1.50	1.47
Passenger Vehicles	0.33	0.33	1.47	0.72	1.50	1.47

Table 1.1: Quantification of Alert Limit requirements in highways and urban scenarios by Reid et al. (2019).

In Reid et al. (2019), the classical integrity concepts used in aviation are transposed to integrity requirements for ground vehicle localization. Using road safety related statistics and geometry of roads and vehicles, Reid et al. (2019) derived bounds for localization error in both highway and urban scenarios. They further distribute the derived total integrity risk to allocate integrity levels to every sub-system present in autonomous vehicles. They start with defining the target safety integrity level for the total navigation system, which is the maximum allowable probability of failures per hour. This is computed based on the statistics of current road safety information and desired improvements, and this is inspired from similar numbers used in aviation and rail navigation. Based on the geometry of the problem, ie. geometry of vehicles, road and lanes in both highway and urban scenarios, they produce Alert Limits for a localization system to meet the defined safety integrity level. This is achieved by allocating integrity risks based on the target safety integrity level to each subsystems of autonomous navigation, which includes perception, localization, planning and control. Their estimations are given in Table 1.1. Although, they have further narrowed down these estimations according to specific road geometry and speed limits in the USA, the values prior to this estimation which are listed in Table 1.1 are the one can be accepted globally.

## 1.4 Conclusion

In this chapter, a detailed discussion of the concept of integrity employed in different fields has been presented. Although arising from the same core definitions, integrity has been defined and treated differently in different scientific fields. Data integrity, information

integrity and system integrity concepts are outlined and their attributes and enablers are listed. The integrity of the information provided by a system is the combination of the integrity of the data and the integrity of the system that processes the data. It has been noted that while system integrity (most of the time integrity of processes) uses similar terminology as that of data and information integrity, their scopes are different. For example, while the availability of data is an enabler for data integrity, availability of a system is a property of that system that is estimated from its integrity. However, when transposed to the output of sensors and sources, data integrity and information integrity are practically the same, as any kind of meaningful representation of data is information itself.

The discussion of integrity concepts applied to intelligent vehicles in state-of-the-art methods provides an important perspective. Majority of the works in this area focus on the integrity of the localization system, specifically GNSS localization. As the measure of trust on the information provided by a localization system, protection levels (based on the confidence of the localization solution) and their upper-bounds (alert limits) are used. Since the ability to provide timely warning about the usability of the system is a requirement for the integrity of localization system in intelligent vehicles, the temporal (time-to-alert) and statistical (integrity risk) aspects of warning generation are also considered. Based on these attributes of the integrity of the localization system, integrity events such as unavailability of system and the presence of misleading information can be inferred. Conventional approaches of localization integrity monitoring like RAIM use physical or theoretical redundancy of measurements to detect, isolate and exclude faults from the measurements to support reliable localization. Here, it is observed that the 'consistency' enabler is quantified through the residuals of measurements and used to estimate the 'accuracy' and 'validity' attributes of the measurement data. It is worth noting that, for consistency check, measurements should be represented in the same domain by applying system-specific transformations. The technique of protection level computation used in RAIM also provides a key insight into the relationship between errors in localization to the position confidence. The sum of variances of errors in localizations from different combinations of measurements is the basis of position confidence which is used to provide protection levels.

Challenges faced by classical integrity monitoring solutions when applied to urban

---

environments have been discussed in this chapter. Since the assumption of single and distinct fault at any given time is not applicable to the urban environment due to multipath and NLOS errors, integrity monitoring methods have been proposed based on modified models for error modelling in both pseudorange domain and position domain. However, integrity monitoring of localization using GNSS in urban environments still remains a largely unsolved problem. On the other hand, few works have proposed the quantified upper-bounds for several components of localization integrity which is required for the safe operation of intelligent vehicles in various environments. This can be used in the performance evaluation framework for upcoming integrity monitoring solutions for localization systems which uses multimodal sensors and data sources. In the next chapter, multimodal localization of intelligent vehicles will be discussed and the different sensors and algorithms involved will be revised.



# Chapter 2

## What Is The Problem?

### Contents

---

<b>2.1</b>	<b>Introduction</b>	<b>24</b>
<b>2.2</b>	<b>Different Types of Localization</b>	<b>25</b>
2.2.1	GNSS Positioning	25
2.2.2	GNSS-Aided DR positioning	28
<b>2.3</b>	<b>Odometry</b>	<b>29</b>
2.3.1	Visual Based Odometry	30
2.3.2	LiDAR Based Odometry	35
2.3.3	Radar Based Odometry	36
<b>2.4</b>	<b>SLAM</b>	<b>37</b>
2.4.1	Filter-Based Approaches	38
2.4.2	Graph-Based Approaches	39
2.4.3	Particle-Based Approaches	40
<b>2.5</b>	<b>Map-Based Localization</b>	<b>41</b>
<b>2.6</b>	<b>Modeling Multimodal Data</b>	<b>44</b>
2.6.1	GNSS	45
2.6.2	Vision	47
2.6.3	LiDAR and Radar	47
2.6.4	Digital Maps	48
<b>2.7</b>	<b>Conclusion</b>	<b>50</b>

---

## 2.1 Introduction

Evolution has equipped each member of animal kingdom with remarkable sensing and planning abilities to navigate themselves in their respective living environment. Along with the five physical senses (perception senses), humans are said to have many self-awareness senses such as inertial sensing, force sensing and direction sensing etc. (collectively known as 'Proprioception') which help us to determine position and movement of our bodies in space. Combined with high-resolution visual sensing abilities, these senses help humans in efficient spatial localization and navigation in both allocentric reference frame (referenced outside of one's current body position, most often to multiple external landmarks) and egocentric reference frame (reference to one's current body position and using personal directional terms.) (Ekstrom and Isham 2017). On an abstract level, successful navigation process is the combined result of answering three basic questions:

- Where am I?
- What is around me?
- How do I safely get to my destination?

Clearly, the response to the third question is highly dependent on the availability and reliability of answers to the first two questions. When expanding this scenario to navigation of intelligent vehicles, these questions correspond to the tasks of localization, perception and path-planning respectively. Analogous to human navigation, different parts of perception are often combined with localization and navigation of intelligent vehicles. For example, lane detection can be used for localizing the vehicle accurately on road while obstacle detection can be used in collision avoidance in navigation. In this chapter, we discuss the different types of localization techniques used for navigation of intelligent vehicles. The primary focus is on the sensors and data sources used in these methods and how the data provided by them are treated and combined in each localization technique.

## 2.2 Different Types of Localization

Positioning a vehicle accurately in its environment is the fundamental element to several applications in Intelligent Transport Systems (ITS). Accurate self-localization of intelligent vehicles improves safety of navigation, efficiency of route planning, and comfort in driving experience by enabling Location-Based Services (LBS) like traffic management, fleet management and emergency response systems (Toledo-Moreo et al. 2018). In this section, we outline different types of localizations, the modalities of data sources used and their working principles.

### 2.2.1 GNSS Positioning

The global navigation satellite system (GNSS) is a radio positioning technology with global coverage based on satellite infrastructure. A GNSS system will be able to provide three-dimensional position (latitude, longitude and altitude) to its users worldwide. Even though several countries and entities have developed their own GNSS systems, on the core of their operation is a constellation of 18-30 Medium Earth Orbit (MEO) satellites that has specific trajectories and orbital planes. The first and most widely adopted GNSS system is American NAVSTAR Global positioning System (GPS), developed by USA and launched in 1978. It became fully operational in 1993 and globally made available in 1994. Since primary intention of GPS was military use, civilian signals were subjected to artificial degradation until May 2000 which reduced the quality of positioning. After this 'selective availability' was lifted, the Standard Positioning Service (SPS) available for civilian, commercial and research applications can provide positioning precision of 5 m. However, due to periodical retirement of old satellites and adoption of new signal bandwidth in the replacement satellites, GPS is capable of providing sub-meter accuracies since 2018 under ideal conditions. As of 2020, there are three more GNSS systems operational on a global level - GLONASS by Russian Federation, Galileo by European Union and BeiDou (COMPASS) by China. GLONASS became fully operational with 24 satellites in 2011 providing a localization precision of 4.5 m to 7.4 m and BeiDou, operational since 2018 with more than 50 satellites provides 3.6 m precision in localization for public use. Galileo positioning system which uses 26 satellites (30 up on completion), is able to provide 1 m precision for civilian uses and is operating since 2016.

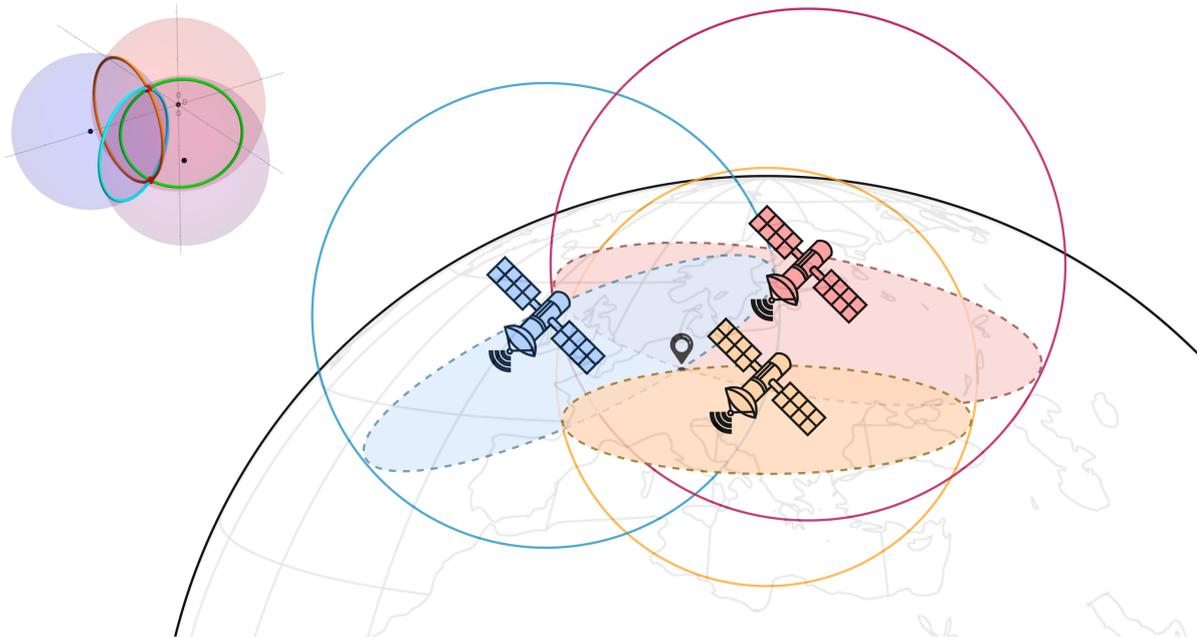


Figure 2.1: GNSS trilateration process

All GNSS systems work based on few common principles and architectures (Toledo-Moreo et al. 2018). The architecture of any GNSS comprises of three segments. *Space segment* consists a constellation of more than 24 satellites such that at least four satellites cover any point on earth at a given time. Ground segment controls the orbits and tracks the positions of satellites in the space segment by maintaining a two way communication. Finally, stand-alone modules (GNSS receivers) containing radio signal receivers and the signal processing units make up user segment that provides navigational data to the users. GNSS satellites typically transmits signals in the spectral range of 1.2-1.6 GHz and uses bandwidths of 2-40 MHz. Every satellite in a GNSS acts as a radio beacon which sends navigation message containing the position and time information using a unique format specific to each of them. These messages also include ionospheric parameters model to compensate the signal degradation and precise clock corrections to offset relativistic effects. GNSS receivers which can receive signals from at least four of such satellites can use the principle of true-range multilateration (trilateration) to produce a localization estimate. A unified global reference system for position and vector referencing is required to correctly position the calculated localization estimates. For example, World Geodetic System WGS-84 developed by the US Defense Department is used as the reference frame for GPS. (Toledo-Moreo et al. 2018).

Trilateration is a geometric problem of estimating the position of a point using its distances from three known points in space. Upon receiving a navigation message, GNSS receiver first computes the signal traveling time using the transmission time information in the message. A *pseudorange*, the relative distance between the receiver and satellite, is calculated by multiplying the signal travel time by the speed of light and correcting the relativistic effects. Measurement of a pseudorange from a satellite reduces the possible localizations on to the surface of a sphere centered at the position of satellite and radius equal to pseudorange. A second pseudorange from another satellite introduces an intersecting sphere which reduces the possibility of localization to a circle and the third one will reduce it to two points as shown in Fig. 2.1. The point near the ground plane is chosen as location and the other point in the space is discarded. However, in the real world conditions, receivers possess clocks which often have offsets which can degrade the signal travel time estimations. This can result in a volume of possible localizations instead of a unique point. To mitigate this ambiguity, trilateration always uses four or more satellites for providing the final localization estimate.

A crucial factor in safety-critical and in liability critical applications is the non-negligible flaws in satellite-based positioning that stay on while executing applications and services. Due to the fact that human life is at stake and full-proof performance is a must for accepting system deployment, GNSS performance problems cannot be overlooked. The problems such as lack of coverage, multipath errors, or jamming/spoofing are almost impossible to overcome without the support of extra information. Automated and cooperative automobiles need to imperatively know their accurate position, velocity, and direction to mindfully stay in the correct driving lane. This is due to the accuracy of the current GNSS positioning system which is well adapted for ITS applications like fleet management and vehicle tracking and control but not for ITS applications coming out in the foreseeable future. Therefore, as an inferred need for the acquisition of new radio access technologies for V2V communications, the accuracy in predicting their position is an indispensable pre-requisite for automated vehicles. The feedback given and inputs delivered to the control algorithm shall be attained by GNSS information, thus, adding a significant contribution to the control loop of the automated vehicles. In order to avoid collisions, GNSS can be an integral part of a short-range situation awareness system of other vehicles if accurate location data is shared among cars. For this reason, hybridized

localization technologies have been developed by combining GNSS systems with other sensors and data sources and used widely in navigation.

### 2.2.2 GNSS-Aided DR positioning

Similar to humans, current intelligent vehicles are also equipped with various sensors for proprioception. Sensors such as wheel and steering encoders can measure instantaneous distances and direction changes in the vehicle movements. These measurements can be integrated over time and used to predict the state of the vehicle (position and orientation) with respect to the previous state. This process is called Dead Reckoning (DR). Inertial sensors such as accelerometers and gyroscopes are also common elements in DR, that can measure motion and rate of motion along different axes. However, all proprioception sensors are susceptible to integration drifts caused by their inherent limitations. Encoders will measure the movement of the vehicle erroneously when wheel slides or slippages occur. Inertial sensors will accumulate minuscule errors in linear and angular acceleration measurements over time which can compound to large errors. Combining DR with GNSS localization systems can complement each other: DR can improve the accuracy of localization when availability of GNSS is poor and GNSS can correct the drift errors in DR process. DR sensors can provide positioning and velocity information at much higher rate than GNSS (often 10-100 times more) and they add a level of redundancy to the localization system.

Integration of instantaneous measurements from DR sensors is done using vehicle motion models. When it comes to conventional road vehicles, they follow nonholonomic motion model, ie., they have continuous closed circuit of the governing parameters for transformation between two states (Munguia 2014). In order to carry out fusion of DR sensors with GNSS, two main architecture classes are used: tightly coupled and loosely coupled (Toledo-Moreo et al. 2018). In loosely coupled fusion, state of the vehicle contains the position and heading and it will be predicted at each time step according to the motion model associated with DR sensors. An observation vector with GNSS fix is used to update the prediction to estimate the final pose. On the other hand, tightly coupled architecture uses pseudorange and Doppler measurements from GNSS receiver and measurements from DR sensors in the same observation vector to estimate pose

(Le Marchand et al. 2009). This allows the fusion to exploit GNSS measurement even if the measurements alone are not enough to provide GNSS fix and to use DR measurement to identify or isolate problems in GNSS measurement.

## 2.3 Odometry

Odometry refers to the process of tracking the pose of a body over time with respect to a reference frame. DR positioning without GNSS assistance is considered as a particular example of odometry in which only proprioception sensors are used. In this section, the focus is on odometry techniques which includes one or more perception sensors (exteroceptive sensors) such as cameras, LiDAR etc. along with the proprioception sensors. Odometry techniques can be classified in multiple ways based on the sensor modalities, motion model and data fusion techniques (Mohamed et al. 2019). Since this chapter focuses on sensors and how the data from them are treated, the classification presented here is based on those criteria.

Visual odometry (VO) uses only cameras and can be differentiated by the image processing method used: direct and indirect methods. In direct methods, whole photometric information in the images is used to compute the camera pose, whereas indirect methods extract the features such as corners and edges for this purpose. Number of cameras used is another classification parameter: one camera is used in Monocular Visual Odometry and two cameras with overlapping views are used in Stereo Visual Odometry. Two or more cameras can also be used without overlapping views as well, which is called Multi-view Odometry. Another important sensor modality is Visual Inertial Odometry (VIO) where any of the VO systems presented above is coupled with an IMU. Several methods are proposed with the additional use of GPS or replacement of cameras with RGB-D sensors (indoor environments), LiDAR scanners and RADAR (for outdoor environments) in VO and VIO systems to increase reliability (Mohamed et al. 2019). The classifications of odometry found in literature along with other self-localization concepts discussed in this chapter are presented in Fig. 2.2.

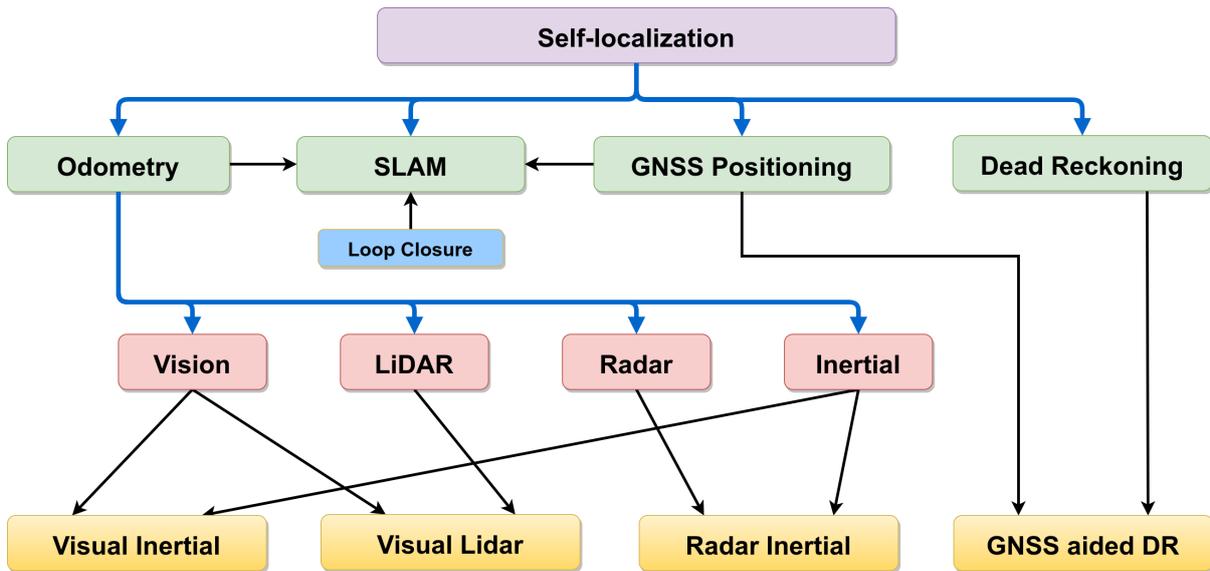


Figure 2.2: Classification of self-localization concepts inspired by the odometry classification in Mohamed et al. (2019)

### 2.3.1 Visual Based Odometry

Visual odometry (VO) is a technique to estimate the ego-motion of a body using one or more camera attached to it. The term VO was coined in Nistér, Naroditsky, and Bergen (2004) by taking inspiration from wheel odometry. Likewise wheel odometry, VO estimates the instantaneous poses using the images obtained from the camera and incrementally adding them up with respect to the initial pose. VO is superior to wheel odometry due to lower drift over time, but like any other vision based process, VO also depends highly on environment illumination and information available in images. VO imposes an additional requirement of having enough overlap between two consecutive frames (Howard 2008). However, VO or VO integrated navigation systems are probably the best solution for odometry in GPS-denied environments, such as underwater and indoor environments.

#### 2.3.1.1 Monocular Visual Odometry

In Monocular VO, 3D position estimation of the points as well as the motion estimation is done using consecutive images from a single camera. Although estimation of absolute scale is difficult in monocular VO without external information (scale ambiguity problem (Strasdat, Montiel, and Davison 2010)), the fact that it is less complex and less expensive makes it a widely used method in many areas. Several sparse feature based VO methods

(Corke, Strelow, and Singh 2004; Lhuillier 2005; Scaramuzza, Fraundorfer, and Siegwart 2009; Tardif, Pavlidis, and Daniilidis 2008), which use distinctive and repeatable features detected from images, employ RANSAC integrated 5-Point minimal solver proposed by Nistér (2004) to calculate the motion from 3D to 2D matching. Here 3D points are computed by triangulating from two consecutive images. Strasdat, Montiel, and Davison (2010) made improvements in regular monocular VO by adopting the key-frame and Bundle Adjustment (BA) (Triggs et al. 1999) optimization. Local windowed-bundle adjustment, which performs bundle adjustment after obtaining a specific number of frames, is used in Tardif et al. (2010) and Sünderhauf et al. (2006). Konolige, Agrawal, and Sola (2010) showcase a 10km VO test on a ground vehicle with windowed-bundle adjustment which showed up to 5 times more accuracy than without it. However, Tardif, Pavlidis, and Daniilidis (2008) decoupled the translation and rotation estimation by using different set of points for each, and proved that robust long term monocular VO odometry is possible without BA.

Since VO has a wide range of applications, state-of-the-art techniques mentioned above are adapted for particular applications, for efficiency and ease of implementation. Monocular VO has been adapted for ground vehicles in several works like Guerrero, Martinez-Cantin, and Sagüés (2005) by exploiting properties of homographies and ground plane estimation. Scaramuzza, Fraundorfer, and Siegwart (2009) adapted ego motion estimation for non-holonomic vehicles using one point RANSAC, which can speed up the estimation to several hundreds of times per second. Some methods have designed specific feature tracking methods for ground vehicles, for instance Scaramuzza et al. (2009) and Pretto, Menegatti, and Pagello (2011). There are a set of other works in the field of VO which focus only on robust scale information by adding minimalistic sensors such as laser distance meter (Wu et al. 2014), a priori knowledge of the movement (Scaramuzza et al. 2009) or distance of camera from the ground plane (Zhou, Dai, and Li 2016) etc.

### 2.3.1.2 Stereo Visual Odometry

The technique for estimating relative camera poses and recovering 3D model of an object from a set of consecutive images was known as SFM (Structure From Motion) in the computer vision community for a long time (Longuet-Higgins 1987). The formulation of VO framework using SFM concepts is first done by Moravec (1980) and initial phase

of VO research was done for Mars rovers by NASA (Moravec 1980; Olson et al. 2000). Moravec (1980) used corner detector in a 'slider stereo' VO system on a planetary rover. The rover would repeatedly move a small distance and stop, and a camera would capture 9 pictures at equidistant positions while sliding through a horizontal rail. The corners are detected and matched across these 9 frames and outliers are removed. The motion between two sets of images are computed as rigid body transformation by aligning the triangulated 3D points observed in each sets. This motion estimation pipeline used by Moravec (1980) paved the foundation of Stereo VO and is still used in several algorithms. By adapting this method Matthies and Shafer (1987) and Matthies (1989) used error covariance matrix for the triangulated 3D points from binocular cameras to model the uncertainty in motion estimation. Results of this procedure showed more accuracy than Moravec (1980) by reducing the error to 2% of the distance traveled. Adding an additional orientation sensor (compass) to VO was proved to obtain better results (error: 1.2%) by reducing non linear increase of accumulation error to linear, in Olson et al. (2000) and Olson et al. (2003).

Unlike all the methods mentioned above, which use 3D to 3D matching for motion estimation, Nistér, Naroditsky, and Bergen (2004) developed a 3D to 2D matching algorithm which also proved to be the first real-time online implementation of VO. Instead of tracking the selected features over frames, they detected Harris corners in each frame and matched between features. Combined with this detection scheme, they used structure to image correspondence to estimate the motion, which is a routine now known as Perspective from N points (PnP). Here, triangulated points from a stereo pair are matched against their projections in the consecutive frame. Another different motion estimation technique which uses image points correspondence (2D to 2D matching) was introduced by Comport, Malis, and Rives (2007). The idea is to use 2 stereo pairs to find the matching points in them and estimate the motion directly without the need of any triangulation. They proved that it is more accurate to use image points than 3D matching but the main disadvantage is that finding the matching points in all four images simultaneously is computationally costly.

### 2.3.1.3 Direct Visual Odometry

Direct Visual Odometry methods are defined as methods for shape or pose estimation using every pixel intensity information available in images (Irani and Anandan 2000). Contrast to feature based (indirect) methods, direct methods minimize an alignment error measure defined on parameters like image brightness, brightness-based cross-correlation etc., to estimate unknown parameters of camera motion. Majority of indirect methods uses points as features, which inherently discard the other significant information such as edges and shapes. Even though indirect methods that use edges (Klein and Murray 2008) or regions (Concha and Civera 2014) exist, they have higher computational cost due to high-dimensional feature space. Direct visual odometry (VO) methods address all these issues by estimating geometry directly on the image intensities, while providing higher accuracy and robustness (Engel, Schöps, and Cremers 2014).

Direct visual odometry has nearly-complete scalability, which means irrespective of environment conditions, they can estimate substantially more information about the geometry of the environment. This property makes it very important for autonomous navigation and augmented reality applications. Kerl, Sturm, and Cremers (2013) used RGB-D sensors whereas Comport, Malis, and Rives (2007) used stereo vision system for direct image alignment on depth images. In monocular VO, Newcombe, Lovegrove, and Davison (2011) developed a dense tracking and mapping (DTAM) framework, which exhibits high quality pose estimation in real time for moderate movements. Forster, Pizzoli, and Scaramuzza (2014) formulated a scheme to combine direct method with indirect method- keypoints-aided dense odometry - with real time capability.

### 2.3.1.4 Visual Inertial Odometry

It is observed that visual odometry often needs an additional sensor to overcome the poor performance in challenging environments. Visual Inertial Odometry (VIO) is developed as an improvement from VO on these grounds, where IMU is integrated with classic VO frameworks. The main challenge in efficient IMU data integration with Visual data, is that IMU's acquisition rate is several times larger than that of cameras. Hence based on this integration, VIO systems can be classified into two, tightly-coupled and loosely-coupled systems.

In general, IMU measurements and visual pose estimations are treated independently in loosely-coupled systems. Konolige, Agrawal, and Sola (2010) and Forster et al. (2017) incorporated pre-integrated IMU data between two image acquisitions to optimize the VO pose estimation. On the other hand, Weiss et al. (2012) uses VO pose estimation to update the pose predictions by IMU using an Extended Kalman Filter (EKF). Loosely-couple methods allow to integrate IMU with existing VO or visual SLAM methods such as LSD-SLAM (Engel, Schöps, and Cremers 2014) or ORBSLAM (Mur-Artal, Montiel, and Tardos 2015). A generalized sensor fusion algorithm using EKF and Unscented Kalman filter (UKF) specifically for multi-rate sensors is proposed by Armesto, Tornero, and Vincze (2007). This avoids the need of pre-integration of IMU data, and holds IMU data by performing only prediction until next vision data arrives for the update. However, Armesto, Tornero, and Vincze (2007) assumed constant linear acceleration model for the system dynamics. At high rates, IMU can sense the dynamics of the system better than assumed models, which inspired Sirtkaya, Seymen, and Alatan (2013) to use an error propagation model for fusion, which allows to model different characteristics of IMU. There are several other methods such as Tardif et al. (2010) and Diel, DeBitetto, and Teller (2005), which used IMU data for bundle adjustment of poses generated by VO, instead of fusing using filters.

In contrast, tightly-coupled methods process the raw information of both sensors in a single estimator, thus all the correlations between them are considered, leading to higher accuracy (Mourikis and Roumeliotis 2007; Leutenegger et al. 2015). One way to achieve this is by incorporating 3D feature positions to the system state used in the filter that estimates pose (Kelly and Sukhatme 2011). But adding features to the state vector when they are observed dynamically increases the dimensionality of state vectors and other matrices involved. This causes high computational cost to implement tightly coupled systems. A sliding window EKF filter is used by Mourikis and Roumeliotis (2007), which estimates a finite number of poses constrained by the features observed in that particular window. Trifocal tensor based sliding window optimization method, which exploits the geometry constraints among three consecutive images, are also demonstrated as viable tightly-coupling method in Hu and Chen (2014). Kong et al. (2015) not only used points but also lines observed from the scene in an EKF based fusion scheme and showed improvement in accuracy. Among direct visual inertial methods, Bloesch et al. (2015)

combines direct VO with IMU in KF framework, while Leutenegger et al. (2015), Jones and Soatto (2011), and Keivan, Patron-Perez, and Sibley (2016) propose energy-function based optimization framework.

### 2.3.2 LiDAR Based Odometry

Light Detection and Ranging (LiDAR) works by emitting coherent and spatial laser light from a laser transmitter and detecting the reflected laser beams on a 2D observation plane on optical detectors. Laser odometry (LO), or LiDAR odometry, is an approach for estimating the position and orientation of a platform by tracking these laser patterns reflected from surrounding objects (Mohamed et al. 2019). The fact that LiDARs are not sensitive to environment illumination and textures introduces a promising prospect in ground vehicle perception systems which are traditionally dominated by vision systems. The adoption of Laser technology is accelerated by modern LiDARs, which have evolved to be cheaper, smaller and more efficient.

The standard iterative closest point (ICP) method is a common method used in LO to identify the motion between two consecutive LiDAR scans by minimizing the distance between corresponding points. Generalized ICP is a combination of standard ICP and point-to-plane ICP proposed by Segal, Haehnel, and Thrun (2009) and used in works like Saito, Watanabe, and Nagai (2015) and Zhuang et al. (2011). However, LiDAR scanning frequency could be relatively slow such that the extrinsic motion of the platform cannot be estimated due to motion artifacts. In this case, other sensors such as cameras and dead reckoning sensors are used to estimate velocity (Zhuang et al. 2011; Yan et al. 2017). In LiDAR Odometry and Mapping (LOAM) proposed in Zhang and Singh (2014), planar surfaces and edges are detected efficiently from point clouds and registered as a visual image so that tracking is accurate and efficient.

Despite the development of efficient LiDAR processing pipelines and matching algorithms, LO remains as a resource-intensive operation to be implemented in on-board computers on vehicles. Also, non-reflective surfaces such as glass and black surfaces pose challenges for LO (Horn and Schmidt 1995). Fusion of LiDAR with vision is considered in literature to overcome the limitation of motion artifacts (Zhuang et al. 2011; Yan et al. 2017; Horn and Schmidt 1995). In Yan et al. (2017), a dense 3d reconstruction

scheme (VISO2) (Geiger, Ziegler, and Stiller 2011) is used to calculate the transformation between two consecutive LiDAR scans and distortions caused by motion. The result of this reconstruction is used as the initialization for localization procedures by LOAM by extracting surface and edge features in the undistorted point clouds. To reduce the computational demand when using 3D LiDAR scans in ICP and 3D data registration, Zhuang et al. (2011) proposes a bearing angle (BA) model to convert the 3D LiDAR data to a two-dimensional BA image, which is an optimal way for feature extraction and matching (Mohamed et al. 2019).

### 2.3.3 Radar Based Odometry

Radio Detection and Ranging (radar) sensors use radio frequency waves to scan and determine the distance, velocity and angle of objects in its surroundings. In the field of localization and navigation by obstacle avoidance, frequency modulated-continuous wave (FMCW) format of radar technology is used. In this format, the sensor transmits a steady stream of linearly modulated continuous radio signals and a detector senses the reflected signals and generates a high resolution image of the environment. Due to low sampling rate, low power consumption and minimum target range, this technology has drawn a lot of attention in the recent years (Quist, Niedfeldt, and Beard 2016). Radar Odometry (RO) attempts to estimate the relative motion of radar sensor based on the environment information provided by these scans (Mohamed et al. 2019). Unlike visual odometry, radar odometry can be performed independently of weather conditions and texture information available in the environment but the steps used in RO are similar to the steps in VO, ie. feature extraction and tracking (Mohamed et al. 2019).

To the feature extraction end, classical visual feature extraction techniques such as SIFT (Quist and Beard 2016) and Hough transforms (Callmer et al. 2011) are used to detect features points and shapes in range-compressed grayscale images created from radar scans (scatters). In order to decrease the effect of noise on scatter images, thresholding is used in Quist, Niedfeldt, and Beard (2016). A density-based spatial clustering technique (DBSCAN) proposed by Ester et al. (1996) is also often used to detect regions of interest (ROI) in radar scans. On the other hand, landmarks are extracted using the power-range spectra of radar scans in Cen and Newman (2018). In the tracking step, extracted features

are tracked using several adapted techniques in the field of object tracking in computer vision. For example, Quist, Niedfeldt, and Beard (2016), used a recursive-RANSAC algorithm to track feature points. A binary annular statistic descriptor (BASD) is applied, in Schuster et al. (2016), to the extracted feature points and tracked based on Hamming distance. Global scan matching techniques inspired from ICP, which minimize a cost function in the alignment of extracted features are employed in works like Vivet, Checchin, and Chapuis (2013). In Cen and Newman (2018), an algorithm to perform data association using a feature descriptor and relationships between features is proposed to track features between multiple scans. Unlike ICP, this approach does not rely heavily on a good initial estimate (Mohamed et al. 2019).

One of the main problems with RO measurements is that, it is highly susceptible to outliers on uneven terrains. This problem has resulted in several outlier rejection schemes, by fusing radar data with other data sources such as IMU and Vision. Combining vision and radar can have complementary effect on understanding the environment well, where vision provides high resolution but inconsistent data (due to weather conditions etc.) from the scene and radar provides consistent data but with a lower resolution (Mostafa et al. 2018). Along with camera and radar, Mostafa et al. (2018) fuses IMU, barometer, and magnetometer in a loosely-coupled EKF framework to estimate position and velocity accurately. In a similar fashion, IMU data is fused with radar in Ghabcheloo and Siddiqui (2018) to obtain the robust forward and angular velocities of a ground vehicle.

## 2.4 SLAM

SLAM (Simultaneous Localization and Mapping) is the inference process of building a globally consistent map of an unknown environment and estimating the pose of the robot with respect to that map at the same time. In fact, several of the works mentioned in Section 2.3 contain sparse or dense mapping modules along with odometry (Jones and Soatto 2011; Kelly and Sukhatme 2011; Newcombe, Lovegrove, and Davison 2011; Vivet, Checchin, and Chapuis 2013; Zhang and Singh 2014). However, the keyword that differentiates modern SLAM techniques from odometry is 'loop-closure', which is the process of identifying previously visited locations in the map and hence improving the consistency of the map and limit drift errors. SLAM architectures possess loop-

closure detection and map correction but odometry does not have any such feedbacks (Cadena et al. 2016). While recent odometry algorithms are consistently able to reduce the drift observed in position tracking to less than 0.6% of the total trajectory [KITTI leader board], the research on SLAM techniques enabled the study of sensor fusion in autonomous vehicles under much more challenging conditions (Cadena et al. 2016). Also, an odometry based navigation system working in a large-scale city environment will eventually fail to understand the actual topology of area due to the lack of loop-closure. The problem will be worsened in urban canyons. Because of the poor GNSS signal availability, there will not be any reliable way to estimate the vehicle's position in global reference frame. Although simple place recognition may seem sufficient to get the right topology of the map, without the metric information provided by the SLAM module, odometry algorithms may run the risk of wrong data association (Latif, Cadena, and Neira 2013; Cadena et al. 2016). Because of the similarities in sensor combinations and implementation of SLAM with different odometry techniques mentioned in section 2.3, this section only focuses on the three basic paradigms used in solution of SLAM problem.

### 2.4.1 Filter-Based Approaches

Filter-based SLAM solutions work on the assumption that a sparse feature-based environment representation is possible. In this representation, features from the environment are points/lines/structures in an applicable parameter space. These features and the estimates of localization of the vehicle will constitute the state vector. The uncertainty in the estimates of both location and features is managed using error covariance matrix. Using the start position (usually known in global frame) as the origin of the reference frame, vehicle moves and senses nearby features/landmarks. An Extended Kalman Filter (EKF) (Jazwinski 2007) is used to update the state and covariance matrix as the vehicle obtains new measurements in the environment according to its position. Odometry computations are carried out between every successive time steps and increase in pose uncertainty is estimated. The measurement uncertainty is combined with pose uncertainty to get the uncertainty of the map, which increases over time. However, the revisit of any of the previously observed features reduces the uncertainty of all the features in the map. This occurs due to the correlation in position estimates of all the features.

This effect, where any measurement that decreases the pose uncertainty of the vehicle in back-propagated through the map, is recognized as the most important characteristic of SLAM posterior (Stachniss, Leonard, and Thrun 2016). EKF-SLAM also uses proximity based reasoning to solve data association problem, in case of unknown identity of detected features. By considering uncertainty associated with each feature and the noise in the measurements, a weighted quadratic distance (Mahalanobis distance) is used to find the most suitable data association. More recent algorithms only add a feature into the long-term map if it is observed with enough frequency. While filter-based SLAM approaches are used extensively in robotic applications in challenging conditions, such implementations are rare in the area intelligent vehicles operating in highly dynamic urban environments. The main problem with scaling filter based approaches to large area is that the size and numerical conditioning of covariance matrix of the system grows quadratically with every new feature added to the map. On the other hand, unique distinguishable features in urban environments are scarce due to the similar architecture of building and other infrastructure.

### 2.4.2 Graph-Based Approaches

While filter-based methods model the SLAM problem as an online state estimation where the vehicle state and some part of the environment is updated 'on the go' as new measurements become available, graph-based techniques estimate the full vehicle trajectories from the complete set of measurements. In recent years, pose-graph optimization has become the standard for most modern SLAM solutions. In this approach, movement of the vehicle is represented using a graph where every node corresponds to a pose of the vehicle during mapping. An edge connects two of such nodes, which corresponds to the spatial constraint between them as suggested by the odometry. As vehicle moves, such odometry constraints connect the consecutive poses and create a map. Since these poses contain uncertainty due to the noise or errors present in odometry measurements, the motion constraints connecting them are 'soft constraints' (Stachniss, Leonard, and Thrun 2016). Likewise, each feature/landmark/obstacle observed at each pose of the vehicle also introduces nodes and soft measurement constraints according to the uncertainty in environment perception. Eventually, when the vehicle observes a previously seen feature,

it generates constraints between non-successive poses. Hence, the heart of graph-based SLAM is the optimization process that finds a node configuration that minimizes the error introduced by all these constraints when a loop-closure is detected (Grisetti et al. 2010). The data association problem is addressed in loop closure by aligning the current map measurements with the existing map and adding the transformation between current pose to a possible previous pose as a constraint. Understandably, most of the graph-based SLAM methods are offline ie. the map generation and robot path optimization is done globally. As the length of vehicle's path increases, the optimization process will be more computationally expensive. However, recent approaches such as ORB-SLAM (Mur-Artal, Montiel, and Tardos 2015) carry out a local bundle adjustment which enables them to perform online.

### 2.4.3 Particle-Based Approaches

Particle based methods use approximation method to represent a distribution instead of probability distribution used in filter based method. A particle filter is a method in which the belief on the position of a vehicle is represented by a set of particles. Each particle in this set corresponds to a localization hypothesis of the vehicle and maintains its own map. The process begins with a uniformly weighted distribution of particles in the area of operation. The motion model of the vehicle is then used to predict the state of each particle based on the control command or dead-reckoning. Due to the noise in these, the uncertainty of the predicted particles increases. The observation of the environment is carried out at this step and the state of the vehicle is updated based on the match between global map and current observation of the environment. Based on the difference between predicted and actual observations, an importance weight is calculated for each particle. The particles are periodically re-sampled according to their importance weight. Similar to filter-based methods, as dimensionality of map and robot path increases particle filters also become computationally expensive due to the increased number of particles required for the representation of all the dimensions (Stachniss, Leonard, and Thrun 2016). However, this problem is addressed in FastSLAM proposed by Montemerlo et al. (2002), by factorizing the posterior into a product of conditional landmark distributions and a distribution over robot paths. This allows the algorithm to

recursively estimate the full distribution over robot pose and landmark locations, while the complexity scales logarithmically instead of quadratically. Montemerlo et al. (2002) claims online operation in large scale environments.

## 2.5 Map-Based Localization

Similar to odometry and SLAM techniques, introduction of Geographic Information Systems (GIS) in navigation systems also started in the field of robotics. Filliat and Meyer (2003) defines two types of localizations strategies when map is available.

Local localization - providing a new position estimate based on a previous estimate and current perception information. Here, the vehicle's position estimate should be corrected in order to get better fit to its perception on the given map information.

Global localization - providing a position estimate based on current perception information. Here, the vehicle position is the location in the map where its perception fits the best. Rather than correcting the position, it is more of a choice between distinct position hypotheses (Filliat and Meyer 2003).

In either case, the technique used is called 'map-matching', where correlating a position generated by a sensor system with information from maps is the key. Poor positioning information and inaccuracies present in maps makes this a rather difficult task (Toledo-Moreo et al. 2018). Based on the type of map data and method used, map-matching algorithms can be divided into three categories: geometric methods, topological methods, and advanced methods (Quddus and Velaga 2012).

Geometric methods use the metric representation of maps and matches a localization to the geometric information available in the map. The earliest method used in geometric map-matching is point-to-point matching (Bentley and Maurer 1980), which finds the nearest map node to a given localization estimate. Despite the speed and ease of implementation, this approach is highly dependent on the digitization of the map. This method will perform better if a section of map has more nodes (eg: curved roads) contrary to a section with fewer nodes (eg: straight highway) (Quddus and Velaga 2012). Point-to-curve matching methods are used in literature, where a location fix from the navigation system is matched to the closest curve on the map. Distances from the position fix to all the line segments that form a curve (road) are computed and

the line segment with shortest distance is chosen. However, urban environments with high road density pose several challenges to this method as the closest line segment doesn't necessarily guarantee the correct solution (Quddus and Velaga 2012). When the history of vehicle localizations is available, the use of curve-to-curve matching is preferred. In this technique, a set of closest nodes to a given location fix is found out using point-to-point matching. Then for each of these nodes, a subset of the curves that includes that particular node is extracted. A curve is created using the previous location fixes (trajectory) of the vehicle and it is matched with curves created with each candidate nodes (White, Bernstein, and Kornhauser 2000). The curves with smallest distance between them is chosen as the road in which vehicle is traveling. Evidently, this method can be inconsistent depending on the performance of initial point-to-point matching, which can be poor when outliers are present in the localization fixes.

Topological representation of the map, which represents the relationship between map elements such as adjacency, link connectivity etc., is used in topological map-matching methods along with the geometric information (Toledo-Moreo et al. 2018). In this method, historical fixes and matching information is used to increase the positioning accuracy and maintain the logical continuity of the trajectory. In these methods, for the correct link identification at junctions and other complex scenarios, additional information such as vehicle speed, heading etc. is also utilized. After the initial phase of map-matching, the candidate nodes are extracted and the links connected to those nodes are weighted according to the similarities in link direction to the vehicle direction, the continuity from the previously estimated link and turn restriction information. Based on the sum of the weights obtained from these different criteria, the correct link is selected and a perpendicular projection from the current positioning point to this link provides the localization of vehicle (Quddus and Velaga 2012).

It is easy to note that in both geometric and topological map-matching methods, similarity and distance based models are used. In the recent years they fell out of favor due to weak performance and new methods have developed which employ multiple models and new information sources (Chao et al. 2020). Such methods, categorized under advanced methods (Quddus and Velaga 2012), are classified further in Chao et al. (2020) based on the matching models used. *State-Transition models* use a weighted graph where vertices represent the possible state of the vehicle and edges represent the transitions between

these states. The weights of the graph are the possibility of states or transitions and optimal path is generated using techniques such as Hidden Markov Model (HMM) and Conditional Random Field (CRF). *Candidate-Evolving models*, as the name suggests, consider a set of hypotheses while map-matching. A candidate set is initiated by a geometric or topological matching and evolved as new position fixes obtained from other sensors. Here, each candidate subset will belong to the possible links in the map and any link without enough candidate support is pruned periodically. The link with most candidate support is chosen as the correct link.

Davidson, Collin, and Takala (2011) used a particle filter (PF) for this purpose and shows superior performance in localization of the vehicle in complex environment with only DR and GPS. Lu et al. (2014) used lane marking detections from vision to match with a map and maintain the set of possible matching locations as the particles in a PF to track the localization over time. Finally, *Scoring models* assigns a set of candidates to each location fixes and maximizes a scoring function to find the best candidate. In Sharath, Velaga, and Quddus (2019), each location fix is used to find possible cells in a virtual grid created using data from map and vision. These cells are scored using a combination of weightings based on proximity, connectivity, kinematic and turn-intent prediction and maximized to find the best localization fix.

In Ballardini et al. (2016), facades of buildings at intersections are detected using vision and fused with building footprints extracted from the digital map to provide better localization. They further extended their work in Ballardini, Cattaneo, and Sorrenti (2019) to achieve localization at intersections using road structures instead of building facades and map data. A map-matching based localization involving lane detection from vision is used in Kang et al. (2020) and Nedeveschi et al. (2013). Similar strategies are employed combining digital maps with features detected from LiDAR data such as, curb detections (Liu et al. 2019), intersection structure detection (Liang, Zhang, and Wang 2017), lane detection (Mueller et al. 2011) etc.

The discussion on map-based localization cannot be concluded without mentioning the recent advancements in enhanced maps and High Definition (HD) maps. As described in the previous sections, localization in complex environment is often solved by introducing data from multiple perception sensors. Sensors such as cameras, LiDARs and Radars, are useful in finding lane markings, road-barrier structures, landmarks, buildings and traffic

signs. However, integration of a classic map with only metric or topological road network information (coarse road geometry) with such an information-rich perception system is generally considered futile. Hence, the concept enhanced maps were developed, in which a classical map will be added with one or more layers including:

- Fine road geometry - road width, position of lanes, lane-width, connectivity of lanes, other road markings, geometry of lane merges and splits etc.
- Navigation constraints - position and contents of traffic signs, traffic flow directions, speed limits etc.
- Surrounding structures - positioning and geometry of sidewalks, dividers, railings
- Infrastructure - Building footprints or their 3D structure.

While works like Jabbour, Bonnifait, and Cherfaoui (2006) focused on identifying the landmarks that can be used for better localization and constructing enhanced maps, most recent works directly uses HD maps. These are commercially available or ready-made (pre-processed) enhanced 2D or 3D maps with detailed representation of physical environment. While Liang, Zhang, and Wang (2017) use fine road geometry from the map with LiDAR curb detection to localize in urban intersections, Nedevschi et al. (2013) and Kang et al. (2020) use lane marking and road-sign detections using camera with road-sign details from an enhanced map to achieve the same. Rabe et al. (2017) combines visual lane marking detections with radar based object detections to match on to an HD map for urban localization.

## 2.6 Modeling Multimodal Data

Up to this point in this chapter, we have established the significance of multimodal data sources in intelligent vehicle localization and the different 'modes' in which they are used for this task. However, an examination of models applied to data from these sources will be beneficial in the understanding of their source level interactions in the total system. A brief definition of 'model' in the context of information science and systems engineering is given here to provide some insights to this discussion. Kühne (2005) observes a 'model' as something that satisfies three features:

**Mapping feature** A model is based on an original (original can be physical, possible or imaginary objects, data and systems)

**Reduction feature** A model only reflects a (relevant) selection of the original's properties (implies simplified representation, abstracted description or interpretation of the original. A copy is not a model).

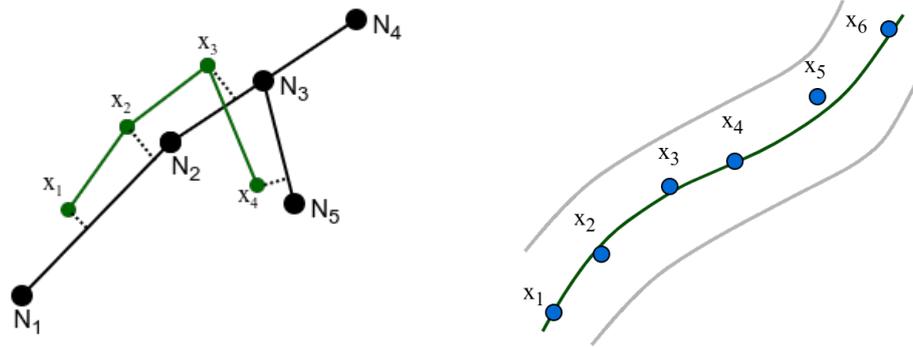
**Pragmatic feature** A model needs to be usable in place of the original with respect to some purpose (depends on the purpose of the original in a specific use case).

Evidently, it is a wide “umbrella” definition of ‘model’ that encloses all possible definitions that may be used in all the domains concerning this work. In this section, we identify the models applied on each data source used in above localization strategies, that fit the definition outlined above and present a review on them.

### 2.6.1 GNSS

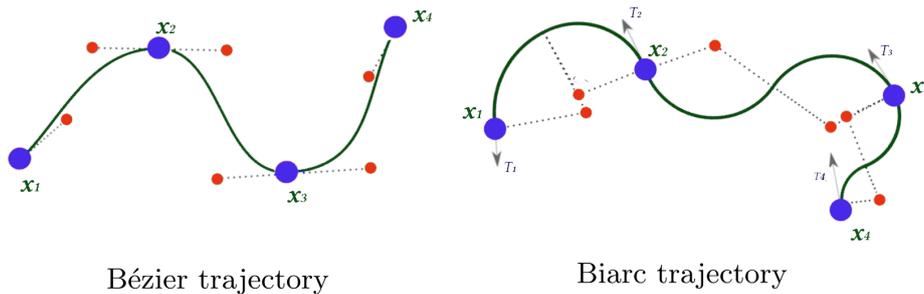
GNSS localization uses several models in pseudorange domain based on satellite constellation, signal propagation parameters, receiver structures and multipath geometry. However, modeling for GNSS localization fixes in spatial (positioning) domain is rare in the literature. The main idea of modeling GNSS tracks comes from the curve-to-curve geometric map-matching methods. Such methods model a history of localization fixes as a piecewise-linear curve and use it to match with similarly piecewise linear curves obtained from map (White, Bernstein, and Kornhauser 2000) as shown in Fig.2.3a. The same concept is used to model GPS trajectories and interpreting road structures in Li et al. (2017) and several algorithms assume piecewise-linearity of GNSS measurement to reduce the size of historical GNSS trajectory data for storage.

Although not directly used in localization, Jeon et al. (2015) notes that piecewise-linear model is not necessary in waypoint modeling and a curve fitting approach can be used. They observe from the literature that every highway road structure and vehicle motion model satisfy clothoid constraints and propose the use of a 3rd degree polynomial curve fitting to reliably model GPS trajectory data as shown in Fig. 2.3b. This forces them to use more than four GPS positions as minimum of four points are required to estimate a 3rd degree polynomial. Vishen, Silaghi, and Denzinger (2015) propose Bezier



(a) Piecewise-linear trajectory models from White, Bernstein, and Kornhauser (2000). Green dots represent GNSS fixes and green polyline is the piecewise-linear trajectory used to match with curves in the map (black polyline).

(b) Clothoid curve fitting trajectory models from Jeon et al. (2015). Blue dots are the GNSS fixes and green third degree polynomial is the trajectory model estimated using curve fitting.



(c) Bezier and Biarc trajectory models from Vishen, Silaghi, and Denzinger (2015). Blue dots are the consecutive GNSS fixes and green curves are the estimated trajectory models using Bezier and Biarc curve estimation. Red dots are the control points obtained from respective curve fitting algorithms.

Figure 2.3: Different GNSS trajectory models

curves and Biarc curves to model GPS fixes (Fig. 2.3c) to interpolate full trajectory. Bezier curve is defined between two consecutive GNSS fixes using two control points which ensure the continuity of total trajectory. Biarc curves, also known as piecewise-circular curves, are determined by two consecutive GNSS fixes and their tangent vectors. Using RMS error between ground truth and estimated trajectory, they show that both models perform accurately if the distance between the consecutive GPS points are less ( $<30$  cm positional error and  $<0.5$  rad orientation error if distance between GPS points are below 5 m).

### 2.6.2 Vision

Since visual data can be used to detect different kinds of features and semantic information, the modeling of this data source is also dependent on particular use cases. Conventionally, visual odometry detects feature points hence camera projection models, 2D to 2D, 2D to 3D and 3D-3D correspondence models are used to model data from vision source. However, recent multimodal localization methods detect complex features such as lane markings, road signs and road geometry using vision data. In case of lane markings, Mammeri, Boukerche, and Tang (2016) and Kang et al. (2020) use hough transform to fit straight lines on to lane markings. Here the assumption is that at short distances, lane markings will be straight and parallel. However this assumption is not valid in many highway scenarios and this model can only perform well for lateral localization and orientation estimation. Lu et al. (2014) use quadratic polynomial for modeling the lane marking and shows better total localization performance including longitudinal directions. On the other hand, Tao et al. (2013) employ 3rd degree polynomial model for lane markings recalling the clothoid constraints of road curvatures in highway scenarios and claims localization accuracy of 1.21 m of horizontal position error.

In the area of mapping using vision, an important model is occupancy grids. Li and Ruichek (2014) address the issue of dynamic urban environment mapping using the data from stereo-vision system to populate occupancy grid. The occupancy grid model is created using U-disparity maps (to remove moving obstacles) and V-disparity map (estimating pitch angle) which are created using stereo camera geometric model. Yu, Cherfaoui, and Bonnifait (2015) further improved this concept by creating two separate grids from U-disparity and V-disparity maps to combine them in an evidential occupancy grid model. U-disparity map is used to create an obstacle confidence map and U-grid is generated. V-disparity map is combined with a ground plane model to create a V-grid based on the map of ground plane estimation confidence. Later, U and V grids are fused using Dempster's fusion technique into a single stereo occupancy grid.

### 2.6.3 LiDAR and Radar

Similar to vision, modeling of LiDAR and Radar data depends on the type of features being detected and used in a process. Zhuang et al. (2011) use point features detected

using SIFT from a 3D LiDAR data, hence they use a bearing angle model. In LOAM method proposed by Zhang and Singh (2014), geometric planar models are used to detect and represent surfaces in the environment. They also detect edges from the point cloud and estimate straight line models to compute motion between each frames. In order to create enhanced map, Jabbour, Bonnifait, and Cherfaoui (2006) detect the edge of sidewalks as landmarks and models them using straight lines. In the localization method proposed by Hsu and Shiu (2019), a 2D grid map is created by segmenting surfaces that are orthogonal to the the road plane. They divide the 2D ground plane into 10 cm x 10 cm cells and determine the occupancy of the each cell based on the number of 3D points available in the vertical column corresponding to the cell. Features are detected from this 2D grid for localizing the vehicle.

Moras, Cherfaoui, and Bonnifait (2011) propose a Scan-grid to represent the LiDAR data by using polar space representation of LiDAR data. The coverage space of LiDAR is divided into angular sectors and each section is divided into a number of cells. This polar scan grid is created using a credibilist approach framework based on Dempster-Shafer theory. Later, using the pose information, polar grid is converted to scan-grid and mobile objects are detected. Capellier et al. (2018) use the approach of evidential occupancy grid construction, by combining grids from segmented ground and obstacles as well as previous grid computed. They named this model Ego-grid and the estimated pose transformation between previous position fix and current fix is used to fuse consecutive ego-grids. They also propose the similar method for vision data using image segmentation and propose asynchronous fusion of LiDAR data and Vision data into a single evidential grid. Likewise, radar data is also represented in the literature using occupancy grid and its variations (Cen and Newman 2018).

#### 2.6.4 Digital Maps

Digital geographic maps contain abstracted representations of the entities in physical world and their relationships. As per the definition presented in the beginning of this section, Kühne (2005) confirms that all maps are a models in a broad sense.

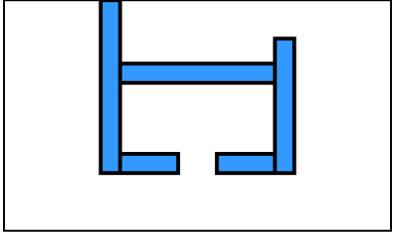
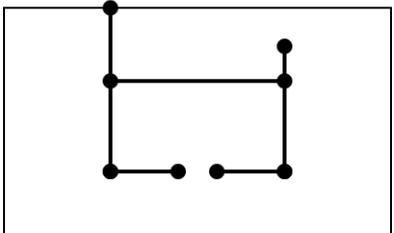
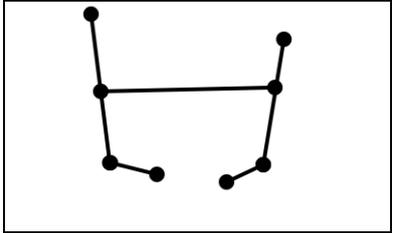
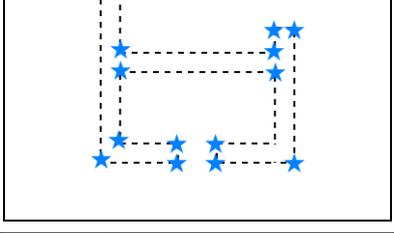
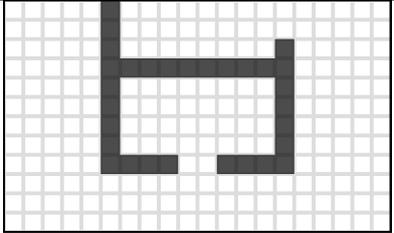
Example of original environment		
Model	Properties	Visualization
Metric	<ul style="list-style-type: none"> <li>- discretization of physical landmarks in a geometric frame</li> <li>- positions, distances and angles are preserved</li> </ul>	
Topological	<ul style="list-style-type: none"> <li>- represents possible states and the transition between them</li> <li>- preserve connectivity, adjacency etc. of links</li> </ul>	
Feature (metric)	<ul style="list-style-type: none"> <li>- representation using features such as corners, edges, areas etc.</li> <li>- accuracy depends on sensor models used to detect features</li> </ul>	
Free-space (metric)	<ul style="list-style-type: none"> <li>- discretization of physical space using labeled cells</li> <li>- represent parts of maps accessible for navigation and obstacles</li> </ul>	

Table 2.1: Different representations used for map data

Hence, without loss of generality, we can establish that different digital map models are nothing but different types of representations of geographical information and relationships. Filliat and Meyer (2003) present two main types of map representations used for navigation - metric representation and topological representation, which are the basis of map-matching techniques presented in sec. 2.5. A summary of all the map representations used in the works mentioned in this chapter is presented in the Table 2.1.

## 2.7 Conclusion

This chapter presents the multimodal localization problem in the field of intelligent vehicles. The principles and techniques of GNSS localization are discussed, and its scope and applications are highlighted. The discussion on multimodal localization begins with the process of combining information from dead-reckoning sensors and the GNSS information, which increases the accuracy of localization due to their complementary nature. Another two major localization approaches, odometry and SLAM, are studied with the focus on different sensors and data sources used in them. It is observed that perception sensors such as camera, LiDAR, radar etc. are extensively used in both SLAM and odometry approaches. Majority of these works detect features such as interest points, lines and the regions from the sensor data to match and combine information from different sources. Both of these approaches also share similar use-cases where they are deemed inadequate. Large urban scenarios are challenging for odometry techniques as they consider their environment as an infinite corridor due to lack of loop closure. On the other hand, the high number of features in the urban environments poses challenges, making the SLAM problem complex to solve. Although few works have managed to address these issues, the challenges faced in odometry and SLAM popularized map-based localization techniques for intelligent vehicles. In this category, the earliest methods are map-matching based GNSS localization, where GNSS position fixes are matched to the road network data available from the map. But for complex scenarios, GNSS alone is not sufficient for reliable map matching. Hence, other perception information such as lane markings, road structure, surrounding infrastructure etc. are used in map-matching based localization techniques. The development of HD maps and their increased availability further enhanced the possibility of using less-common perception information (road signs, curbs, etc.) in map-matching based localization. In the final section of this chapter, a survey on different types of models used to represent data from the sources is presented. In effect, this chapter gives us a perspective on why and where the data integrity concepts learned in Chapter 1 should be applied to the current multi-modal localization problems. Based on the insights obtained from these two chapters, the next chapter proposes an integrity monitoring framework for multimodal data sources used in localization problems.

# Chapter 3

## The Proposed Solution

### Contents

---

<b>3.1</b>	<b>Introduction . . . . .</b>	<b>52</b>
<b>3.2</b>	<b>Why Integrity of Sources? . . . . .</b>	<b>53</b>
<b>3.3</b>	<b>Modeling of Data Sources . . . . .</b>	<b>55</b>
<b>3.4</b>	<b>Integrity Assessment Technique . . . . .</b>	<b>57</b>
<b>3.5</b>	<b>Integrity Markers . . . . .</b>	<b>59</b>
<b>3.6</b>	<b>Conclusion . . . . .</b>	<b>60</b>

---

## 3.1 Introduction

The second half of the last decade has seen a significant emergence of commercially available vehicles with autonomous driving capabilities. We can confidently say that the status of autonomy in vehicles is well into the realm of SAE level 2 (International 2018). While the researchers and industries are rapidly moving towards SAE level 3 systems that can dramatically improve driving safety and efficiency, monitoring the integrity of sources and process used in such systems can often pose challenges (Velaga et al. 2012).

As discussed in Chapter 2, the solutions for the localization problem appear in a multitude of flavors. Evidently, as the technology progresses and manufacturing costs reduce, commercially available vehicles will be equipped with a large collection of sensors and ample computation power on-board. Since every sensor comes with its own advantages and disadvantages, it is safe to assume that combining data from multiple sensors, i.e., sensor fusion, will be a fundamental process in all navigation related tasks of intelligent vehicles. At the same time, as the number of elements in a system increase, the probability of system failure also increases. Hence the resultant reliability of the system depends on the reliability of the individual elements and their number and mutual arrangement (Mencik 2016). From Chapter 1, we know that advancements have been made in monitoring the faults in some of these elements (eg.: GNSS and digital maps) so that the total system reliability (and integrity) for a particular application is under the required limits. But the open question is: **how can we address the integrity of all components used in any multimodal localization system in a generalized framework so that any unexpected behaviors of that system is well explained and if possible, supervised?**

A simplified block diagram of the typical localization module in intelligent vehicles which fuses the information from multiple sensors and data sources to provide a localization estimate is shown in Fig. 3.1. In order to provide an accurate integrity indicator associated with the estimated localization output, integrity of sources and processes should be combined. Our work addresses the integrity monitoring of data sources and in this chapter, we begin by presenting the specific challenges identified in regards to the data integrity monitoring in multimodal localization problems. Based on the insights gained, we present the hypothesis of a generalized and scalable framework to address

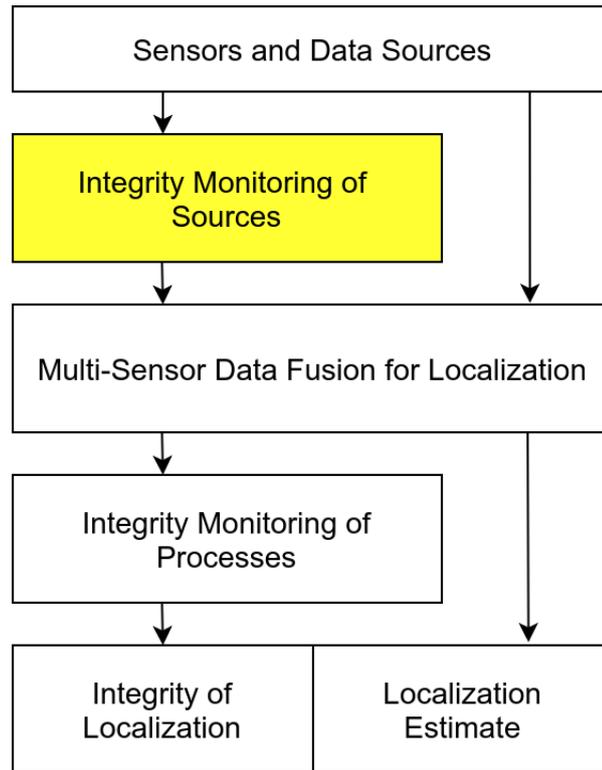


Figure 3.1: Localization module in Intelligent Vehicles

this problem. We also provide evidence-based support and argument for the design and choices made for the proposed solution. We define the indicators - qualitative and quantitative - used in this work to represent the integrity evaluation results generated by the framework. Finally, we outline the comparison schema used for the validation of results and the performance of the proposed framework.

## 3.2 Why Integrity of Sources?

Localization estimates are susceptible to uncertainty due to the inherent errors present in positioning information. In order to detect, isolate and remove any faults or anomalies in localization estimates, it is important to measure the confidence in positioning information. Introduction of redundancy and integration of positioning information with a different modality are the most common ways to achieve this task. The European Geostationary Navigation Overlay Service (EGNOS) and the Wide Area Augmentation System (WAAS), are developed to form a redundant source of information for the GNSS in order to perform integrity monitoring by providing correction information. However,

recent requirements of highly reliable localization systems in urban environments have inspired the inclusion of data with other modalities to monitor integrity of positioning information. While Velaga et al. 2012 propose a map-aided integrity monitoring of GNSS positioning system, Toledo-Moreo, Betaille, and Peyret (2010) combine map and dead reckoning to realize the same. Likewise, Shytermeja, Garcia-Pena, and Julien (2014) proposed an architecture for integrity monitoring of GNSS positioning system with the help of a Fisheye camera in urban environments. Integrity monitoring of GNSS localization using vehicle trajectories from repetitive journeys (Zinoune, Bonnifait, and Ibañez-Guzmán 2016) or using another source (often high-quality digital maps) as reliable ground truth (Worner et al. 2016; Le Marchand et al. 2009) can also be found in the literature. While adding data redundancy (often different GPS receiver for map-matching and sensor fusion Zinoune, Bonnifait, and Ibañez-Guzmán 2014) can monitor the integrity of processes, the integrity of these additional data sources has to be largely assumed. Only a small number of works like Li, Quddus, and Zhao (2013) and Zinoune, Bonnifait, and Ibañez-Guzmán (2016) consider digital maps as a source with probabilities of error. However, an important observation can be made here: the integrity of these complimentary data sources is overlooked while using them for the integrity monitoring of positioning information.

A multimodal localization system that uses lane markings detections from camera (Lu et al. (2014)) needs certain level of confidence associated to the available lane marking detections, in order to fuse this information with GPS. In works such as Liang, Zhang, and Wang (2017), inaccurate curb detection from LiDAR point cloud can greatly impact the localization in complex urban environments. In the same way, other LO and RO techniques as well as SLAM approaches described in Chapter 2 are based on the assumptions that the outlier removal methods (specifically RANSAC, model estimation) employed in those works can effectively eliminate any error in the input data. But both of these can often fail in certain complex cases, where ratio between inliers and outliers increases. On the other hand, lack of information is also a major concern to such systems. In a system which applies sensor fusion on 3 data sources (such as Kang et al. (2020) and Lu et al. (2014)), consistent absence of information from one source can reduce the positive impact that sensor brings to the fusion and ultimately causes the system to perform poorly. In the last decade, most of detection and tracking algorithms used in

Source	Type	Modality	Frame of reference	Frequency
GNSS	EC	ground-based beacons - 3D points	fixed, external reference frame (ECEF)	1-5 Hz
DR	PC	velocity, acceleration orientation - angles, values	ego frame of each sensors	10-500 Hz
Vision	EC	Visible light imaging - RGB, grayscale images	Camera frame	10-60 Hz
LiDAR	EC	Active ranging - cloud of 3D point	LiDAR frame	5-20 Hz
Radar	EC	Active Ranging - location of obstacles	Radar frame	10-20 Hz
Maps	-	Semantic data - points, lines, shapes	fixed, external reference frame (ECEF)	-

Table 3.1: Modality and properties of data sources

intelligent vehicles rely on machine learning (ML) techniques. But as shown by Szegedy et al. (2013), ML-based object detection from vision can have adverse effects even by slight shift in smoothness or contrast of an image. Explicit attacks on such data sources are also possible, where attackers can make slight physical modifications on the scene that can have huge impact on the accuracy of traffic sign and object detections using vision (Papernot et al. 2017) or employ a LiDAR spoofer to introduce 'fake' obstacles on to the scene to fail the navigation system (Cao et al. 2019). Though no works have directly examined the effects of such spoofing, jamming or disturbances of data sources on vehicle localizations, it is evident that the reliability of data sources needs considerations in the future of autonomous navigation.

### 3.3 Modeling of Data Sources

The main data sources used for localization are GNSS, maps, DR, vision, LiDAR and Radar. Table 3.1 illustrates the differences in modality, representation and rate of data provided by each of them. There are several possible classifications for these data sources based on the modality of the output data. As we know, sensors can be proprioceptive (PC) and exteroceptive (EC). But digital maps cannot be considered as sensors, since it is not measuring anything from the environment in real-time. Due to the focus on the data

coming from different sensors rather than the sensors itself, this work treats sensors and maps as data sources and defines a classification for data sources. *Spatial data sources (SDS)* are sources which provide positional data in 3D space (GNSS, vision, LiDAR, radar, maps) and *Motion data sources (MDS)* are sources that provide measurements of movement of the sensor (IMU, encoders). Exteroceptive sensors can be disturbed by external conditions, jammed or spoofed by external agencies while proprioceptive sensors do not suffer from any of these issues. Since maps can also contain intentional or unintentional errors and lack of information, the class of spatial data sources requires performance measures and quality indicators. Hence, spatial data sources are considered for the integrity monitoring framework developed in this work.

One of the main advantage of modeling of SDS is that strictly geometric modeling is possible without considering temporal attributes and physical parameters of the data measurements. For example, to model the 3D point cloud from a LiDAR sensor, the physics of laser light reflection or the time of measurement are not requirements. At the same time, since the geometries of data from each SDS and their respective frame of references are specific to the sources, their modeling often become source-specific as well. This disadvantage makes the simultaneous analysis of multiple SDS complex. Hence the first proposition of this work is to develop geometric models that can represent multiple SDS in a common frame with a common geometric model. Following the discussion presented in 2.6, we propose two models for data sources that can be used for integrity monitoring of several data sources.

In Chapter 4, we use a 2nd degree polynomial model (quadratic polynomial model - QPM) to address the integrity monitoring of three sources, GPS, vision and map in highway scenarios. Although Jeon et al. (2015) propose 3rd degree polynomial for GPS trajectories, in strictly-highway scenarios and with short GPS history, 2nd degree polynomial is a better choice. This is due to the fact that, 3rd degree polynomials cause over-fitting even at low levels of noises in GPS. As supported by Lu et al. (2014), lane markings in highways can also be modeled using QPM. Since nodes available from map data can be considered as a discretized version of road structures (which follow polynomial characteristics in highway scenarios), we propose QPM for modeling of map nodes as well. In the interest of including varied and complex urban scenarios to this framework, we propose a grid-based model in Chapter 5, which is inspired from the concept of

occupancy grids proposed by Yu, Cherfaoui, and Bonnifait (2015) and Moras, Cherfaoui, and Bonnifait (2011). We focus on urban and semi-urban environments using this model, where LiDAR, vision and map are used in state-of-the-art localization methods to detect various features from the environment. Instead of the free-space representation used in conventional occupancy grids, we propose Feature Grid (FG), where cells are populated with the labels corresponding to the most probable feature they contain. In both of these approaches, we have made the choice of ego-frame of the vehicle as the common reference frame for these models. Except GPS and maps, transformation between other SDS frames and ego-frame of vehicle are known or calculated using extrinsic calibration. Transformation of GPS and maps to ego-frame of the vehicle can be observed from map-matching or DR sensors. The methods used for this transformation are explained in the respective chapters.

### 3.4 Integrity Assessment Technique

From the research presented in Chapter 1, we concluded that accuracy and validity are the two core attributes of integrity that can be applied to the navigation of intelligent vehicles. Out of the listed enablers of information integrity shown in Fig. 1.1, we propose consistency and availability as the main enablers of data source integrity in navigation. Since accuracy is the quantification of how well the information represents the real scenario, consistency between the same information from multiple sources is enabler for accuracy. Verifiability can also be treated as the enabler for accuracy, but in real-time operations like navigation of a single intelligent vehicle, it is not possible to exactly replicate the data from a source under same conditions to verify its accuracy. However, this is possible to an extent with redundancy (RAIM algorithms) or history of multiple journeys through the same environment (Zinoune, Bonnifait, and Ibañez-Guzmán 2014). Since in this work we do not require redundancy of sensors and adhere to real-time considerations, we regard consistency as the enabler for accuracy. As far as the attribute of validity is concerned, we propose availability, understandability and dependability of information as its enablers. Information that is not available for a process is definitely invalid. If a road is missing in an outdated map but exists in the real world, map is invalid for a vehicle which travels through that road. On the other hand, understandability and

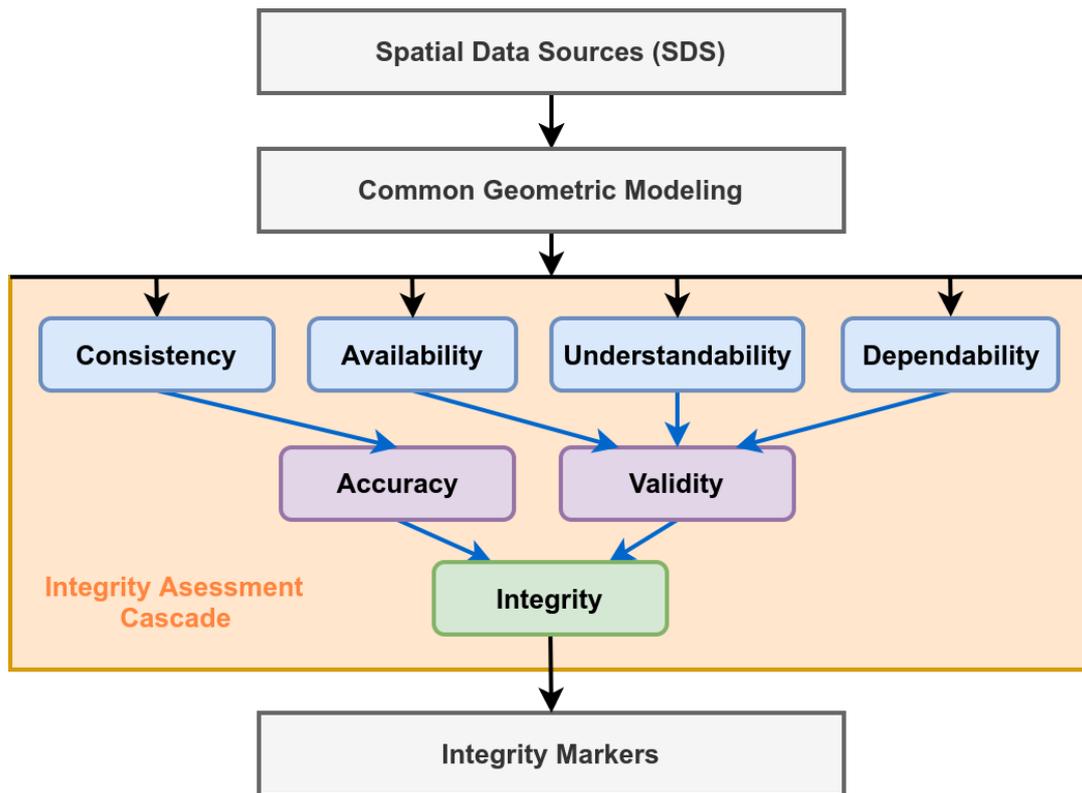


Figure 3.2: Framework for integrity monitoring of sources

dependability are application-specific enablers. If the information from a data source does not meet the requirement of the process, understandability of the data source is less. For example, in a system where history of GPS fixes are used for map-matching (Sharath, Velaga, and Quddus 2019), lack of required length of this GPS trail can reduce the understandability of the data. Likewise, if we have any prior information that a particular abnormality from a data source can cause a specific failure or degradation of performance in a system, dependability of that data source for that process can be monitored. In case of lower understandability and reduced dependability (beyond a threshold), we can treat information from the data source as invalid.

Developing on this premise, we propose an integrity assessment technique that estimates the above mentioned enablers and addresses accuracy and validity of information from data sources as shown in Fig. 3.2. Data from different SDS are processed and the information collected are modeled using a common geometrical model. Models from each source are compared with other to estimate the consistency between them. These consistency estimates are combined to evaluate the accuracy of each source. Availability of required data for the modeling, understandability of the available data and prior

knowledge based dependability of data are also estimated at this stage. A decision making step generates validity indicators according to the applications and sensor modalities available. The accuracy estimate and validity indicators are combined to provide integrity attributes associated with each source.

## 3.5 Integrity Markers

In Sec. 1.3, we have seen the conventional integrity markers used in intelligent vehicles such as PL, AL, TTA and IR. The integrity events (SA, SU, MI and HMI) are defined based on these markers and they are used in FD, FDI, FDE or FDIA approaches to monitor integrity of localization estimation. However, the inclusion of multiple sources and the shift in focus to the integrity of data sources rather than the localization process makes conventional integrity markers quite limited or not applicable in our work. Hence, a new set of integrity markers has to be defined with respect to the integrity concepts adopted from data sciences to formulate our integrity assessment techniques. When it comes to the data from the sources, it is more useful to answer the question '*what is the extent of trust that can be placed on the data?*' rather than '*can the data be trusted?*' because the system that follows can have multiple modes of operation. Consider this scenario: a system uses road structure from the map and lane marking detection from vision to correct the localization provided by GPS. At an instant, the vision system mis-identified a white patch on the pavement as the left lane marking instead of the correct one, while right lane marking detection and road structure obtained from map are accurate. In this case, localization system does not need to discard the entire visual data, but reduce the importance of visual data in the localization process.

On the other hand, even if the system is not designed to perform such an adaptive integration of data, it can still provide the localization output without data integrity consideration, but with the reasoning of which source or data combination may have caused the degradation in localization. The former approach will be analogous to FDIA and the latter to FDI. Hence, our integrity markers are a combination of two concepts - a variable value called Integrity Weight (IW) assigned to each source based on accuracy of data and a logical value called Fault or Feasibility Predictor (FP) for sources or situations based on the validity of data. In Chapter 5, since GNSS position fixes are not directly

observable using Feature Grid model, we propose a classical protection level computation technique for GNSS fixes using the integrity assessment framework proposed in this work. This is compared to the existing protection level definitions and calculations for vehicle localization in the literature.

## 3.6 Conclusion

In this chapter, the proposed solution to the integrity monitoring of data sources used in the multimodal localization problem is presented. The reasoning behind the necessity of integrity analysis of data sources is discussed. A new classification of data sources called spatial data sources (SDS) is introduced and their modelling schemes are outlined. For highway scenarios, a quadratic polynomial model (QPM) is identified as a suitable model for representing GPS, vision and map data. A new feature grid (FG) model inspired from the occupancy grid concept is proposed to represent LiDAR, vision and map data used in urban scenarios. Data integrity concept presented Chapter 1 have been transposed and the attributes and enablers of data integrity that are persistent to data sources used in this work are identified. The framework of data integrity monitoring is developed based on these choices. Finally, two markers - Integrity Weight (IW) and Fault and Feasibility Predictor (FP) - are formalized to quantify the estimated integrity. Hence this chapter serves as a precursor by providing a common framework for the methodologies presented in the following chapters.

# Chapter 4

## The Proof Of Concept

### Contents

---

<b>4.1</b>	<b>Introduction</b>	<b>62</b>
<b>4.2</b>	<b>Data Handling</b>	<b>62</b>
4.2.1	Detection and Outlier Removal	63
4.2.2	Transformation Optimization	69
<b>4.3</b>	<b>Integrity Assessment</b>	<b>71</b>
4.3.1	Cross-Consistency estimation	71
4.3.2	Integrity Weights	73
4.3.3	Fault and Feasibility Predictors	73
<b>4.4</b>	<b>Experiments and Discussion</b>	<b>74</b>
4.4.1	Dataset 1 - 2011_09_26_drive_0028	75
4.4.2	Dataset 2 - 2011_09_26_drive_0029	76
4.4.3	Dataset 3 - 2011_09_26_drive_0042	77
4.4.4	Dataset 4 - 2011_09_26_drive_0070	79
4.4.5	Performance Evaluation	80
<b>4.5</b>	<b>Conclusions</b>	<b>83</b>

---

## 4.1 Introduction

In this chapter, we present the framework for integrity monitoring of data sources intended for highway and semi-urban scenarios. The generalized functional block diagram of this approach shown in Fig. 4.1, is inspired from Peynot (2012). Consider  $N$  number of data sources which can continuously provide spatial information relevant to the localization of the vehicle in the driving scenario. As a preprocessing step, outliers are removed from each source according to the prior knowledge about their models and data acquisition technique. Then all data are transformed to a chosen reference frame using different relative geometric transformations, scale and units. A common model is applied to each of the data sources represented in a pre-defined reference frame and these models are compared using a cross-consistency analysis routine which in turn provides integrity markers for each source.

For the development and analysis of the proposed concept, a multimodal perception system with three different data sources is used in this work. This section describes the approach to sensor integrity assessment in which digital maps are considered as a data source along with GPS receiver and cameras. GPS coordinates, detected lane markings from vision and road structure from digital maps are considered in a common framework and their models are extracted. Quadratic polynomials are used as models, assuming that road structures and driving tracks can be approximated locally using single convex functions (Toledo-Moreo et al. 2007). In order to robustly estimate the models, M-estimators are used and outliers are removed from each source. Obtained models from each data source are then compared by estimating their fit across other data sources. This enables a cross-consistency analysis of each data sources, thus providing an effective and quantifiable marker to represent their reliability.

## 4.2 Data Handling

Due to the importance of accurate data representation in the proposed method, this section is devoted to describe the evolution of data from each source considered. Choice of a common reference frame is the first step as it defines the source-specific transformations. Ego-frame of the vehicle is chosen as the common reference frame where X axis is along

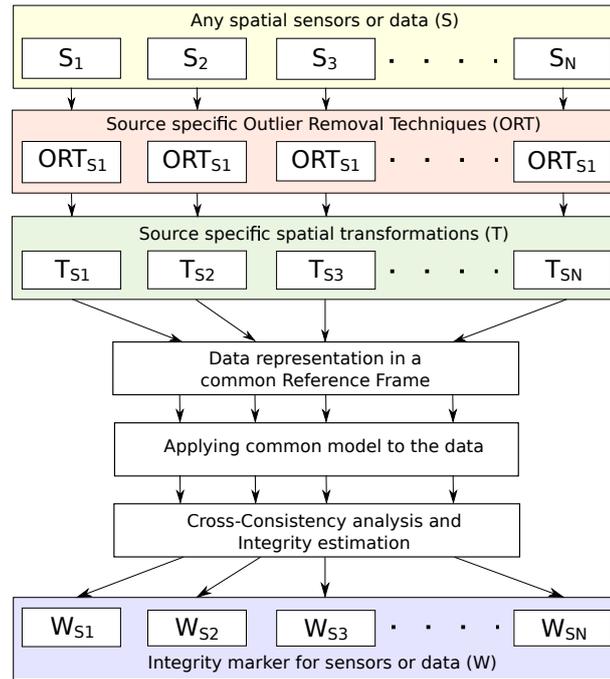


Figure 4.1: Functional block diagram of the methodology proposed

the lateral direction and the Y axis is along the longitudinal direction of the road as shown in Fig. 4.4.

### 4.2.1 Detection and Outlier Removal

The visual data used in our framework is coming from a front-facing camera on-board the vehicle. We assume the extrinsic and intrinsic calibration parameters of the camera are known and their uncertainties are known. The first step is to transform the image formed on camera plane into Bird-Eye-View (BEV), which is a virtual top view where image plane is parallel to ground plane. This transformation is called Inverse Perspective Transform (IPM), which applies a homography based on the projection matrix of the calibrated camera. IPM transformations have three fundamental assumptions: 1) the position of camera is fixed respective to the road, 2) road is planar and 3) the scene only contains road (only the ground plane). These three assumptions may not be true all the time in real world situations. This is highlighted in Fig. 4.2b, the BEV of railings on the sides of the road does not represent the real dimensions and the parallel nature of the lane markings are gradually lost as the distance increases from camera due to the slight errors in planarity of the road. Also as shown in Fig. 4.2b, accuracy of IPM decreases



(a) The original image from camera



(b) Image transformed to BEV space using IPM

(c) Cropped image in BEV space

(d) Segmentation of lane markings from image in BEV space

Figure 4.2: Vision data processing pipeline

as distance from the camera increases and pixels are largely interpolated at distances. Hence, a valid IPM transform is only applicable at reasonably short distances ( $< 25$  m) and only information available on the road surface within this range are reliable to use. This leads to cropping the BEV view to a region of interest (ROI) as show in Fig. 4.2c. Since we focus on highway scenarios in this part of the work, all available lane markings from the road. A typical 3-channel color-based thresholding function generates possible lane marking regions from the image as shown in Fig. 4.2d. Clearly, such a segmentation does not guarantee accurate lane markings. However, the output represents the relevant parts of data available from the source (vision data in camera-frame,  $V_{cf}$ ), from which lane markings are obtained by applying models and removing outliers (Rabe et al. 2017; Kang et al. 2020; Mammeri, Boukerche, and Tang 2016). Parameters such as centroid, major-axis, minor-axis, orientation, area etc. are calculated using blob analysis method for every candidate region. Less probable candidate patches are removed using filtering criteria based on area, orientation and minor-axis length. Valid candidate patches ( $V'_{cf}$ ) obtained for the situation in Fig. 4.2a are shown in Fig. 4.3a with their respective centroids and bounding boxes. Pixels included in these valid patches are reprojected to

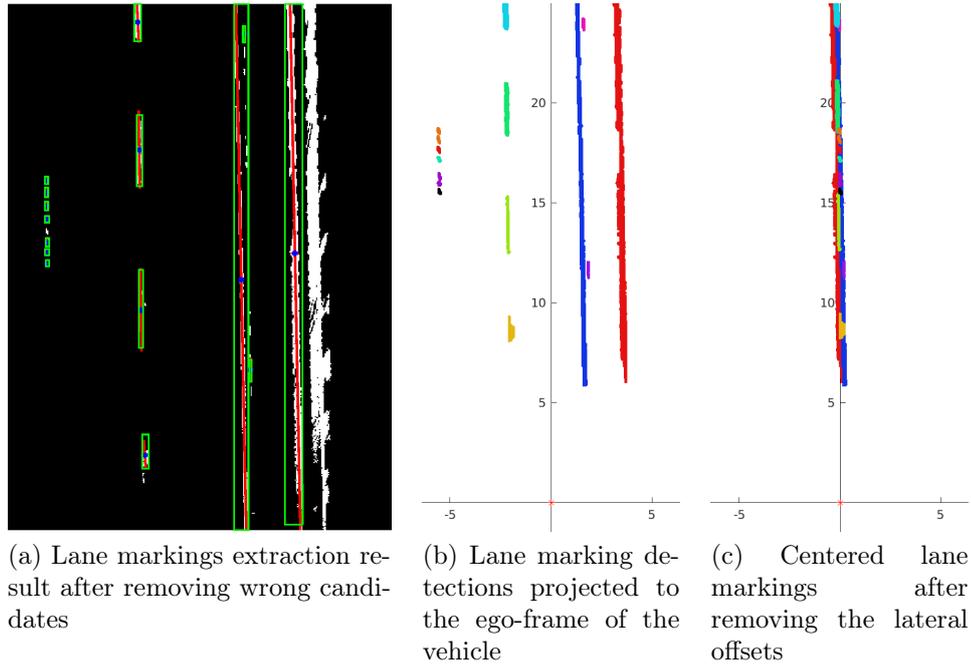


Figure 4.3: Lane markings detection and representation

ego-frame of the vehicle as vision data points in real-world distance units, using camera calibration and BEV parameters as shown in Fig. 4.3b. Finally, the lateral offset between the Y-axis and the centroid of each patch is calculated and the offset is removed to align the vision points to the Y-axis as in Fig. 4.3c, thus providing outlier removed vision data,  $V'_{ego}$ .

GPS receivers provide data in the form of (*latitude, longitude*) which are in degrees of angles in World Geodetic System (WGS) coordinates, specifically its latest version, WGS84. Since WGS84 is a spherical coordinate system, projection into a Cartesian coordinate system is required to directly apply analytic geometric functions and operations. At every time step, current GPS fix is used as the reference and the GPS trajectory ( $G_{wgs}$ ) consisting of  $n$  previous fixes are transformed into East-North-UP (ENU) frame using Mercator projections (GPS data,  $G_{enu}$ ). The different frames involved in this are shown in Fig. 4.4. Here we assume that the origin of ego-frame and ENU frame created based on current GPS fix are the same. In other words, we neglect the extrinsic transformation information of the position of GPS receiver with respect to the vehicle.

Geographic Information System (GIS) is a framework for capturing, analyzing and visualizing geographic data. GISs are the core of every modern location-enabled service, that depends on spatial and semantic information from environments. Although GIS

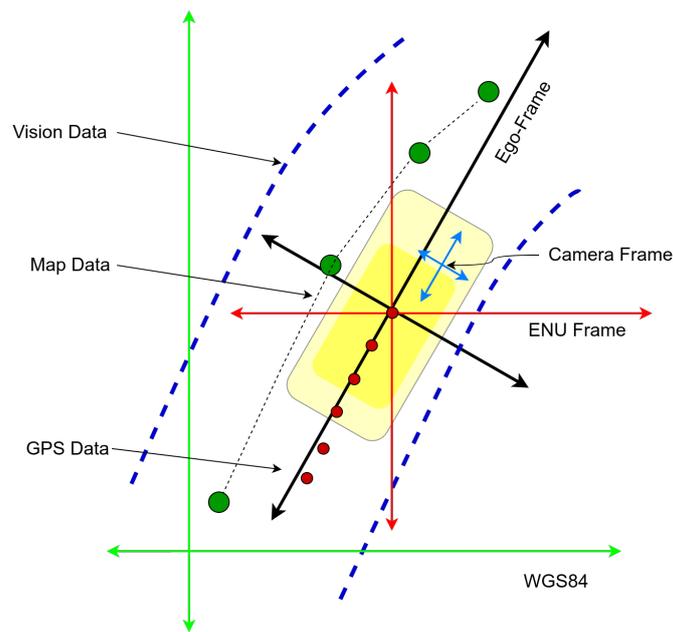


Figure 4.4: Different frames associated with the analysis

based data services can be categorized based on several parameters, there are two main categories based on the access to data. Proprietary GISs own the data by copyrights and allow restricted access to use the data based on subscriptions or purchases (eg. Google Maps, Apple Maps, Bing Maps, Here Maps etc.). On the other hand, open-source GISs allow users unrestricted copyright-free access to the data (OpenStreetMap, MapBox). While both of these can have free web-based visualization tools and possibility of user feedbacks and modifications, only open-source GISs allow the access to the core data structures to be downloaded offline or modified by users. While several GISs have been used in the literature in the context of vehicle localization, OpenStreetMap (OSM) has gained huge popularity due to its data accessibility and community support. OSM is a leading example of Volunteered Geographic Information (VGI), where the data is crowd-sourced, user-generated and community-verified. Unlike other VGI projects, OSM revolves around the construction of a vector data set representing the entire planet, not just annotations on an existing map, and emphasizes the openness of its datasets Ballatore, Bertolotto, and Wilson 2013. One of the reasons for the adoption of OSM in intelligent vehicle navigation is the simplicity of the structure data used for map data storage. This data can be downloaded in XML format from OSM website, which gives an offline copy of all information in an area of interest specified by a bounding box. Road

```

<node id="2480372183" visible="true" version="1" changeset="18192"
  lat="41.3888277" lon="2.1628747"/>
  ...
  ...
<way id="37143032" visible="true" version="5" changeset="25493" >
  <nd ref="2480372183"/>
  <nd ref="432594825"/>
  <nd ref="432594826"/>
  <nd ref="432594831"/>
  <nd ref="2480372183"/>
  <tag k="barrier" v="wall"/>
</way>
  ...
  ...
<relation id="165992" visible="true" version="3" changeset="20363" >
  <member type="way" ref="37143032" role="outer"/>
  <member type="way" ref="37143031" role="inner"/>
  <tag k="leisure" v="garden"/>
  <tag k="type" v="multipolygon"/>
</relation>

```

Listing 4.1: An example code snippet from offline OpenStreetMap data file

structures, buildings and other spatial features are represented using an entity-tag data model. There are three basic entities in OSM:

- Node - a single point in space which is defined by latitude, longitude (with optional altitude information), referenced with an id.
- Way - an ordered list of nodes that constitutes a unique spatial feature, referenced with an id.
- Relation - an ordered list of nodes, ways and/or relations which collectively define a geographic structure, referenced with an id.

Tags are used to denote the specific features of any of these map elements. An example of the data from OSM snippet is given in code listing 4.1.

In code listing 4.1, a relation (id:165992) consisting of two *ways* that defines a multi-polygon (tag:type=multipolygon) feature corresponding to a garden (tag:leisure=garden). One of these *ways* (id:37143032) represents the wall tag (tag:barrier=wall) of the garden and has five nodes. Since the first and last node of this way are the same, it can be understood as a closed wall. Finally, the nodes (eg.- id:2480372183) give the geographic position of the wall with their latitude and longitude. Since in this chapter we focus on highways and road structures, we extract a subset map from the offline OSM map

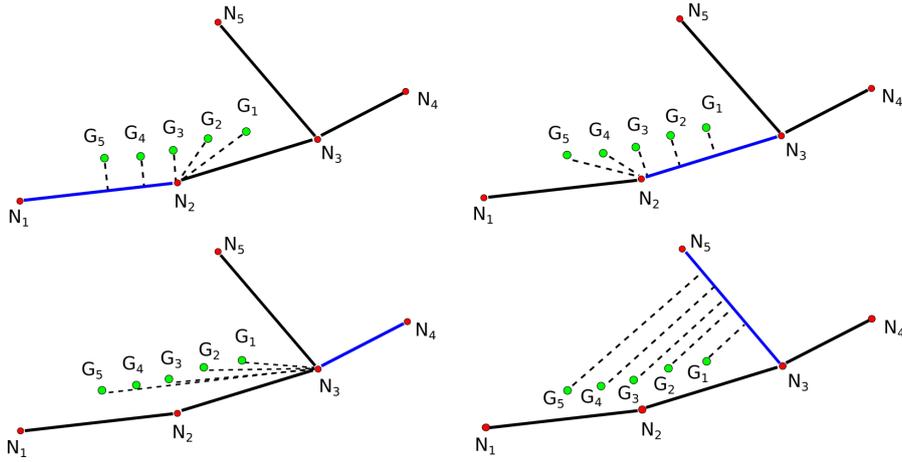


Figure 4.5: Point-to-Curve map matching algorithm

in the area of interest using 'highway' tag. All classified roads (except residential roads) and link roads are selected using the values available for 'highway' tag. Cross-referenced look-up-tables (LUT) are created for nodes and ways that represent the road structure in the area.

Since OSM employs the WGS84 standard as the coordinate system,  $G_{wgs}$  can be used to extract the relevant information from map using basic map-matching techniques. We use a basic point-to-curve matching scheme to identify the specific road on which vehicle is traveling. Based on the current GPS fix, a bounding box is created with pre-defined size and all map nodes inside ( $M_{wgs}$ ) are extracted. The distance between every GPS fix in  $G_{wgs}$  and all the possible line segments (according to *way* information) created by the nodes in  $M_{wgs}$ , is calculated as shown in Fig. 4.5. If a GPS fix has perpendicular projection on the line segment, the perpendicular distance is calculated. If it is not the case, the distance to the nearest node on the line segment is calculated as distance. The segment that generates the least sum of distances is chosen and its corresponding *way* is selected as the road on which the vehicle is currently present. All the member nodes of the estimated way which are inside the bounding-box are selected and transformed to ENU frame using current GPS fix as the reference, resulting in map data ( $M_{enu}$ ).

$G_{enu}$  and  $M_{enu}$  have similar outlier characteristics. Outliers in these sources are the coordinates which do not fit the characteristic model of them. Least squares method is used to fit quadratic polynomials of the form

$$y = p(x) = ax^2 + bx + c \quad (4.1)$$

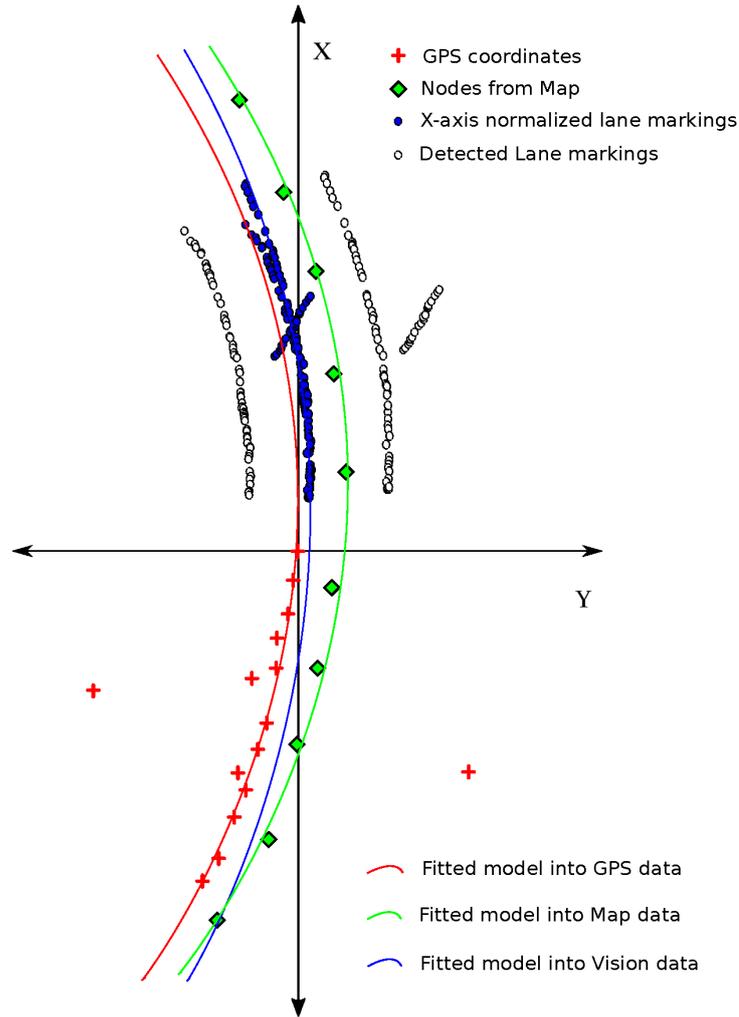


Figure 4.6: Representation of data from different sources in ego-frame during typical scenario

to  $G_{enu}$  and  $M_{enu}$ . Once the models are applied ( $p_g$  and  $p_m$  for  $G_{enu}$  and  $M_{enu}$  respectively), the residual distributions of each data source are computed ( $r_g, r_m$ ). To model the distribution of residuals, mean and standard deviation are computed. Using these as initial estimates, Huber's robust parameter estimation method is employed to get the final model of the distribution. Outliers are detected and removed using  $2\sigma$  threshold on  $r_g$  and  $r_m$ . The resulting data points provide valid GPS data  $G'_{enu}$ , and map data  $M'_{enu}$ .

### 4.2.2 Transformation Optimization

To proceed further,  $G'_{enu}$  and  $M'_{enu}$  need to be transformed to ego-frame. Since current GPS fix is the common origin of ego-frame and ENU frame, translation component of this transformation is already addressed. Thus, this transformation is reduced to rotation

using the estimated orientation of the vehicle. A straightforward option to estimate this information, is to use integrated IMU or DR sensor data. However, drift observed in such systems should be considered in this case. Another option, at least in theory, is to estimate the orientation by calculating the angle between the current and previous GPS points. But noisy GPS data does not allow this approach. Although mean of the angles between current and few immediate previous points could improve the reliability of GPS-based estimation of orientation, roads with large curvature could pose a problem to this approach. While leaving the choice of orientation estimation process open-ended, in our work, we propose a method that is robust and less-reliant on the quality of GPS data.

Let  $G'_{enu}$ ,  $M'_{enu}$ ,  $V'_{ego}$  be the outlier removed dataset. To estimate the orientation,  $G'_{enu}$  and  $V'_{ego}$  are used in our work due to two reasons. Firstly,  $G'_{enu}$  and  $V'_{ego}$  are both directly related to the vehicle because of their respectiveness sensors are on the vehicle while  $M'_{enu}$  is an external data source. Secondly, in a chronological point of view, a set of previous GPS coordinates represents the history of the vehicle's trail and detected lane markings represents the prediction of the trail. Intuitively, these two demand consistency under all normal driving scenarios.

At this point, we introduce the operation *RFS* (Rotate-Fit-Sample), which is defined as

$$P_{D\theta} := \{(x_1, y_{k1}), (x_2, y_{k2}), \dots, (x_N, y_{kN})\} = RFS(D, \theta, N) \quad (4.2)$$

where  $D$  is the set of data points from a source represented in ego-frame. *RFS* rotates  $D$  about the origin by angle  $\theta$ , then robustly fits second degree polynomial to the result and sample the polynomial into  $N$  2D points equidistantly along X-axis.

The initial estimate of orientation of the vehicle ( $\theta'$ ), is the mean of the angles between current and 5 previous GPS positions from  $G'_{enu}$  (outlier removed GPS data).  $RFS(G'_{enu}, \theta', 20)$  and  $RFS(V'_{ego}, 0, 20)$  gives the initial framework for angle estimation method. The angle is found out by iterative minimization of the Euclidean distance between these two set of points.

$$\hat{\theta} = \arg \min_{\theta} (\|P_{G\theta} - P_{V0}\|_2) \quad (4.3)$$

$G'_{enu}$  and  $M'_{enu}$  are rotated using  $\hat{\theta}$  to get  $G'_{ego}$  and  $M'_{ego}$ . The final representation of all the data from different sources as illustrated in Fig. 4.6.

### 4.3 Integrity Assessment

In this section, the integrity assessment of the three data sources using the extracted polynomial models is presented. The technique used to check the consistency between these models is named as cross-consistency estimation.

#### 4.3.1 Cross-Consistency estimation

In this section, we propose a method to analyze the consistency between different data sources. If the data from each source is accurate, their models should exhibit similar properties such as similar curvature and tangent angles. After the representation of data in ego-frame, quadratic polynomial models  $p_{gg}$ ,  $p_{mm}$  and  $p_{vv}$  are estimated from  $G'_{ego}$ ,  $M'_{ego}$  and  $V'_{ego}$  respectively, as shown in Fig. 4.6. Next step is to assess the 'Goodness-of-Fit' (GoF) of each of these polynomial models across all possible data sources. The characteristics of a quadratic polynomial are defined by its second and first order parameters and the constant parameter represents the Y-axis offset of the polynomial at the origin. Hence, to check the GoF of a model to a different set of data points, it can be shifted to the location of those data points. This is achieved by changing the constant parameter of the model under consideration, by the same of the model of the data source to be considered. Hence three data sources considered in our proposed method will provide 9 combinations as follows.

$$\begin{array}{lll}
 p_{gg} : (a_g, b_g, c_g) & p_{mg} : (a_m, b_m, c_g) & p_{vg} : (a_v, b_v, c_g) \\
 p_{gm} : (a_g, b_g, c_m) & p_{mm} : (a_m, b_m, c_m) & p_{vm} : (a_v, b_v, c_m) \\
 p_{gv} : (a_g, b_g, c_v) & p_{mv} : (a_m, b_m, c_v) & p_{vv} : (a_v, b_v, c_v)
 \end{array}$$

Where  $P_{ij}$  represents the polynomial model used to analyze GoF of  $P_{ii}$  to  $j'_{ego}$  data source ( $i, j \in \{g, m, v\}$ ). GoF of each of these models are defined based on normalized sum of weighted residuals. In order to weight the residuals, a measure of curvature  $\omega_l$  is

computed for each point on the polynomial function using in Eq. 4.4.

$$\omega_l = \frac{2a}{\left(4a^2x_l^2 + 4abx_l + b^2 + 1\right)^{\frac{3}{2}}} \quad (4.4)$$

Where  $a$ ,  $b$  are polynomial model parameters and  $x_l$  is the x-axis value of the point at with curvature needed to be estimated.

Weighing residuals using measure of the curvature has particular significance. The curvature of the model is most significant and observable at smaller values along the x-axis. Likewise, while estimating a polynomial model from a set of data points, points that are near origin will determine the curvature of the model more than the farther ones. This property intrinsically incorporates more importance for the data points near the vehicle's current position, which is preferred for cross-consistency analysis. On the other hand, if the vehicle is on a straight road segment, none of the sources will provide a model with a considerable curvature (i.e, close to zero value). Hence incorporating the measure of the curvature into the GoF estimation will automatically compensate for the limitations of this approach. Finally, we define the marker for cross-consistency analysis as

$$e_{ij} = \frac{1}{N_j} \sum_{l=1}^{N_j} \left( \omega_l (y_l - p_{ij}(x_l)) \right) \quad \forall (x_l, y_l) \in j'_{ego} \quad (4.5)$$

where  $e_{ij}$  is marker associated with  $p_{ij}$  and  $N_j$  is the total number of data points in  $j'_{ego}$ . All possible  $e_{ij}$  values are arranged in an evolving square matrix  $E$  as given below.

$$E = \begin{bmatrix} e_{gg} & e_{mg} & e_{vg} \\ e_{gm} & e_{mm} & e_{vm} \\ e_{gv} & e_{mv} & e_{vv} \end{bmatrix} \quad (4.6)$$

In order to obtain the total cross-consistency marker associated with the source, we can also combine all its individual cross-consistency markers  $E_i$  as,

$$E_i = \sum_{\forall j} e_{ij} \quad (4.7)$$

### 4.3.2 Integrity Weights

In section 4.3.1, we formulated a quantifiable marker to express the consistency of each source with the other sources. At this point, if  $e_{ij}$  is a low value, it can be inferred that  $i'_{ego}$  is consistent with  $j'_{ego}$ . In the case of  $i = j$ , a low  $e_{ij}$  value means that model estimation is reliable. However, to estimate the integrity of a source, we have to combine these notions in a sensible way. The idea is to provide a weighing parameter for each source, which can be easily incorporated into sensor-fusion or localization improvement algorithms used in autonomous driving process. Combining the GoF of a polynomial model to its own source and the GoF to other sources will represent the consistency of that model, hence the integrity of the respective data source. Hence the integrity weights (IW) are defined as below.

$$IW_i = 1 - \frac{E_i}{\sum_{\forall i,j} e_{ij}} \quad (4.8)$$

These weights,  $IW_g$ ,  $IW_m$  and  $IW_v$ , represent the integrity of data obtained from GPS, map and vision respectively. Higher values of these weights correspond to higher integrity of their respective sources and vice versa. The formulation of this marker tells us how much of the total error observed in the consistency analysis is caused by each data source.

### 4.3.3 Fault and Feasibility Predictors

The cross-consistency estimation process assumes that the models for all data sources are always present and their consistency is an observable quality at all the instances. There can be several scenarios where these assumptions could not be held true. In this section, we try to address such scenarios with Fault and Feasibility Predictors ( $FP$ ) defined for each of them.

- $FP_g$ : Fault Predictor for GPS - If GPS fix is not available or the current GPS fix is an outlier at any given time. Under these cases, we will use the previous reliable GPS localization and proceed with the process, but to account the unreliability of the process  $FP_g$  is raised.
- $FP_m$ : Feasibility Predictor for Map - If there are two or fewer nodes in the map cache, a reliable fitting of the polynomial model is not possible, thus raising  $FP_m$ .

This happens when road sections do not have a considerable curvature because cartographers tend to use fewer nodes for constructing maps in such cases.

- $FP_v$ : Fault prediction for Vision - If there are no long and continuous lane markings detected. Under certain challenging situations, characteristic lane markings may not be present or may not be detected. The model fitting may work on detected noises in visual data, but its validity could be seriously affected.  $FP_v$  addresses these conditions.
- $FP_s$ : Feasibility predictor based on speed - If the vehicle comes to a stop or moving significantly slow, the history of GPS fixes used to extract the model from GPS source could not represent the situation. Hence if the GPS history does not span over a fixed threshold length,  $FP_s$  is raised.
- $FP_t$ : Feasibility predictor based on turning - If a vehicle is taking a hard turn from one road to another discontinuous road,  $FP_t$  is raised. In these scenarios, there will not be any consistency between what vehicle observes in front and the history of vehicle localization.  $FP_t$  will be kept raised until the history of vehicle localization completely belongs to the new road section.

When any of these  $FP$ s are raised, the output of the cross-consistency analysis should be taken with expected unreliability . Though  $FP$ s need some hard thresholds to be defined, they need not be highly empirical and can be easily derived using pragmatic understanding of the situations.

## 4.4 Experiments and Discussion

Experiments are conducted with datasets available in KITTI benchmark suit (Geiger et al. 2013) to establish proof-of-concept. RTK GPS fixes available in these datasets are modified with additive white noise to emulate poor localization fixes as well as random outliers provided by generic GPS sensors (Lu et al. 2014). This section considers the use of four scenarios to test proposed method. The result of integrity assessment is reported as three scores at each time-step (in this case frame number, since GPS and visual data are sampled synchronously) : total error observed ( $E_t$ ), ratio of errors due to each data

source in the total error observed ( $R = E_g:E_m:E_v$ , such that each are in normalized range (0,1)) and reliability of integrity assessment ( $FP$ ).  $FP$  encloses the set of all fault detection and feasibility criteria for integrity assessment discussed in section 4.3.3. The following discussion on results is structured in a specific format. Sections with high total error observed are identified from each dataset and a representative scenario from each section is examined to explain the cause and confirm the validity of the contribution by each data source to the total error.

#### 4.4.1 Dataset 1 - 2011\_09\_26\_drive\_0028

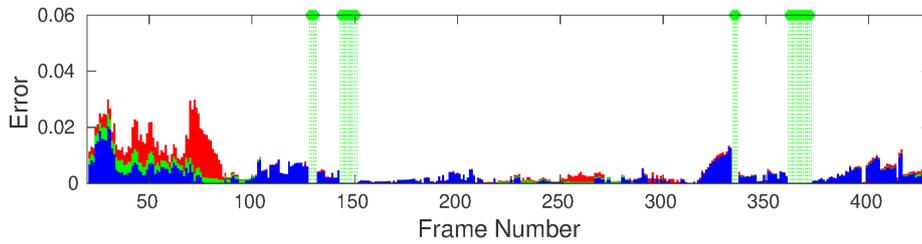
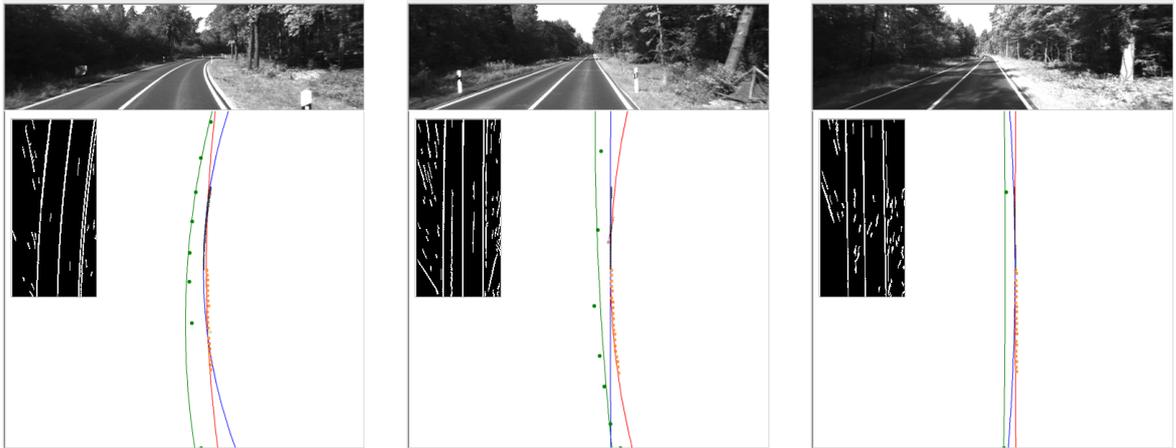


Figure 4.7: Error evolution for Dataset 1. Red: GPS error, Green: Map error, Blue: Vision error, Green dotted lines:  $FP_m$



(a) Scenario at frame number 28 (b) Scenario at frame number 77 (c) Scenario at frame number 365

Figure 4.8: Different scenarios from dataset 1. Top: Camera view of the scenario, Bottom: Data representation and model fitting in ego-frame, Inset: lane detection result in BEV. Red: GPS data and its model, Green: map data and its model, Blue: vision data and its model.

This dataset contains images that are polluted with multiples shadows from surroundings which reduces the performance of lane marking detection. Evolution of errors observed in

the integrity estimation process on this dataset is given in Fig. 4.7. Despite the low error observed throughout this dataset, it is interesting to notice the contribution of different sources in the total error in several conditions. In the scenario at frame number 28 (Fig. 4.8b), a larger curvature is observed in vision due to the entry of vehicle into a curve from a straight road segment, which results in the ratio of errors 0.191:0.183:**0.625**. At frame number 77 (Fig. 4.8b), the vehicle exits the same curve to a straight road segment, causing higher curvature in GPS which results in the ratio **0.883**:0.108:0.011. Two or less nodes are available from the maps at all the straight road scenarios marked with green dotted lines ( $FP_m$ ) in Fig. 4.7. An example of such a scenario is shown in Fig. 4.8c.

#### 4.4.2 Dataset 2 - 2011\_09\_26\_drive\_0029

In dataset 2, a compound and a reverse curve follows a left turn from a red light stop at the end of straight road segment. Four types of  $FP$  markers are observed in this process. In the straight line segment, few  $FP_v$  (frame number 89,91) and  $FP_m$  (frame number 104-111) are observed due to lack of information. As the vehicle switches the lanes to the left most lane for the turn, an increase in  $E_t$  is observed due to poor model fitting on GPS data until the  $FP_s$  is produced from frame number 186 to 265. In Fig. 4.10a, a situation during the stop is shown. Note that the model is ill-fitted on all the data sources. GPS points are concentrated around vehicle's stop location, causing unstable polynomial models. While vision observes the straight road ahead, moving vehicles present in front of the vehicle cause faulty lane marking detection, thus resulting in unstable polynomial models.  $FP_s$  encompasses all these unreliability in the integrity assessment process until the vehicle starts to move again from 265. But from 259 onward  $FP_t$  is produced indicating the hard turn that results in wrong lane detection and inconsistent GPS history. The turning section is then followed by vision data outages due to challenging light conditions and shadows present along the curves (similar to the situation shown in Fig. 4.10b and Fig. 4.10c). Integrity assessment works well on compound curve section due to highly stable curve models estimated. Even though  $E_t$  is high, combined with the  $R_e$ , integrity assessment gives behavior of each data source accurately. Fig. 4.10c shows the limitation of the proposed method due to the presence of a reverse curve. The curve vehicle had traveled and the curve vehicle observes ahead are in opposite direction in this situation

so a second degree polynomial fails to model the map data.

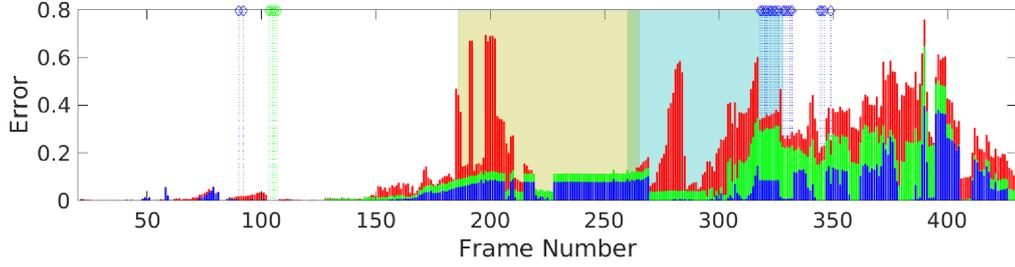
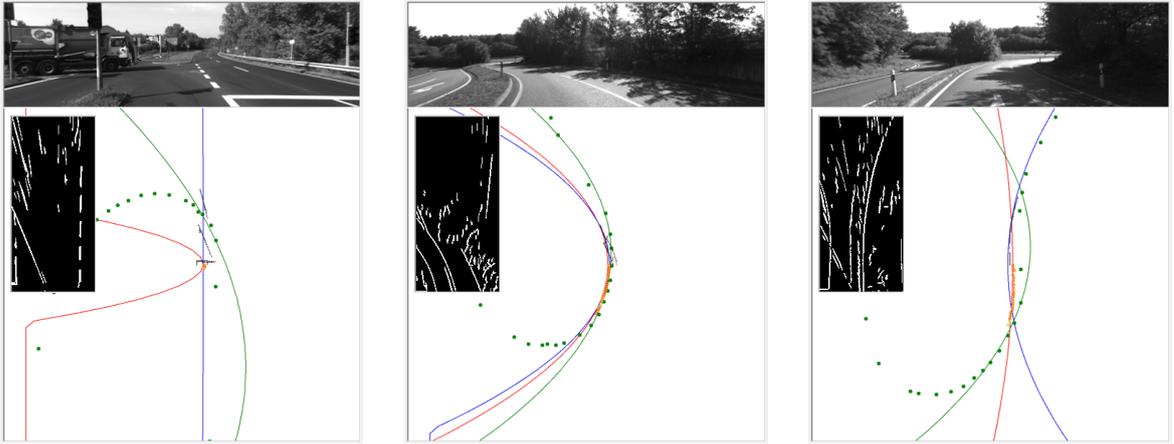


Figure 4.9: Error evolution for Dataset 2. Red: GPS error, Green: Map error, Blue: Vision error, Green dotted lines:  $FP_m$ , Blue dotted lines:  $FP_v$ , Light brown:  $FP_s$ , Light blue:  $FP_t$



(a) Scenario at frame number 207 (b) Scenario at frame number 362 (c) Scenario at frame number 390

Figure 4.10: Different scenarios from dataset 2. Top: Camera view of the scenario, Bottom: Data representation and model fitting in ego-frame, Inset: lane detection result in BEV. Red: GPS data and its model, Green: map data and its model, Blue: vision data and its model.

This gives the peak  $E_t$  at frame number 390 (the inversion point of the reverse curve) and a ratio of 0.147:0.329:0.524, where this indicates vision is the most unreliable source since other two models are coherent following the same direction.

#### 4.4.3 Dataset 3 - 2011\_09\_26\_drive\_0042

The error evolution observed in this dataset is shown as result in Fig. 4.11. Because of the long straight highway section between frame number 50 and 890 provides low consistency errors.  $FP_v$  and  $FP_m$  are observed in this section and most of the error in this section is

from vision due to challenging light conditions and presence of markings from road splits and road merges.

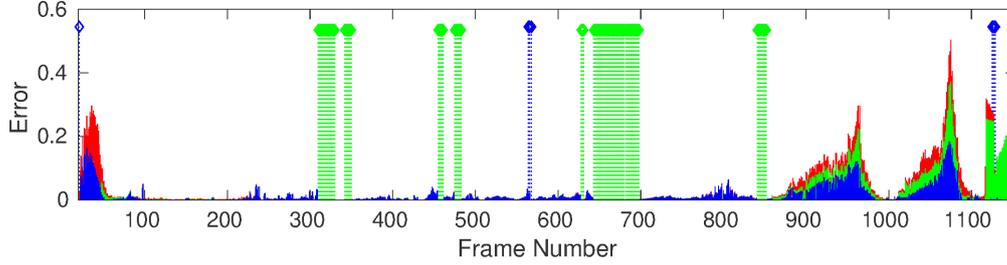
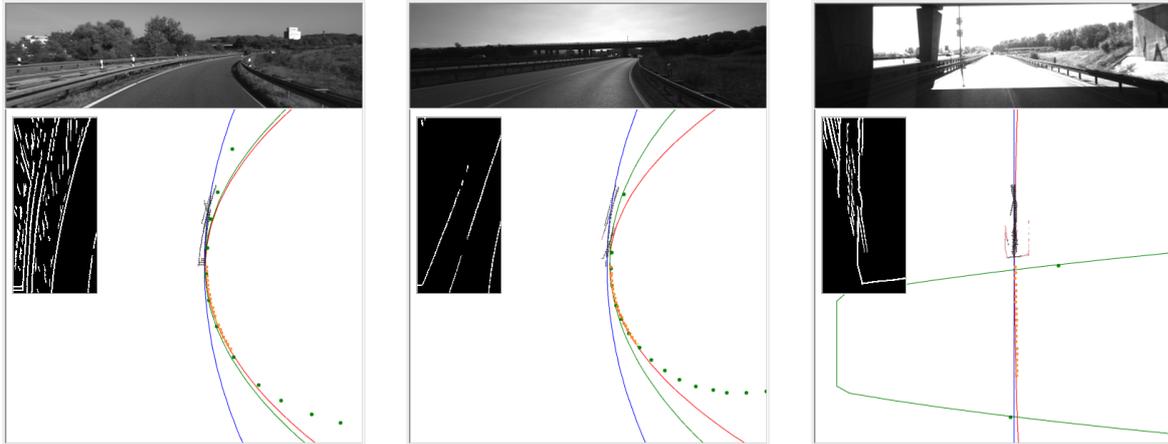


Figure 4.11: Error evolution for Dataset 3. Red: GPS error, Green: Map error, Blue: Vision error, Green dotted lines:  $FP_m$ , Blue dotted lines:  $FP_v$



(a) Scenario at frame number 962 (b) Scenario at frame number 1075 (c) Scenario at frame number 1140

Figure 4.12: Different scenarios from dataset 3. Top: Camera view of the scenario, Bottom: Data representation and model fitting in ego-frame, Inset: lane detection result in BEV. Red: GPS data and its model, Green: map data ad its model, Blue: vision data and its model.

After Frame number 890, the vehicle enters a continuous curved road split where increase in  $E_t$  is observed. As evident from Fig. 4.11, majority of the error is observed in vision data due to the limitation of IPM transformation used to create BEV images for lane detection. Since these are large radius curves, the IPM transformation fails to map the curvature of the road farther from the vehicle position. A typical example of this is shown in Fig. 4.12a. In Fig. 4.12b, vehicle leaves the curve to enter a straight road as part of a smooth road merge. Here but vision again fails to observe the change in curvature resulting in a high  $E_t$  with a  $R = 0.441:0.343:0.216$ . After frame number

1137, another section of high  $E_t$  is observed with majority of which is due to  $E_m$ . The vehicle is approaching an overpass with multiple lanes, which caused an erroneous map matching result. The integrity assessment method manages to identify this because of the inconsistent model obtained from the map data as shown in Fig. 4.12c. It is also worth noting that few  $FP_v$ s are also detected in this section when the vehicle was completely under the shadow of the overpass where no lane markings were visible.

#### 4.4.4 Dataset 4 - 2011\_09\_26\_drive\_0070

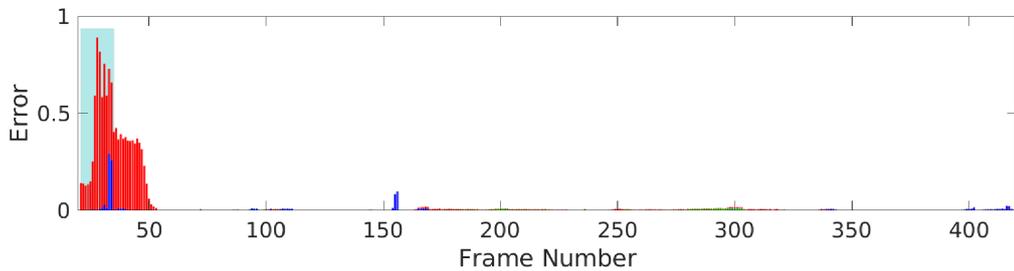
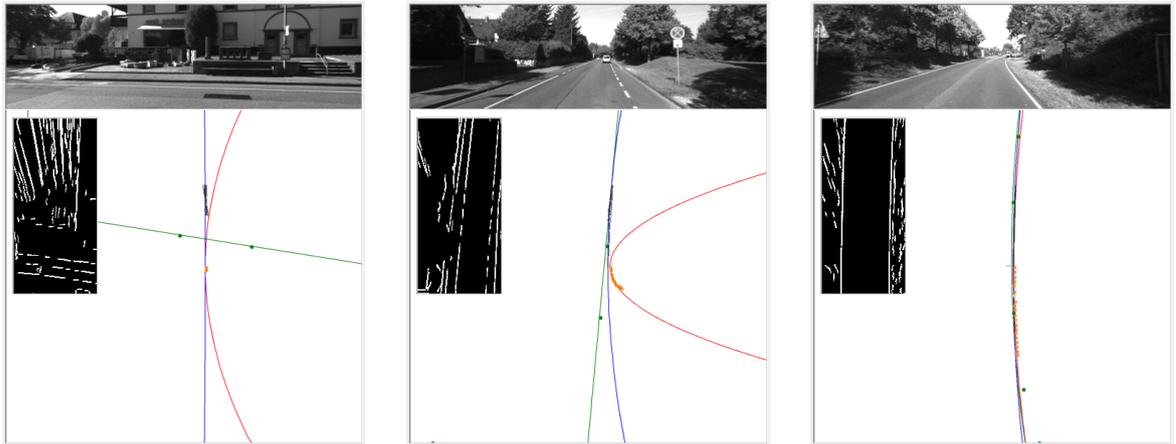


Figure 4.13: Error evolution for Dataset 4. Red: GPS error, Green: Map error, Blue: Vision error, Light blue:  $FP_t$



(a) Scenario at frame number 10 (b) Scenario at frame number 40 (c) Scenario at frame number 200

Figure 4.14: Different scenarios from dataset 4. Top: Camera view of the scenario, Bottom: Data representation and model fitting in ego-frame, Inset: lane detection result in BEV. Red: GPS data and its model, Green: map data and its model, Blue: vision data and its model.

In dataset 4, vehicle enters a straight main road from a curved side road as shown in Fig. 4.14a. Here the vision data is unusable, as the outlier removal technique employed in

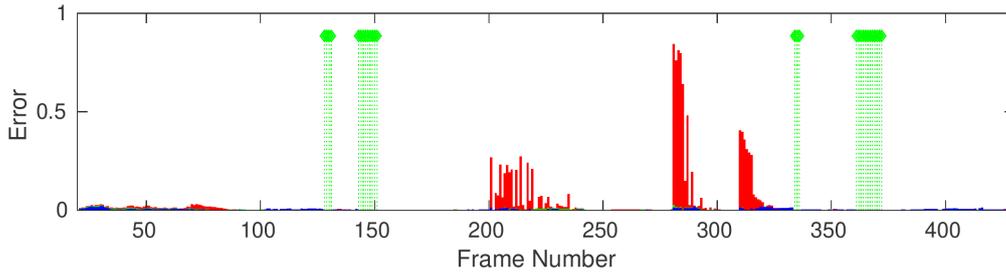


Figure 4.15: Performance evaluation on Dataset 1. Red: GPS error, Green: Map error, Blue: Vision error, Green dotted lines:  $FP_m$

this work treats lateral white patches as outliers. This resulted in discarding the actual lane markings on the road for model estimation. However, until the vehicle is completely entered on the main road,  $FP_t$  is produced, which helps to avoid the inexplicable errors observed from each source. But from Frame number 35, although vehicle has completely entered the new road after the turn, a high  $E_g$  is observed until frame number 50. In this scenario, GPS history is not along the path in which vehicle is currently moving while vision and map data are consistent as shown in Fig. 4.14b. It can be observed that  $E_g$  is continuously decreasing as the GPS history increasingly represent the actual road on which vehicle is moving as shown in Fig. 4.14c. Except errors observed in lane marking detection due to shadows around frame number 155, this dataset exhibited data sources of high integrity due to the consistency in models as shown in Fig. 4.13.

#### 4.4.5 Performance Evaluation

In order to assess the performance of the proposed integrity assessment method, errors and faults are introduced to each source under controlled conditions to characterize the behavior of the proposed method. To simulate errors in GPS, outliers are introduced whereas random road sections are selected from the map to simulate erroneous map inputs. Obtained vision markings are rotated by randomly generated angles, to serve as instances of poor detections from vision systems. In Table 4.1, an example of performance evaluation on Dataset 1 is shown. The errors mentioned above are applied to different sets of frame numbers in the dataset. Each set contains frames at fixed intervals (every 50 frames). Errors introduced in GPS are measured as the distance between ground-truth GPS position and the generated noisy GPS position. The angle between the actual road section and randomly selected road sections characterizes the errors introduced in the

map.

The angle used to rotate all lane marking detections is used as the measure of error introduced in vision data. Total error  $E_t^*$  and the ratio of errors  $R^*$  obtained from the modified dataset are compared to that of the original dataset ( $E_t$  and  $R$  respectively). Not only we observe higher  $E_t$  in each case, but also the increment observed is correlated to the amount of error introduced. The same can be stated for  $R$  as well. At frame number 50, the introduced GPS error is 1.684 m, which constitutes an increment of 0.0071 in  $E_t$  and changes  $R$  from 0.511:0.175:0.314 to 0.628:0.130:0.242. But for an introduced GPS error of 9.862 m at frame number 350,  $E_t$  has increased 0.2852 and a distinctive change in  $R$  from 0.122:0.048:0.830 to 0.965:0.008:0.026 is observed. Similar associations can be observed for other sources and the behavior can be explained using the measure of the errors introduced.

In the next experiment, GPS source is focused to establish the performance of the proposed method on two different GPS localization scenarios. Two sections of errors are introduced in Dataset 1: i) continuous outliers for a trajectory length (30 fixes) more than the length of GPS history used in this work (i.e, 20 fixes) and ii) a continuous added lateral offset (length: 30 frames) to actual GPS localization fixes. The former can simulate the behavior of GPS in a tunnel, whereas the latter can be approximated as the behavior of GPS in an urban canyon due to multi-path effects. The result of this experiment is shown in Fig. 4.15. Frame numbers from 200 to 230 contain introduced outliers and depending on the magnitude of them, the process immediately responded and higher  $E_t$  and higher ratio for GPS in  $R$  is observed. But when offset is introduced between frame number 280 and 310, two distinct high  $E_t$  sections are observed instead of one. When the offsets lasts more than the length of GPS history used, entire GPS history used for the process can contain only measurements with offsets. Since this method uses curvature consistency between models and discards consistency in offsets, such cases may observe good cross-consistency thus providing lower  $E_t$ . After frame number 290, we can observe the decrease in  $E_t$  due to the fact that more than half of the points in GPS history are with offset, hence resulting in a consistent model to other sources. The opposite is observed once the offset is removed at frame number 310. Until frame number 320, the model estimated corresponds to the GPS points with offset present in the history which leads to high  $E_t$  values. This is a limitation attributed to the proposed method. This

Frame no.	GPS error (m)	Map error (rad)	Vision error (rad)	$E_t$	$E_t^*$	$R$ $R^*$
50	1.684	0	0	0.0217	0.0288	0.511 : 0.175 : 0.314 0.628 : 0.130 : 0.242
60	0	1.741	0	0.0076	0.5683	0.268 : 0.263 : 0.467 0.003 : 0.992 : 0.004
70	0	0	1.434	0.0295	0.0927	0.690 : 0.080 : 0.229 0.202 : 0.023 : 0.774
100	7.339	0	0	0.0023	0.3063	0.037 : 0.230 : 0.733 0.968 : 0.012 : 0.020
110	0	2.048	0	0.0066	0.0017	0.034 : 0.006 : 0.960 0.058 : 0.335 : 0.607
160	0	-1.643	0	0.0006	0.2528	0.091 : 0.037 : 0.872 0.002 : 0.997 : 0.001
170	0	0	-1.135	0.0004	0.1336	0.206 : 0.040 : 0.754 0.001 : 0.001 : 0.998
200	9.339	0	0	0.0037	0.3246	0.030 : 0.021 : 0.948 0.955 : 0.003 : 0.042
210	0	1.048	0	0.0021	0.0016	0.074 : 0.045 : 0.880 0.085 : 0.483 : 0.431
220	0	0	1.124	0.0009	0.0889	0.338 : 0.584 : 0.078 0.003 : 0.004 : 0.993
270	0	0	1.079	0.0032	0.0532	0.413 : 0.156 : 0.431 0.019 : 0.007 : 0.974
300	4.637	0	0	0.0016	0.0991	0.474 : 0.014 : 0.512 0.962 : 0.002 : 0.036
310	0	1.515	0	0.0014	0.2750	0.099 : 0.014 : 0.887 0.0001 : 0.998 : 0.001
320	0	0	0.529	0.0037	0.0082	0.105 : 0.007 : 0.887 0.044 : 0.004 : 0.952
350	9.862	0	0	0.0011	0.2852	0.122 : 0.048 : 0.830 0.965 : 0.008 : 0.026
360	0	1.768	0	0.0026	0.1795	0.020 : 0.026 : 0.954 0.001 : 0.9987 : 0.002
400	3.088	0	0	0.0074	0.0859	0.112 : 0.016 : 0.872 0.879 : 0.007 : 0.112
410	0	-0.083	0	0.0067	0.0089	0.149 : 0.031 : 0.820 0.141 : 0.080 : 0.779
420	0	0	0.644	0.0040	0.0134	0.163 : 0.068 : 0.769 0.043 : 0.016 : 0.941

Table 4.1: Performance evaluation on Dataset 1

can be avoided by using an offset tracking mechanism between data sources.

## 4.5 Conclusions

In this part of the work, a proof-of-concept is provided for the proposed novel approach to the integrity assessment of data sources used in autonomous navigation. The method is adaptable to a variety of sensor modalities and data sources that can provide spatial information about the surroundings of a vehicle. It also does not depend on the choices of preprocessing methods, detection algorithms or data association methods. There are only two requirements to apply this integrity assessment method to any sensor combinations and situations: i) possibility of representing all considered data sources in a common frame and ii) formulating a model that can be used to represent all individual data sources in the chosen common frame.

In our experiments, vision data from a camera, localization fixes from a GPS sensor and geographical information from a digital map are used as data sources. A second-degree polynomial is used as the model for all the data sources. By conducting experiments on different standard datasets, it has been shown that the proposed method could provide a reliable integrity marker for each data source by exploiting the cross-consistency between them. Integrity marker for vision data can account for the effect of quality of lane markings detected, erroneous detection, lack of information and challenging lighting conditions. Integrity marker for GPS data monitors the quality of localization fixes and its suitability in using them at a particular situation, whereas integrity marker obtained for map data encompasses the effect of map matching process and adequacy of geographic information available from the map. Along with integrity markers, a set of fault predictors is also monitored in the proposed process which acts as a self-assessment marker for the process itself. The novelty of the proposed method lies in the fact that instead of a fault detection and isolation (FDI) framework currently used in integrity monitoring of localization applications, a set of more generalized and handcrafted fluid integrity markers can be used, that can account for a variety of situations without the use of empirical thresholds or rigid logic systems. Hence, instead of bi-state or tri-state fault predictions on each data source, the proposed method provides a weighting scheme for data sources which corresponds to the accuracy of data they provide. Along with the

Fault and Feasibility Predictors which corresponds to the validity of data sources, to the best of the author's knowledge, the proposed method is an improvement upon the current state-of-the-art sensor data integrity assessment tools. In the next chapter, we expand the scope of this framework to accommodate complex urban scenarios, and propose the integrity monitoring of perception data sources used in such cases.

# Chapter 5

## The Generalized Prototype

### Contents

---

<b>5.1</b>	<b>Introduction</b>	<b>86</b>
<b>5.2</b>	<b>Methodology</b>	<b>89</b>
5.2.1	Detection	90
5.2.2	Map Handling	94
5.2.3	Data Representation	95
5.2.4	Integrity Analysis	98
5.2.5	Transformation Optimization	100
5.2.6	Calculation of Protection Levels	103
<b>5.3</b>	<b>Experiments and Discussions</b>	<b>105</b>
5.3.1	Integrity Markers comparison	106
5.3.2	Complex Situations	110
5.3.3	Performance of Integrity Monitoring	112
<b>5.4</b>	<b>Conclusion</b>	<b>114</b>

---

## 5.1 Introduction

Considering the multi-modality of data provided by the wide varieties of sensors used in urban vehicle localization, finding a framework to evaluate integrity is a challenging yet crucial task. In Chapter 4, we addressed this task using a cross-consistency based integrity monitoring framework well-suited for highway scenarios. In this chapter, we propose a framework to apply on complex semi-urban and urban scenarios in a generalized way, thus providing context-awareness to a multi-modal vehicle localization system. Here, generalization implies that the framework developed does not depend on the number of sensors or the features used. Semi-urban and urban environments often contain multitude of intersections, roundabouts, road-splits and merges compared to highway scenarios. As discussed in Chapter 2, the use of multimodal data sources is the key strategy to achieve reliable localization in urban and semi-urban scenarios considering the usability of GNSS is limited. Developing upon the framework presented in Chapter 3, finding a generalized common model for the representation of data from all sources is the primary aim. Even though works like Ballardini, Cattaneo, and Sorrenti (2019) and Mueller et al. (2011) propose geometrical models for several types of intersections, they are limited to single perception data source and digital maps. They also require prior classification of intersections to reliably fit the predefined models to the data. On the other hand, sensors adopted in intelligent vehicles and considered in this study have considerably different properties (field of view, precision, measurement noise) and output in such scenarios. Hence, the rest of this section is focused on how data from different sources are used in complex scenarios. We also examine the possible errors associated with these use-cases and discuss the applicability issues of a common geometrical model (eg. polynomial model) in these situations.

Traditionally, vision data is used to detect road structure elements such as ego lane markings and/or lanes parallel to the ego lane using a curvature based model. In urban scenarios, such lane detection models fail due to different types of lane markings (eg.: stop lines, road separation markings etc.), orientation (eg.: lane markings from other road sections in the junctions) complex curvatures (eg.: splitting and merging lane markings) and occlusions due to traffic. Another approach using visual data is to detect the drivable road region in front of vehicle. But due to the unforeseeable shapes of

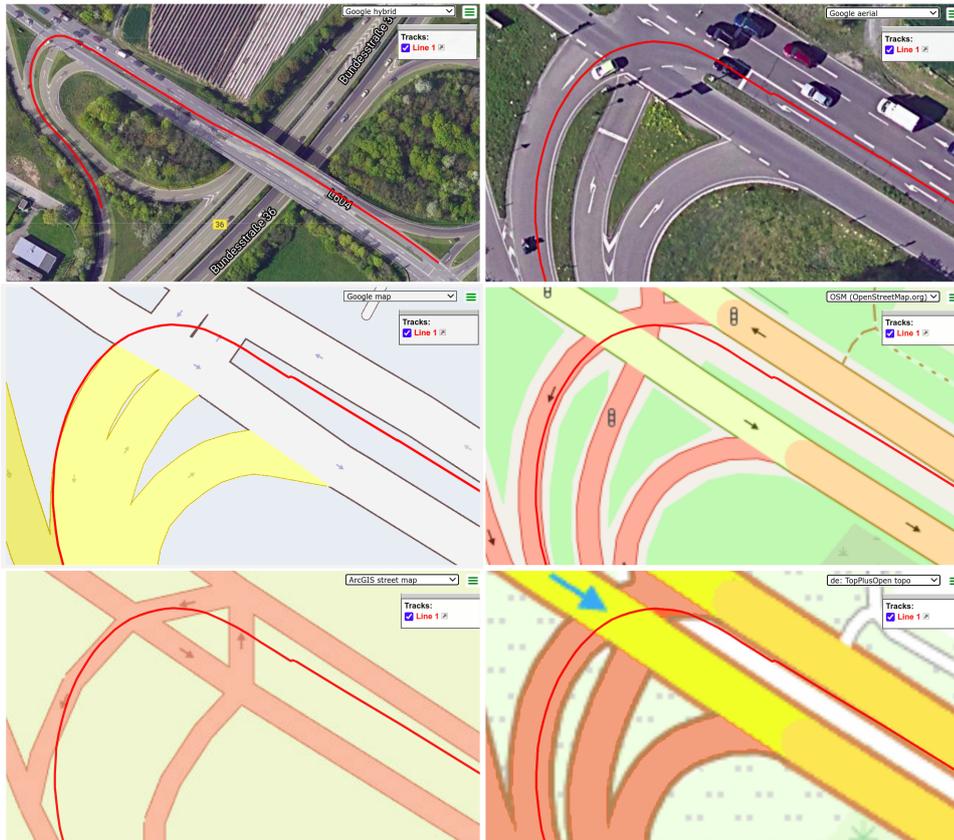


Figure 5.1: Integrity issues in map sources. Top-left: An example of GPS track (in red) from KITTI dataset projected on satellite map from Google. Top-right: Zoomed aerial view of the track at an intersection. Middle-left: The intersection in street map from Google. Middle-right: The intersection in street map from OpenStreetMap. Bottom-left: The intersection in street map from ArcGIS. Bottom-right: The intersection in street map from Federal Agency for Cartography and Geodesy (BKG) of Germany.

possible road segment detections, modeling of such output with a geometrical model is difficult. Intersections with multi-lane branch roads can have a large common region at the center, which can limit the observability of other road branches through visual inputs.

It is reasonable to assume that vehicles travel slowly and stop more often in semi-urban and urban scenarios than highways. GPS receivers are proved to have poor performance in slow-moving vehicles (Toth et al. 2017). Combined with the fact that the presence of buildings and other obstructions can cause multi-path effects or even outages of signals, GPS receivers experience classical localization problems in urban environments as described in Chapter 1.

With the exception of a few advanced and proprietary Geographic Information Systems (GIS, eg.: Google maps), publicly available GIS sources lack accurate road properties (lane or road widths, locations of lane splits and merges at junctions etc) and

strongly depend on rule based rendering to display maps. The discrepancies observed while overlapping the satellite view and rendered map structures from different GIS as shown in Fig. 5.1 are examples of the limitation of this approach. GPS trajectory of the vehicle is accurate in satellite view of the junction, which includes a lane change to the left-most lane of the highway for a left turn and a smooth turn through the left side of link road. However, from the rendered road structure view of all the map sources, the track section corresponding to lane change appears to be wrong as it is outside the boundary of the road structure. It is also worth noticing that none of the GIS sources shown in Fig. 5.1 represent roads with their actual width, but with rule-based dimensions. It is evident from the same width of two highway sections despite different number of lanes in each of them. Likewise, modeling of junctions is also considerably different in each map source, particularly between Google Maps and OpenStreetMap. Hence, inclusion of map data in localization process is sub-optimal in urban and semi-urban scenarios and forces us to consider it as a data source with associated instantaneous integrity rather than a baseline reference source.

While data from vision, GPS and maps add complexities and impose limitations, LiDAR on the other hand, can provide useful data in urban and semi-urban environments. Here we consider top-mounted 360°3D LiDAR sensor specifically. It can observe ego road and other road branches efficiently. By using the reflectivity information available in LiDAR data, we can detect bright surfaces like lane markings and curbs (Liu et al. 2019). Though LiDAR poses challenges in detection and modeling of features like in the case of vision, the accurate 3D information available makes it an important source for representing the structure of a large urban scenario.

The integrity monitoring method in Chapter 4 provides a weighting scheme for data sources which infers the cause of inconsistencies observed in the localization method at a given time. For any data source combination that can be represented in a common frame and with a common model in that chosen frame, the cross-consistency analysis proposed in Chapter 4 can be applied. However, discussion presented in here shows that developing a common model is significantly difficult as different sensor modalities and diverse features are introduced to the system to accommodate urban scenarios. To this extent, we could not find any integrity assessment solution from the literature study, that can handle more than two perception data sources and wide varieties of scenarios. This

chapter presents three main contributions with regard to this problem statement.

1. Defining a common reference frame and formalizing common model to represent all data sources in all scenarios.
2. Prototyping an integrity assessment framework using the common model.
3. Analyzing the performance of the proposed framework using publicly available datasets and comparison with other state-of-the-art integrity monitoring solutions from the literature.

## 5.2 Methodology

The framework proposed for the integrity assessment developed is given in Fig. 5.2. The main prerequisites and specifications for developing this framework are listed below.

- Four SDS sources that are commonly adopted for localization in urban environments are considered - vision, LiDAR, digital map and GPS.
- The position and orientation of the vehicle is available with enough precision to roughly perform reliable map-extractions and its transformation to ego-frame. Hence, there is a minimal localization requirement to achieve the integrity of the sources.
- The data sources are synchronous, ie., at any given time step, all sources provide data corresponding to the environment at that time step.

The Detection Block includes sensor-specific routines to detect features that are relevant to different localization algorithms described in Sec. 5.1. The Rendering Block uses GPS position to extract data from surrounding map region and applies rule-based rendering to reconstruct the geometrical structure of the area. Obtained information is represented in a common frame using a common model. Common reference frame is chosen as ego-frame of the vehicle as the transformations between ego-frame, camera-frame, LiDAR-frame and GPS-frame can be determined by calibration procedures (Pusztai, Eichhardt, and Hajder 2018). A decision algorithm is used to decide whether optimization of localization is required in case of unknown or unreliable transformations between frames of data

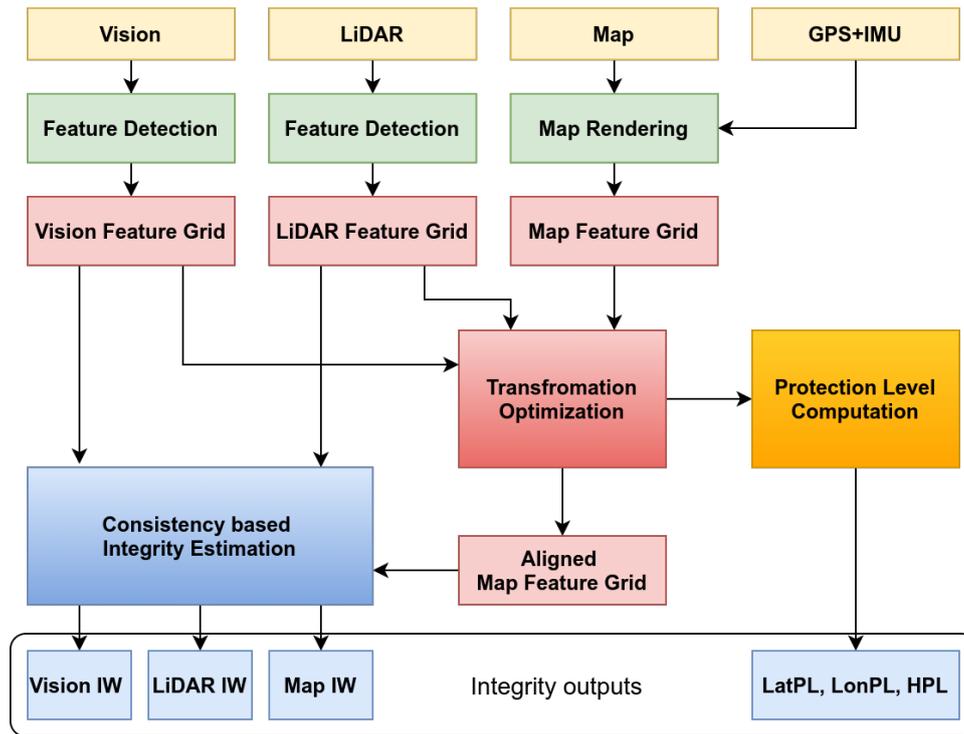


Figure 5.2: Framework for integrity assessment of multi-modal data sources

and the common reference frame (in our case, map-frame to ego-frame). Once the required optimization is achieved, consistency between data representations is evaluated and integrity is assessed for each source. In this section, the specific techniques and concepts used in the framework presented in Fig. 5.2 are detailed.

### 5.2.1 Detection

The purpose of the feature detection blocks is to extract common and representative information (semantic features) from each data source. From the literature review presented in Chapter 2, we identify three classes of semantic information that are most commonly used in state-of-the-art localization methods in urban scenarios - markings on the road, drivable road region and structure of the surroundings of the vehicle. The definitions and properties of each of these classes and examples of features in each class used in this work are as follows:

- **Markings on the road** - All kinds of painted portions on the road. Includes lane markings (both continuous and discontinuous), stop line markings, pedestrian crossing markings, direction signs on the road. Generally, the size, color and

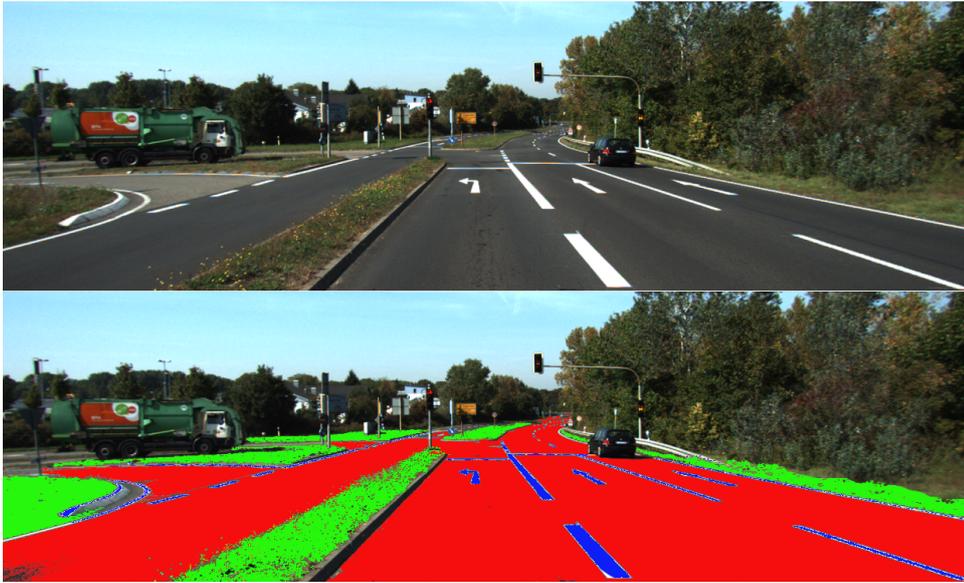


Figure 5.3: Example of features used in our work, detected from a complex scenario. Red: drivable road regions, Green: vegetation/grass-patches, Blue: markings on the road

purpose of these markings are strictly defined by national-level (in some cases, international) standards and in our work this is assumed as prior knowledge. These features have distinct colors (bright white or yellow) and high reflectivity.

- Drivable road regions** - The obstacle-free space on all the roads appear in the FOV. Once the map nodes are represented in ego-frame, rule-based rendering algorithm is used to create geometrical sub-map for the ROI. Number of lanes, lane width (when of a sensor. The assumptions here are that the roads are flat (same ground plane) and it is parallel to X-Y plane of the vehicle's ego-frame. Roads typically have black and other darker gray colors (asphalt, tarmac, concrete etc.) and their reflectivity will be less. We also assume that we have access to the standardized rules on road geometry by the country of interest.
- Structure of the surroundings** - While this can include all possible permanent infrastructure and road structures such as buildings, road railings, barriers, dividers, curbs and vegetation etc., we limit our work to the last one due to ease of detection. Vegetation/grass patches near and around the road structure have two main properties that can be easily exploited for detections. They have colors close to shade of green and they are at higher heights than the road plane, often providing the boundaries for roads.

The detections are also limited to a pre-defined 3D region of interest (ROI) around the vehicle in ego-frame. The methods used to detect these features from each source are explained hereafter.

### 5.2.1.1 Vision

To accommodate varieties of lane markings present in different scenarios, all possible markings are detected. Images from camera are transformed to bird eye view (BEV) using camera calibration. Intensity based segmentation is used to detect all possible lane markings as shown in Fig. 5.4. After detection of all the candidate lane markings, blob analysis is used to estimate their width, size and orientation (Nedevschi et al. 2013). Based on the prior knowledge about the properties of lane markings in the real world, unreliable candidate detections are removed. In Fig. 5.4, markings that are narrower or wider than the width of typical lane markings are removed. Seed-based fast-marching method proposed by Sethian (1998) is used to detect dark road regions and regions with grass/vegetation patches. To accomplish this, a set of seed points are equidistantly distributed in the image and based on the color properties of their neighborhood, they are classified into seed points for roads (darker shades of gray) and seed points for grass/vegetation (shades of green) as shown in Fig. 5.4. These seed points are used to grow connected regions based on similarity in color, using fast-marching segmentation algorithm. An example of the segmentation result is shown in Fig. 5.4. Note that wrong seed categorization can affect the quality of the segmentation, as misidentified seed points caused some part of the green-patches to be detected as road in Fig. 5.4. Collectively, detection of these three features ensure that every pixel in the BEV image classified into members of one or more features, or unclassified.

### 5.2.1.2 LiDAR

A subset of LiDAR data containing points which are inside the 3D ROI is selected. Points on the road and on the edges of the road are classified using 3D gradients using the method proposed by Liu et al. (2019). Along every scan-line of LiDAR data, local 3D gradients are calculated based on consecutive points. Since a certain height difference between vegetation and road plane is present, a sudden gradient change is observed at point of

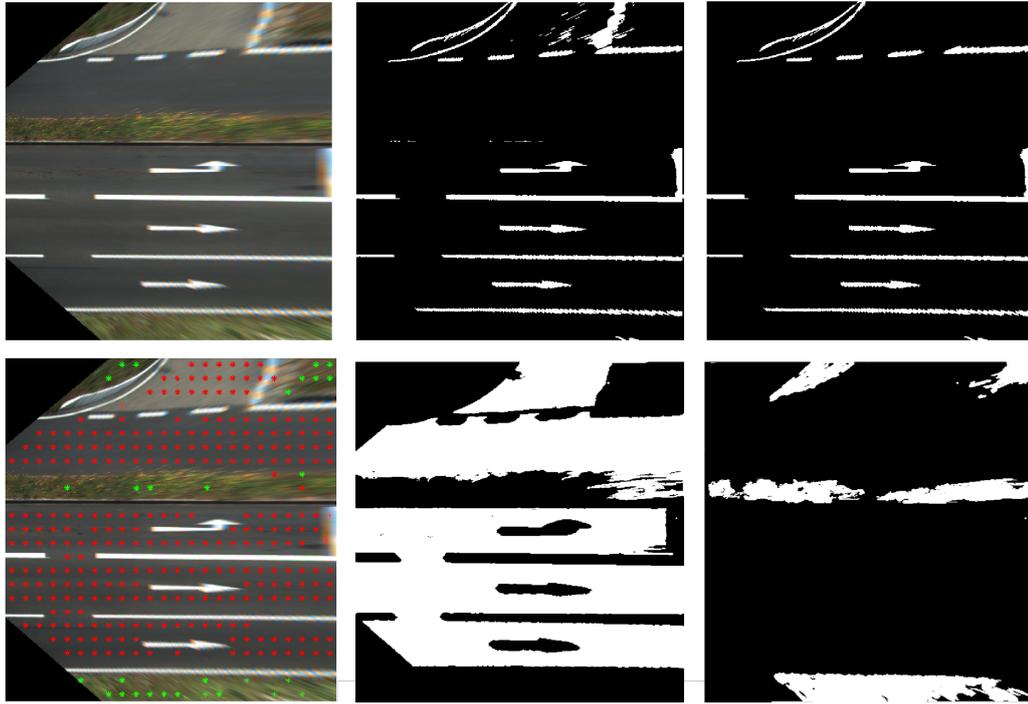


Figure 5.4: Steps for detecting features from vision. Top row from left - 1) original image in BEV, 2) candidate lane marking detections extracted using thresholding, 3) outlier removed lane marking detection. Bottom row from left- 1) seed-points detected for road and vegetation, 2) segmentation result for road, 3) segmentation result for vegetation

discontinuity as illustrated in Fig. 5.5. Hence, points that have similar 3D gradients until point of discontinuity are classified as points corresponding to road. All the points between points of discontinuities are selected as candidate points for vegetation/green-patches. Then, the 3D ROI is divided into smaller blocks (size of the blocks decided according to the requirement of data representation presented later in this section) in XY plane and the candidate points belonging to each block are examined for their Z values. The candidate points are classified based on their Z values (heights) into groups. This helps to differentiate between road segment and vegetation using the technique presented in Xu et al. (2019).

Finally, for the lane marking detections, LiDAR data points with high reflectivity are selected as they correspond to the bright surfaces such as lane markings, railings and other reflective surfaces. Since we are only interested in markings on the road, only the points with high reflectivity and are on the road plane (based on the road detection obtained) are selected as final candidates. As the result of these detection steps, every point in the ROI is either categorized into road markings, drivable road regions and

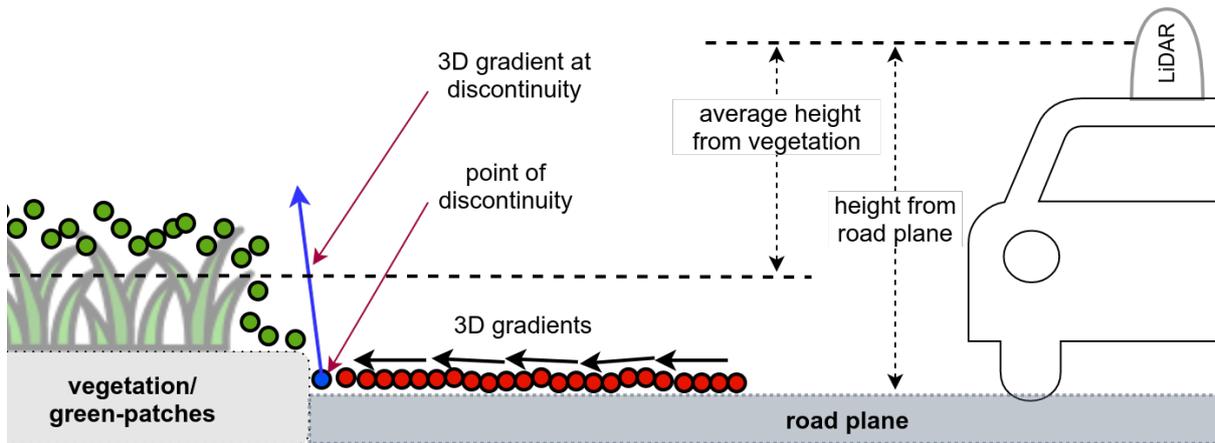


Figure 5.5: Detection criteria for LiDAR data

vegetation classes or stays unclassified.

## 5.2.2 Map Handling

OpenStreetMap (OSM) is used in this work as GIS source. OSM provides nodes corresponding to ways, grass-patches and railings etc. However, finding relevant geometrical information in a vehicle's surroundings from maps involves two key components - location and orientation of the vehicle (Li, Quddus, and Zhao 2013). Using the available localization and orientation estimate, all the relevant map nodes in the ROI are selected and the map data is transformed into vehicle's ego-frame using the orientation of the vehicle. The location estimate is provided by the GPS sensor whereas the orientation estimate is given by the on-board Inertial Measurement Unit (IMU). Once the map nodes are represented in ego-frame, rule-based rendering algorithm is used to create geometrical sub-map for the ROI. Number of lanes, lane width (when available), location of road boundaries, boundaries of vegetation etc. are used in the rendering process, producing enriched geometrical model of the environment from OSM as illustrated in Fig. 5.7c. In works like Kang et al. (2020) and Nedevschi et al. (2013), custom-made High-Definition Maps (HD Maps) which contain lane marking information and accurate road structure information are used. Even though the exact location or type of lane markings are unavailable in OSM, assuming continuous lane markings on left side of the leftmost lane, right side of the rightmost lane, and dashed lane markings for the lanes in the middle, approximate lane level information can be produced. In case of missing lane width information, standardized road construction guidelines of the country are used to render

the map. However, it is evident that errors in GPS positioning or orientation estimation can greatly affect the accuracy of map data extraction and cause uncertainties in map rendering (Li, Quddus, and Zhao 2013), especially in exact location of lane markings.

### 5.2.3 Data Representation

One of the key observation from Sec. 2.6 is the use of 2D spatial grids to model the data from vision, LiDAR and map. As data representation models, occupancy grids (Li and Ruichek 2014; Yu, Cherfaoui, and Bonnifait 2015; Capellier et al. 2018), scan-grids (Moras, Cherfaoui, and Bonnifait 2011) and free-space representation of maps are similar concepts. They are discretized spatial representations of obstacles and free-space using binary states or probability. While this concept is generally used in data fusion/association for building environment maps from multiple sensor inputs, it is recognized in the literature as one of the most efficient way to represent spatial data (modeling spatial data) from various sources. At this point, it would be beneficial to note the difference in using grids for data fusion and integrity monitoring.

Fundamentally, data fusion process combines the data from various sensors to define the occupancy of every cell in a grid. However, when used for integrity monitoring in this work, no such fusion of information between data sources are executed. Each source produce separate grids containing the information about features detected strictly using its own data as shown in Fig. 5.6. Since detected features are the data from the sources considered in the work, we substitute 'state of features' instead that of obstacles, with a labeling scheme. To be able to deal with the features and geometries of different types and shapes, a 2D Feature Grid (FG) is proposed as the model. FG consists of array of cells where each cell represents a block of pre-defined size in the real world as illustrated in Fig. 5.6. Four feature labels (LB) are assigned to cells in FG according to the type of the feature:

1.  $LB_r$  - drivable road
2.  $LB_l$  - lane markings, signs on the road
3.  $LB_o$  - other surfaces such as dividers, green patches, road barriers
4.  $LB_u$  - unclassified/unidentifiable.

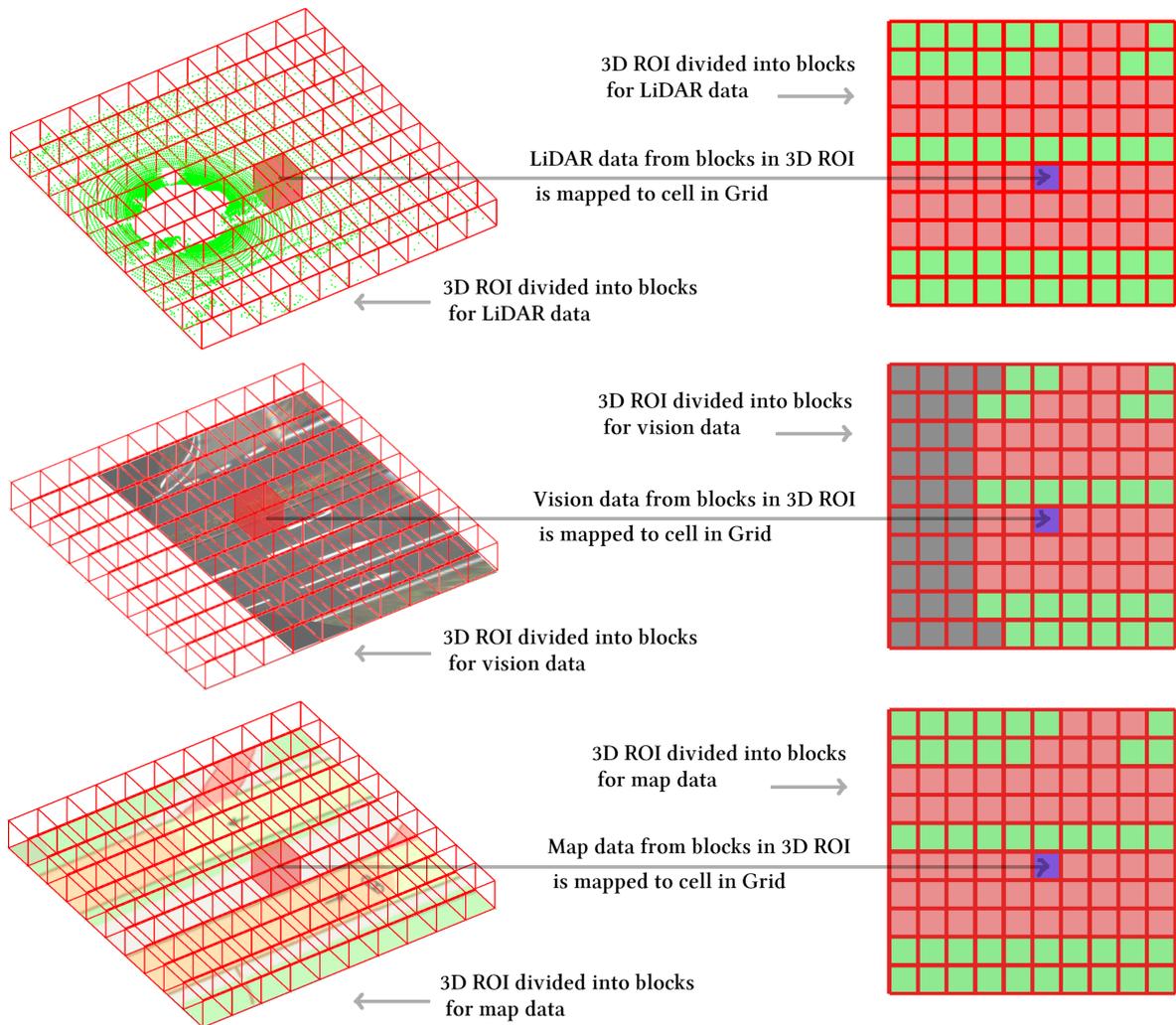
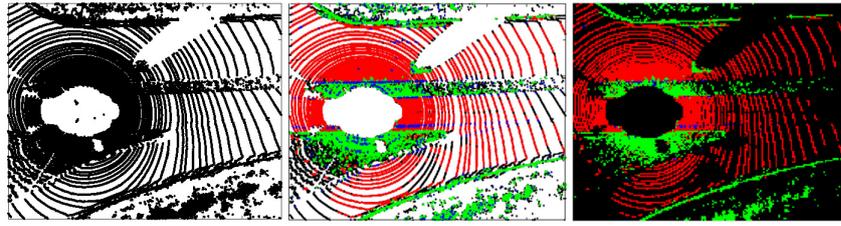


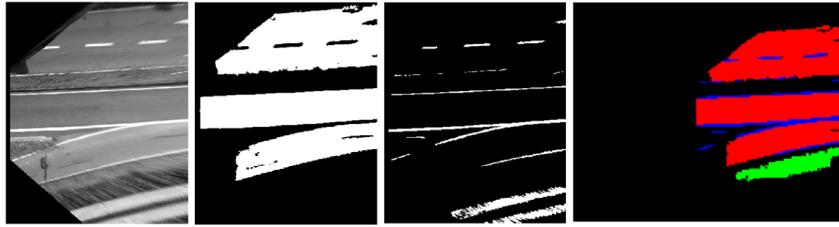
Figure 5.6: 3D ROI to grid representation for integrity monitoring of sources

The blocks are examined for the information they contain. The type of feature with highest contribution to a block is used to assign respective label to the cell. An example of the creation of FG from each data source following this criterion is shown in Fig. 5.7.

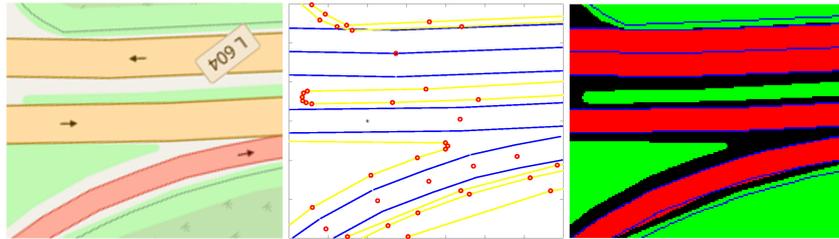
Along with labels, it is important to model the intrinsic confidence of sources (independent of use-conditions such as exposure, weather etc. ) of data provided by each sensor. Accuracy of LiDAR data decreases as the distance from the sensor to the measurement location increases Zheng et al. 2008. On the other hand, the Inverse Perspective Mapping (IPM) transformation used to create the bird-eye view images from actual images, increasingly introduces deformation as the distance from camera increases due to camera calibration errors. To account for these facts, a confidence function is proposed drawing inspiration from curvature-based weighting used in Sec. 4.3.1 for all



(a) Left: LiDAR data in ROI. Middle: Detection of road (red), lane markings (blue), other surfaces (green) and unclassified points (black). Right: Feature Grid from LiDAR



(b) From Left, 1- BEV image. 2- Output of road detection. 3- Output of lane detection. 4- Feature Grid from vision



(c) Left- Rendered OSM map from official website. Middle- Available map data in ROI. Right- Feature Grid from Map

Figure 5.7: Example of modeling data from different sources using Feature Grid representation: cells with road labels (red), cells with lane marking labels (blue), cells with other surfaces labels (green), cells with unclassified labels (black).

relevant FGs. Using the concept of Inverse Distance Weighting (IDW) function presented in Feng et al. 2017, the weights are computed as,

$$\omega_{ij} = 1 - \left\langle \sqrt{(x_{ij}^2 + y_{ij}^2 + h_s^2)} \right\rangle_{min-max} \quad (5.1)$$

where  $\omega_{ij}$  is the weight associated to the cell  $C_{ij}$ ,  $x_{ij}$  and  $y_{ij}$  are the distances to the center of  $C_{ij}$  from sensor position and  $h_s$  is the height of the sensor. The min-max normalization operator  $\langle \hat{x} \rangle_{min-max}$  is defined as

$$\hat{x}_{norm} = \frac{\hat{x} - \hat{x}_{min}}{\hat{x}_{max} - \hat{x}_{min}} \quad (5.2)$$

Hence, the total representation of data from sources will have two components: the

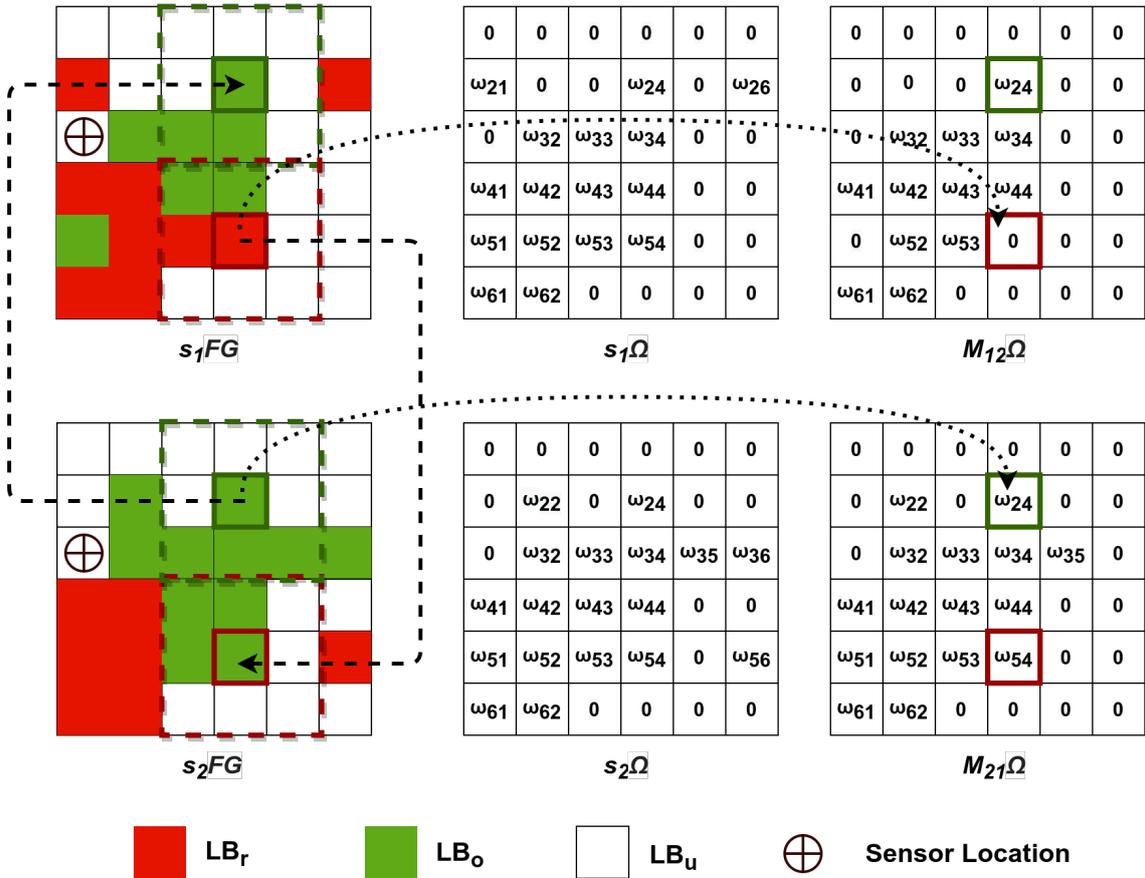


Figure 5.8: Matching operation between FGs of two sources

labels and their associated uncertainty which is denoted by the weights. Other source-specific weighting functions using homography of image transformations and LiDAR data acquisition model can also be used for this purpose. However, data sources like maps use uniform weights for all the cells in their FGs due to the fact that they are not measured but just extracted. In fairness, the map rendering process can contain certain amount of uncertainty which needs to be incorporated in its FG. However, this is out of the scope of this work and discussed further in Conclusion and Perspective chapter.

## 5.2.4 Integrity Analysis

Semantic data representation of the multi-modal data sources with a common-frame of representation and same dimensionality, allows us to transpose the definitions of integrity presented in Chapter 3. As hypothesized in Chapter 3 and supported by the results from Chapter 4, consistency between representations (model) of data from different sources enables to estimate accuracy of data, which is a key attribute of integrity. Data sources

with high consistency can be treated as reliable and their integrity can be expressed as a function of their consistency with other sources.

Let  $S = \{s_1, s_2, s_3, \dots, s_N\}$  be the set of  $N$  sensors and  $s_pFG$  be the feature grid provided by each sensor where  $p \in \{1, 2, 3, \dots, N\}$ . Labels are denoted as  $LB_x$  where  $x \in \{r, l, o, u\}$  (r- road, l- lane marking, o- other surfaces, u- unclassified). The cells  $C_{ij}^{s_p\Omega}$  in the weight grid  $s_p\Omega$  associated with  $s_pFG$  is computed using Eq. 5.3 (refer Fig. 5.8).

$$C_{ij}^{s_p\Omega} = \begin{cases} 0 & \text{if } C_{ij}^{s_pFG} = LB_u \\ \omega_{ij} & \text{otherwise} \end{cases} \quad (5.3)$$

Two different FGs  $s_pFG$  and  $s_qFG$  are taken for the consistency analysis at a time. One cell  $C_{ij}^{s_pFG}$  with feature label  $LB_x$  from  $s_pFG$  is defined as consistent if there is at least one matching cell with  $LB_x$  in a  $3 \times 3$  neighborhood around the cell  $C_{ij}^{s_qFG}$  in  $s_qFG$ . This operation is denoted as a consistency-check operator  $\triangleright$  and gives a binary output as,

$$C_{ij}^{s_pFG} \triangleright C_{ij}^{s_qFG} = \begin{cases} 1 & \text{if Consistent} \\ 0 & \text{otherwise} \end{cases} \quad (5.4)$$

In ideal cases, matching between two grids to check their consistency, implies cell to cell equivalency. However, accounting to the possible errors in blocks-to-cell label association, the neighborhood of the cells are considered which relaxes the spatial constraints for matching. It is worth noting that the consistency-check operator  $\triangleright$  is not commutative, as illustrated in Fig. 5.8. Hence the match of  $s_pFG$  to  $s_qFG$  will be different from match of  $s_qFG$  to  $s_pFG$ , thus providing different consistency estimations. Based on this consistency examination, a matched weight grid  $M_{pq}\Omega$  as shown in Fig. 5.8 is computed, where its cells  $C_{ij}^{M_{pq}\Omega}$  are computed using Eq. 5.5.

$$C_{ij}^{M_{pq}\Omega} = \begin{cases} \omega_{ij} & \text{if } C_{ij}^{s_pFG} \triangleright C_{ij}^{s_qFG} = 1 \\ 0 & \text{otherwise} \end{cases} \quad (5.5)$$

By extension, a matching operation  $f_m$  between FGs is defined as,

$$f_m(s_p FG, s_q FG) = \frac{\sum_{\forall i,j} C_{ij}^{Mpq\Omega}}{\sum_{\forall i,j} C_{ij}^{sp\Omega}} \quad (5.6)$$

After computing the matches between all the possible combinations (six combinations in our case: LiDAR-map, vision-map, map-vision, map-LiDAR, vision-LiDAR and LiDAR-vision), the integrity weight (IW) associated with a source  $IW_p$  is computed using Eq. 5.7.

$$IW_p = \frac{\sum_{\forall q, p \neq q} f_m(s_p FG, s_q FG)}{\sum_{\forall p, q, p \neq q} f_m(s_p FG, s_q FG)} \quad (5.7)$$

### 5.2.5 Transformation Optimization

The integrity analysis mentioned in Sec. 5.2.4 assumes that the localization of vehicle is accurately known, ie., the location and orientation information used in map extraction and transformation are reliable. But in real world applications, GPS positioning - even from inertial/dead reckoning coupled GPS receiver - can have errors due to multipath effects, outages or drifts. Inherently, error in localization and orientation affects consistency analysis of map data and other sources, hence impacting the integrity of whole system. Hence, a transformation optimization procedure is proposed that uses semantic level information from data representations of sources. It can efficiently allow integrity assessment and also identify particular defaults such as map offsets or inconsistent map sections.

In this work, a particle filter based on the work of Sandhu, Dambreville, and Tannenbaum (2008) is developed for map-matching to optimize the transformation. The steps required for the transformation optimization in ego-frame of the vehicle with decision criteria is given in Algorithm 5.1.

**Algorithm 5.1** Algorithm for localization optimization

---

Inputs: Current localization:  $X_{state}$ , GPS+IMU localization measurement:  $x_m$ , FG of LiDAR:  $LFG$ , FG of Vision:  $CFG$ , FG of Map:  $MFG$ , Minimum consistency limit:  $limit$

```

if  $X_{state}$  and  $x_m$  are consistent then
  if  $f_m(LFG, MFG) > limit$  and  $f_m(CFG, MFG) > limit$  then
    Output: Integrity markers
    Update  $X_{state}$ 
  else
    Compute:
     $y_L^* = \arg \max_y (f_m(LFG, t(MFG, 0, y, 0)))$ 
     $y_C^* = \arg \max_y (f_m(CFG, t(MFG, 0, y, 0)))$ 
     $(x_L^*, \theta_L^*) = \arg \max_{x, \theta} (f_m(LFG, t(MFG, x, y_L^*, \theta)))$ 
     $(x_C^*, \theta_C^*) = \arg \max_{x, \theta} (f_m(CFG, t(MFG, x, y_C^*, \theta)))$ 
    if  $(x, y, \theta)_L^*$  and  $(x, y, \theta)_C^*$  are consistent then
      Output Integrity values
      Update  $X_{state}$ 
    else
      if  $f_m(LFG, CFG) > limit$  then
        Output: Integrity markers
      else
        Output: Error in map
      end if
    end if
  end if
else
  Output: Error in GPS
end if

```

---

In the first step, new position and orientation measurements from GPS and IMU are compared with the current best localization estimate. If the new measurements ( $x_m : [x_m, y_m, \theta_m]$ ) are not within the non-holonomic constraints of current state ( $X_{state} : [x_{state}, y_{state}, \theta_{state}]$ ) of the vehicle, they are detected as an outlier (Roysdon and Farrell 2017). Conversely, consistent position and orientation measurements are used to render map from the database and the consistency between FGs of map and other sources is computed. If sufficient consistency is observed (greater matching than empirically derived threshold for  $f_m(s_iFG, s_jFG)$  considering different sensors and scenarios), transformation

optimization is not performed and the data representations from each source are used for integrity assessment.

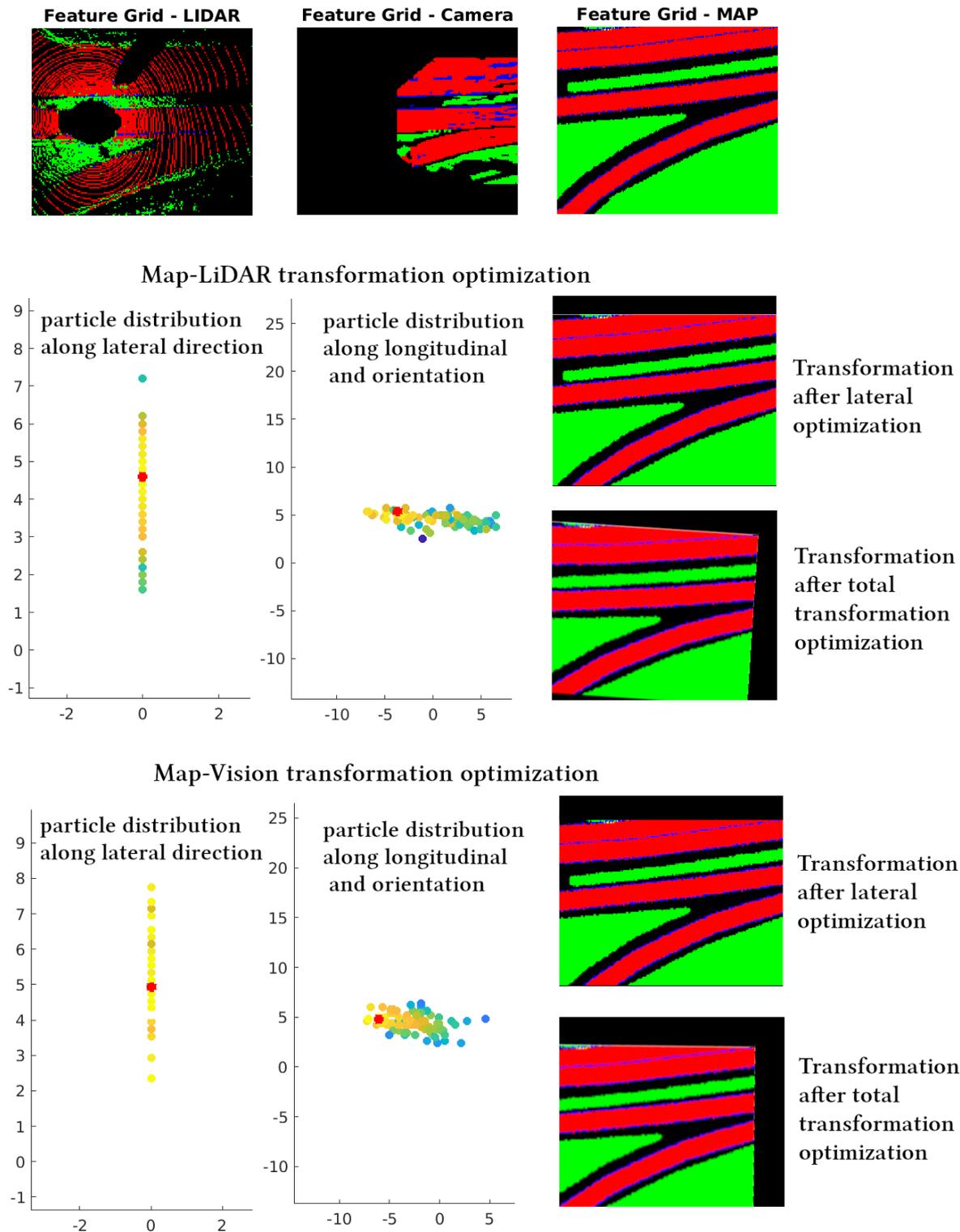


Figure 5.9: Transformation optimization process using sequential particle filters

In case of poor consistency between the combinations, a sequential localization optimization using particle filters is performed. The transformation function  $t$  on map FG ( $MFG$ ) used to maximize the consistency between sources is defined as,

$$t(MFG, x, y, \theta) = R(\theta) * MFG + T(x, y) \quad (5.8)$$

Where  $R(\theta)$  is the 2D rotation matrix constructed using  $\theta$  and  $T$  is the 2D translation vector constructed using  $x$  and  $y$  translations.

In the sequential transformation optimization, consistency between map ( $MFG$ ) and each of the other sources ( $s_pFG$ ) is maximized in ego-frame along  $y$  direction (lateral) at first by iteratively distributing particles around the best match positions. The lateral offset estimation  $y^*$  and the final distribution of particles from this step is used for initializing the second particle filter which maximizes the match along the  $x$  (longitudinal) and  $\theta$  (heading) dimensions as illustrated with an example scenario in Fig. 5.9. The resulting optimized transformation  $(x^*, y^*, \theta^*)_{s_pFG}$  for each  $s_pFG$  is checked for their consistency by thresholding the distance between them. If they are not consistent, the consistency between all  $s_pFG$  is computed. Issue with map structure is identified if the consistency between other sources (sources which have fixed, known extrinsic transformations, eg. LiDAR-vision) is good but still the transformation optimization of these sources could not produce consistent localizations (within  $2\sigma$  uncertainty bounds). If the transformation estimates for each sensor combination are consistent, the estimation which gives the best consistency is chosen and integrity assessment is carried out. This estimation is also used to update the transformation estimation for the next time step.

### 5.2.6 Calculation of Protection Levels

To evaluate and compare the proposed integrity framework to the integrity concepts transposed from civil aviation concepts, Horizontal Protection Level ( $HPL$ ) is computed. According to Zhu et al. 2018, HPL is the radius of a circle in the horizontal plane which describes the region assured to contain the indicated horizontal position. It is the statistical bound for horizontal position error with a confidence level derived from the integrity risk requirement of an application. We also compute Lateral Protection Level ( $LatPL$ ) and Longitudinal Protection Level ( $LonPL$ ) as proposed in Reid et al.

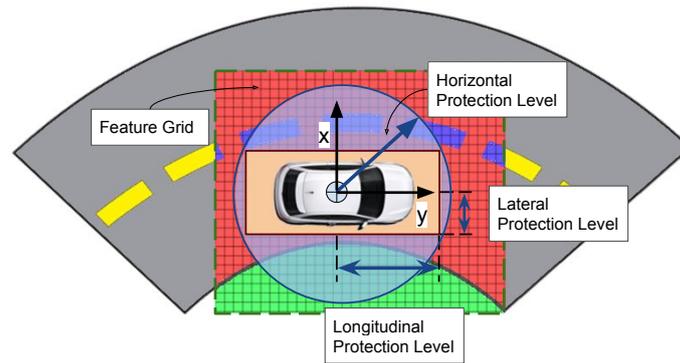


Figure 5.10: Illustration of Protection Levels for localization of ground vehicles.

2019. The illustration given in Fig. 5.10 shows the geometrical interpretations of these Protection levels with respect to the ego-frame of the vehicle and Feature Grids.

Extending these concepts, we use the final distribution of the particles from the transformation optimization particle filter described in Sec. 5.2.5, to compute  $LatPL$ ,  $LonPL$  and  $HPL$ . The support for this approach comes from the PL computation using RAIM presented in Sec. 1.3.2. According to this classical PL computation method, the confidence of positioning is multiplied by an inflation factor determined by the missed detection probability. Since localization estimation of the vehicle is not in the scope of the work, a direct adoption of this concept will be of less significance. However, we can define protection levels for information provided by the data sources based on the confidence on our consistency estimation method.

Since the key requirement for our integrity estimation framework is the transformation of all sources to a common frame, the confidence of consistency estimation depends on the accuracy of this transformation. In the proposed method, the accuracy of transformation (map to ego-frame) is determined by the particle filter based transformation solution obtained using transformation optimization step. Since the final transformation for the map data is determined by the particle distributions obtained from its matching operation with other data sources (LiDAR and Camera), the properties of these distributions are combined to obtain confidence of our consistency estimation method. Lateral and longitudinal positions of all the final particles from each optimization result (obtained from map-LiDAR and map-vision combinations) are modeled using Gaussian distribution.  $LatPL$ ,  $LonPL$  and  $HPL$  are then computed using the average variances along  $X$  and  $Y$  of particle distributions from each sensor combinations used to optimize transformation

as:

$$LatPL = K_Y \sqrt{(\sigma_{CY}^2 + \sigma_{LY}^2) / 2} \quad (5.9)$$

$$LonPL = K_X \sqrt{(\sigma_{CX}^2 + \sigma_{LX}^2) / 2} \quad (5.10)$$

$$HPL = K_H \sqrt{(\sigma_{CX}^2 + \sigma_{CY}^2 + \sigma_{LX}^2 + \sigma_{LY}^2) / 4} \quad (5.11)$$

where  $\sigma_{CX}^2$  and  $\sigma_{CY}^2$  are the lateral and longitudinal variance of particles from vision-map optimization result and  $\sigma_{LX}^2$  and  $\sigma_{LY}^2$  are the lateral and longitudinal variance of particles from LiDAR-map optimization result.  $K_X$ ,  $K_Y$  and  $K_H$  are the inflation parameters for the confidence of consistency estimation lateral, longitudinal and horizontal dimensions respectively. The proposed value for these parameters in our work is 2, since  $2\sigma$  bounds ensure the reliability of 95 percent of the transformations solutions.

### 5.3 Experiments and Discussions

Experiments are conducted with scenarios available in KITTI benchmark suite (Geiger et al. 2013) to establish proof-of-concept. RTK GPS fixes in these datasets are added with noise generated (standard deviation  $\sigma = 61$  cm) using the GPS-noise simulation model proposed by Backman et al. (2010) to simulate poor GPS localization fixes. Outliers which are higher than  $2\sigma$  of the GPS-noise simulation model, are used to replace RTK GPS fixes at random sections of the trajectory. Finally, 5% of the RTK GPS fixes are removed from the trajectory randomly to emulate GPS outages as they may occur in generic GPS receivers. Since different data sources have different spatial ranges, a 3D region of interest (ROI) in vehicle's ego-frame is established. 25 m in front of the vehicle (positive X axis), 15 meters behind (negative X axis) and 15 m at each side (Y axis) are chosen as its limits in XY plane. Since vision can't provide data in the back of the vehicle as well as until the front bumper of the vehicle, the ROI of vision is limited from 3.5 m to 25 m along positive X axis. Even though vision data used in this work does not cover the back view of the vehicle, the other two major sources - LiDAR and Map - can

provide information in the back of vehicle, hence justifying the choice of limit in negative X axis. The FGs illustrated in Fig. 5.7 are generated according to these ROI definitions.

The discussion on the result has three parts. Firstly, comparing the performance of the FG-based method to the polynomial-based method in Chapter 4. This includes comparison of integrity markers in the datasets presented in Chapter 4 and showcasing the improvements provided by the new method in handling Fault and Feasibility Predictors (*FP*) produced by the previous method in Sec. 4.3.3. *F**P*s are the markers generated when the fitting of common model to the data sources is not possible or feasible. These markers suggest the limitations of the polynomial model, which mainly arise when the method is applied on non-highway scenarios. The set of five *FP* markers defined are:

- *FP<sub>m</sub>*: Not enough nodes in the map for model fitting.
- *FP<sub>v</sub>*: Not enough lane markings for model fitting.
- *FP<sub>g</sub>*: GPS measurement is not available or an outlier.
- *FP<sub>s</sub>*: Vehicle not moving or moving very slow.
- *FP<sub>t</sub>*: Vehicle performing a hard turn

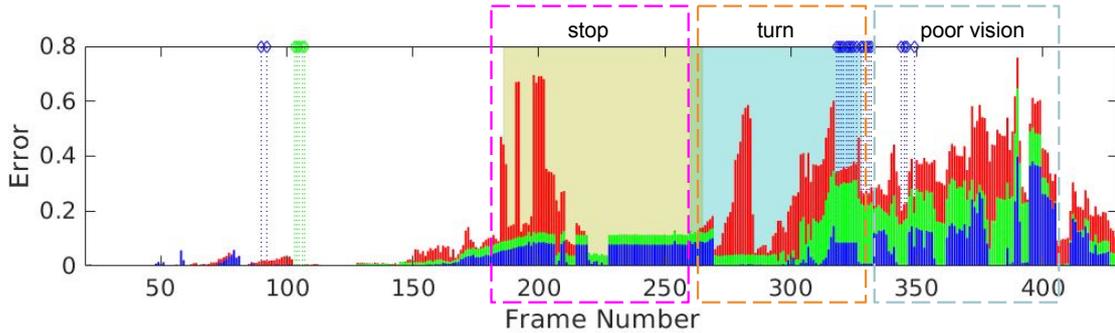
The second part of this discussion considers more datasets in semi-urban and urban scenarios to evaluate the integrity estimation of sources in complex situations such as junctions, road-splits and merges, etc. In the final part, we compute protection levels based on consistency of data sources estimated from our framework and compare them with values presented in Reid et al. 2019.

### 5.3.1 Integrity Markers comparison

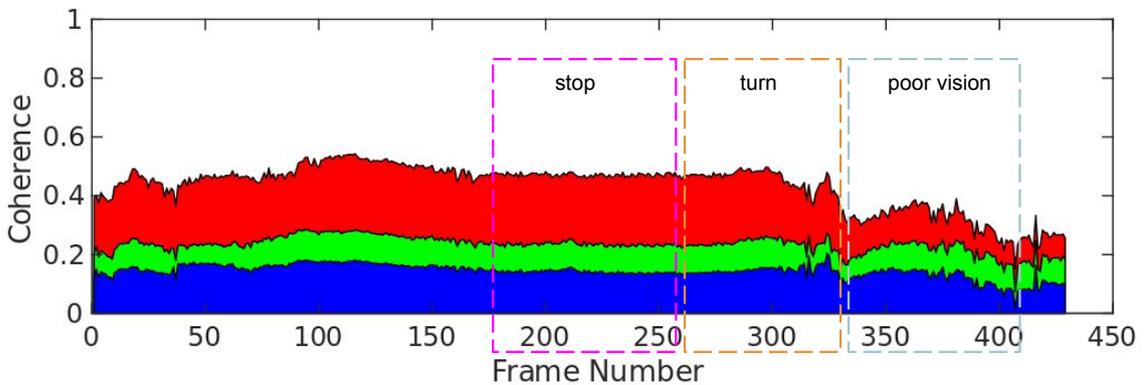
The key difference between polynomial-based method and FG-based method is the parameter they use for the integrity computation. The former uses the error observed in model fitting to evaluate integrity, whereas the latter uses consistency between data representations to achieve the same. Hence, the contribution of error by each sensor and the contribution of consistency by each sensor are used for this analysis of results of these methods respectively. Note that, these parameters are complimentary in nature, ie. when the error is high, consistency will be low and vice versa. This is evident in the section

between frame number 1-170 in Fig. 5.11. The same errors are introduced in the GPS for each algorithm and the results obtained from the dataset 2011\_09\_26\_drive\_0029 are shown in Fig. 5.11.

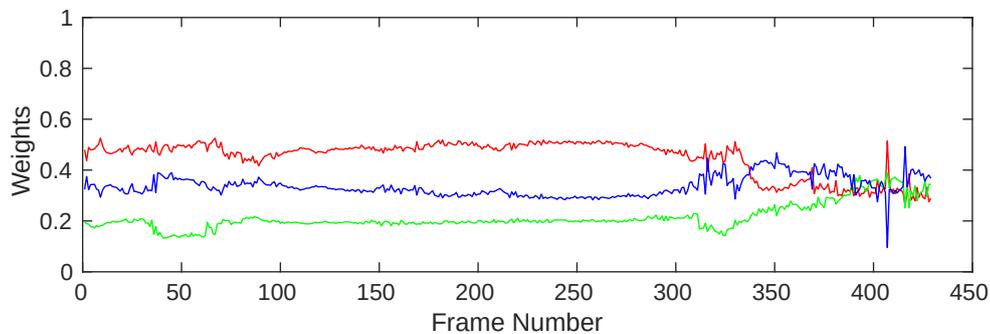
The primary advantage of the FG-based method is the ability to evaluate the integrity during the conditions where  $FPs$  are produced due to the limitations of polynomial-based integrity analysis employed in Chapter 4.



(a) Error from each sensor used in dataset 2011\_09\_26\_drive\_0029 using previous method. Red: GPS error, Green: Map error, Blue: Vision error, Green dotted lines:  $FP_m$ , Blue dotted lines:  $FP_v$ , Light brown:  $FP_s$ , Light blue:  $FP_t$



(b) Consistency observed for each sensor used in dataset 2011\_09\_26\_drive\_0029 using new method. Red: LiDAR, Green: Map, Blue: Vision



(c) Integrity Markers for dataset 2011\_09\_26\_drive\_0029 using new method. Red: LiDAR, Green: Map, Blue: Vision

Figure 5.11: Comparison Results of dataset 2011\_09\_26\_drive\_0029

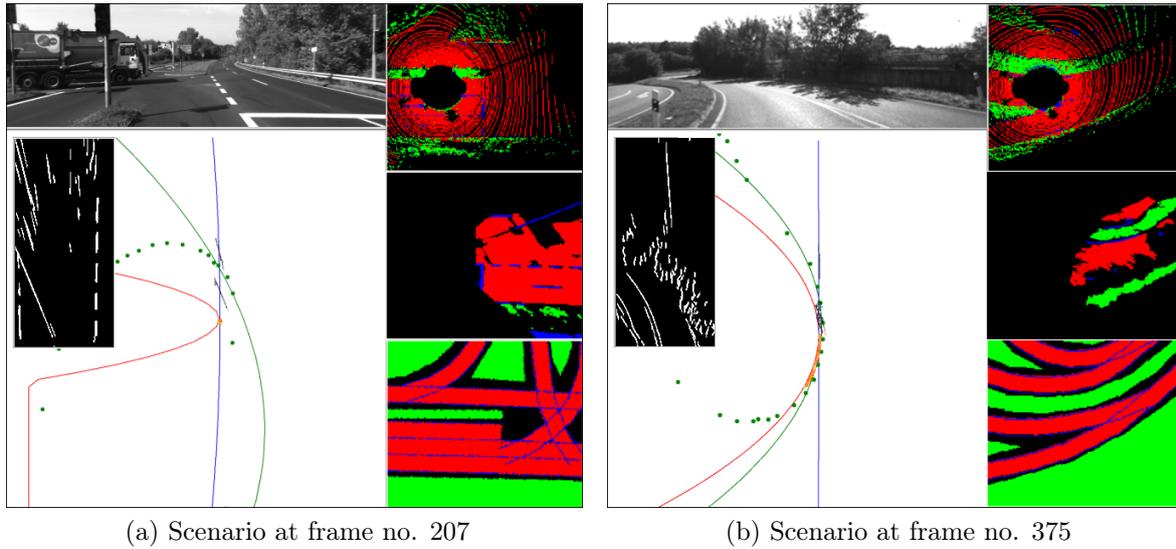
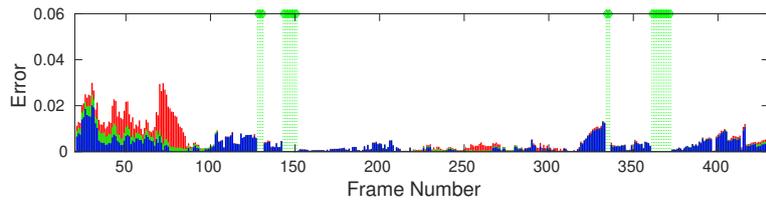


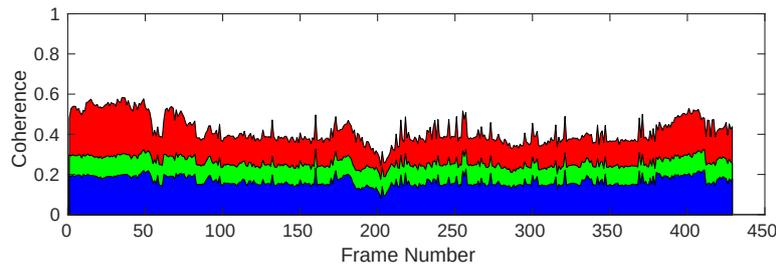
Figure 5.12: Specific scenarios from dataset 2011\_09\_26\_drive\_0029. Left-top: view of the scenario, Left-bottom: model fitting, Left-inset: lane marking detections, Right-top: FG of LiDAR, Right-middle: FG of vision, Right-bottom: FG of map

The stopping of vehicle between frame numbers 187 and 265 and a hard left turn at the junction from 265 to 330 cause poor model extraction using the polynomial-based method resulting in unusable integrity evaluation. High consistency is observed during the same scenario as shown in Fig. 5.11b using the FG-based method providing meaningful integrity estimation. Fig. 5.12a shows an example frame (207) during this section where polynomial model estimation fails to represent data from sources. On the other hand, the FGs are able to represent the scenario well. After frame 330, the vehicle enters a curved link road with challenging light conditions such as shadows and oversaturated road sections as shown in Fig. 5.12b, causing large model fitting errors in vision shown in Fig. 5.11a. Though a decrease in the consistency is observed, the addition of LiDAR and introduction of new features help the FG-based method to provide reliable integrity markers.

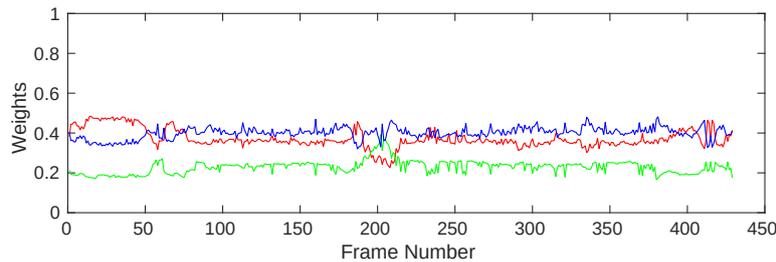
In Fig. 5.13, results of integrity assessment in a highway scenario are presented, where the polynomial-based method reliably performed. The  $FP_m$  instances observed in this dataset are due to the lack of map nodes to reliably fit the polynomial model in straight line road sections. In the FG-based method, the model fitting is replaced with FG data representation, which tackles such errors in modeling. Comparison of integrity markers in specific cases presented in Chapter 4 with the integrity markers provided by FG-based



(a) Error from each sensor used in dataset 2011\_09\_26\_drive\_0028 using previous method. Red: GPS error, Green: Map error, Blue: Vision error, Green dotted lines:  $FP_m$ , Blue dotted lines:  $FP_v$ , Light brown:  $FP_s$ , Light blue:  $FP_t$



(b) Consistency observed for each sensor used in dataset 2011\_09\_26\_drive\_0028 using new method. Red: LiDAR, Green: Map, Blue: Vision



(c) Integrity Markers for dataset 2011\_09\_26\_drive\_0028 using new method. Red: LiDAR, Green: Map, Blue: Vision

Figure 5.13: Comparison results of dataset 2011\_09\_26\_drive\_0028

method is given in Table 5.1.

A general tendency of improved integrity values is observed across all datasets and scenarios. For example, in second row of Table 5.1, integrity weight of vision computed using polynomial method was lower due to the improper detection of curved lane markings as straight lane markings. This resulted in an inconsistent polynomial model compared to other two data sources, causing a low integrity weight of 0.175. But using the new method, drivable road detection along with surrounding structure detection improved the consistency of vision data with other sources, resulting in a higher integrity value of 0.612. The proposed method is proven to be able to handle every situation where  $FP$  was provided by the old method. In first row of Table 5.1, lack of enough map

Dataset-Frames	Integrity (Polynomial model)	Integrity (Feature Grid)	Situation
Dataset 1-150	$FP_m$	map- 0.422	not enough nodes from the map
Dataset 1-21	vision- 0.175	vision- 0.612	no good quality lane markings
Dataset 2-390	map- 0.087	map- 0.374	road with multiple curvatures
Dataset 3-562	$FP_v$	vision- 0.573	partial occlusion in vision due to vehicles
Dataset 3-1117	map- 0.006	map- 0.381	wrong map extraction
Dataset 4-22	vision- 0.214	vision- 0.681	road with multiple curvatures
Dataset 4-260	vision- 0.651	vision- 0.629	highway road with single curvature

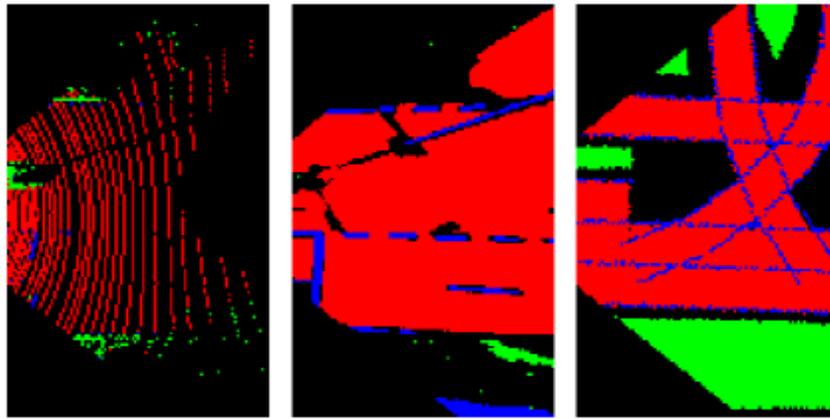
Table 5.1: Comparison of results obtained using FG-based method and polynomial-based method

nodes on a straight road segment made model-based integrity estimation not possible as confirmed by the  $FP_m$  flag. The new approach enables integrity estimation and provides an integrity weight of 0.422. It is worth noting that a high integrity value is not observed because of poor map rendering due to lack of correct lane width information from the map.

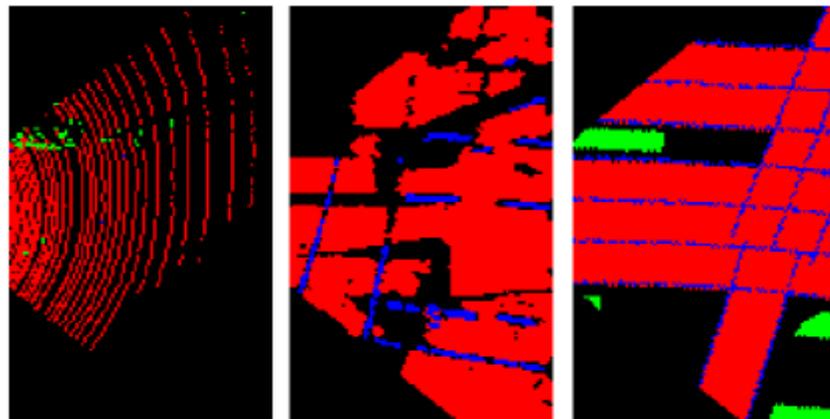
### 5.3.2 Complex Situations

This section is dedicated to analyze the behavior of integrity assessment system in some of the selected complex scenarios present in the KITTI dataset. In Fig. 5.14a, an example of a semi-urban road junction is shown. Due to the lack of information from the map, the rendering process failed to reconstruct the continuity of lanes at the intersections.

On the other hand, vision and LiDAR data detected all the branch roads at the junction and managed to perceive the width of each of these road sections accurately. This results in a lower integrity value for map at this junction (Frame numbers: 310 - 320) compared to other sources as shown in Fig. 5.11c. One of the main reasoning behind the proposed data representation is the fact that it is an improvement over other existing geometrical models for intersections which fail to accommodate partially correct data. Fig. 5.14b shows a partial road detection from LiDAR due to the difference in elevation



(a) Multi lane junction from dataset 2011\_09\_26\_drive\_0029.



(b) Partially consistent data from a junction from dataset 2011\_09\_26\_drive\_0011.

Figure 5.14: Examples of complex scenarios - cells with road labels(red), lane marking labels(blue), other surfaces labels(green), unclassified labels(black)

of one of the road branches in the scenario. Even though data available from LiDAR is not complete, the part that is detected is coherent with both vision and map. In fact, LiDAR has more integrity than vision in comparison not only thanks to its consistency in road detections, but also the available grass-patch detection compensates the partial road detection. The integrity values at this scenario (Frame numbers: 120-200, dataset 2011\_09\_26\_drive\_0011) are computed around 0.456, 0.349 and 0.165 for LiDAR, vision and map respectively. Map data has low integrity in this situation due to the errors in rendering the junction.

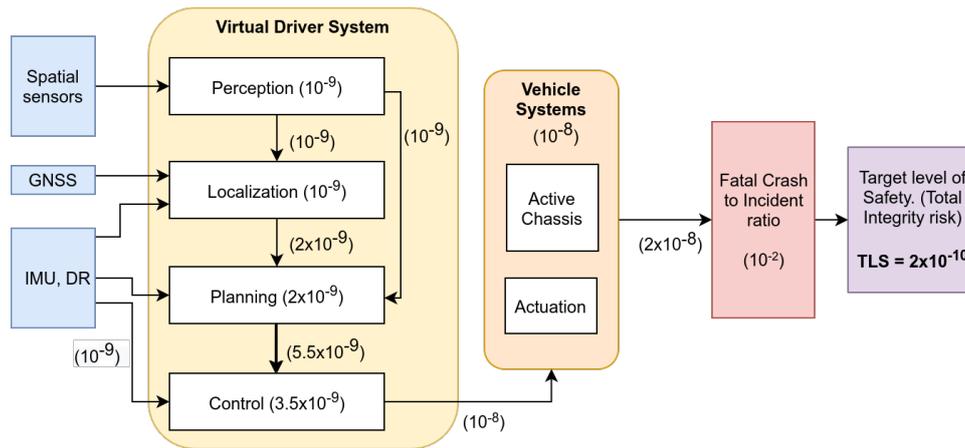


Figure 5.15: Integrity risk allocation from Reid et al. (2019)

### 5.3.3 Performance of Integrity Monitoring

As the main work used in this comparison, the integrity quantifications by Reid et al. 2019 should be recalled from Sec. 1.3.4 and explained here. To recall, Integrity Risk (IR) is the probability of providing a localization estimate that is out of Alert Limits (AL) without warning the user within the Time-To-Alert and it is typically represented as number of such occurrences per-hour. Based on the statistics of fatal road accidents, Reid et al. 2019 defines the total integrity risk (Target Level Safety, TLS) of a navigation system as  $2 \times 10^{-10}$  fatal crashes per vehicle mile. To arrive at the risk allocation for the navigation system as shown in Fig. 5.15, they back-propagate the integrity risks from TLS to the subsystems. Considering every integrity risk event does not result in fatal crashes (only 1 in 100), the vehicle navigation system can have  $2 \times 10^{-8}$  integrity risk events per vehicle mile. This integrity risk allocation is equally divided ( $10^{-8}$ ) to the two main parts of an intelligent vehicle: Virtual driver system (intelligence) and vehicle systems (hardware). Within the virtual driver system, localization block should have low failure probability as its performance directly propagates downwards to planning and hence control of the total system. Perception and Localization blocks get a tenth of the total integrity risk ( $10^{-9}$ ) of the virtual driver system and the rest is distributed for planning and control. The computation of ALs are done by considering the size of different classes of vehicles and road properties such as lane width, number of lanes, curvatures etc. Since the allocation of the integrity risks is available for each block, the ALs computed for the whole system by Reid et al. 2019 can be extrapolated to ALs of its individual block. The vehicle in

the datasets used in this work is a full-size car, we can take the corresponding total AL values,  $AL_t$  from Table 1.1 and calculate the AL for perception block  $AL_p$  as,

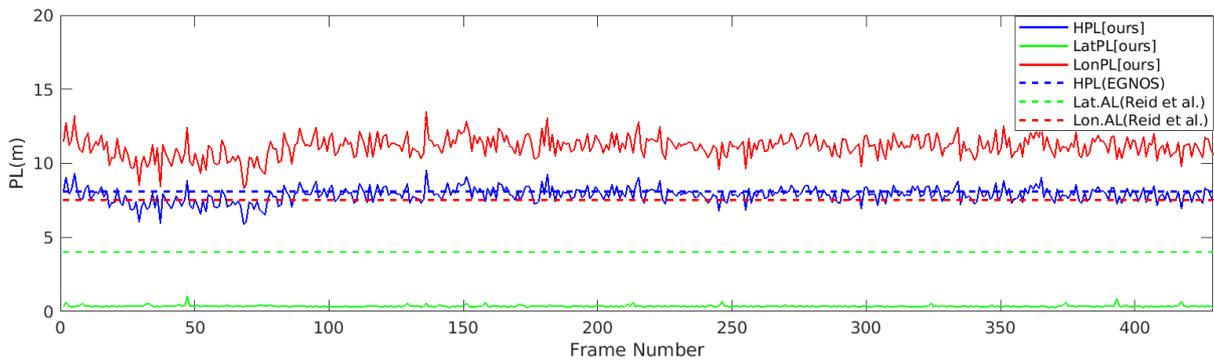
$$AL_p = AL_t \frac{IR_t}{IR_p} \quad (5.12)$$

where  $IR_t$  and  $IR_p$  are the integrity risk allocation for the total system and perception block respectively. Using Eq. 5.12, the different ALs for perception block is calculated as,

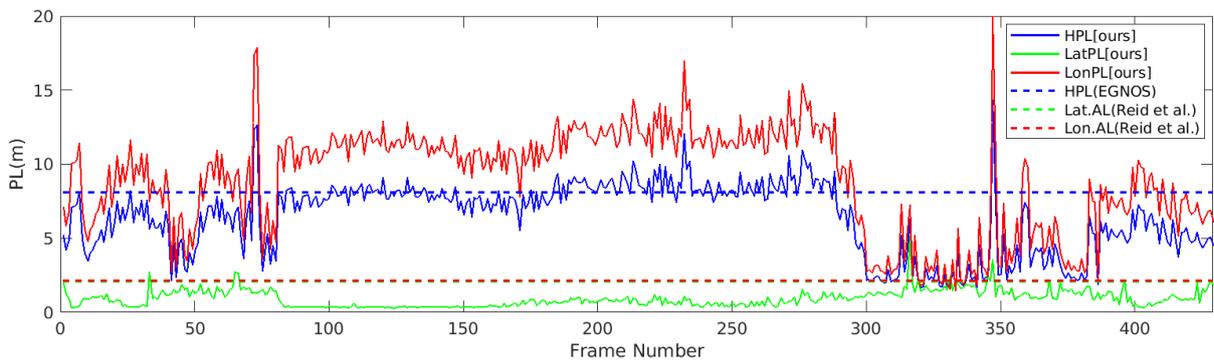
	Urban Scenarios	Highway scenarios
Lat. AL	2.1 m	4.0 m
Lon. AL	2.1 m	7.5 m

European Global Navigation Satellite System Agency (EGNOS) provides the historical performance of GNSS systems in terms of HPL and VPL in their website ([www.egnoss-user-support.essp-sas.eu](http://www.egnoss-user-support.essp-sas.eu)). Using this information, the average value of  $HPL$  over last 5 years (from 01-2015 to 07-2020) for the nearest zone (Zurich, distance 178 km) to the dataset location (Karlsruhe) is calculated as 8.1 m. In the rest of this section we present integrity analysis of several datasets and their comparison with these quantifications of PL and ALs. The results obtained from PL evaluation of two of the datasets presented in Sec. 5.3.1 are shown in Fig. 5.16

In highway scenarios, the  $LatPL$  computed using our method is completely within the  $Lat.AL$  limit derived by Reid et al. 2019 whereas in urban scenario, most of the time  $LatPL$  from our method is under the limit. On the other hand, HPL computed using our method shows good consistency with the historical  $HPL$  calculated using EGNOS historical data. However, the  $LonPL$  computations are outside the limit of  $Lon.AL$  derived by Reid et al. 2019 most of the time in both the scenarios. This is due to the fact that the sensors considered in this work are better at providing lateral information (Reid et al. 2019) than longitudinal information. This is evident from highway scenario in Fig. 5.16a, where the road is straight without any other significant information to bound the sensor data in longitudinal direction. In Fig. 5.16b, sections where the  $LonPL$  computed from our method is closer to the  $Lon.AL$  of 2.1 m, contain curved road sections or distinguishable other surfaces as shown in Fig. 5.12b, which helps to reduce  $LonPL$  considerably. Hence, the results presented in this section demonstrate the capability of



(a) HPL evaluation result of dataset 2011\_09\_26\_drive\_0028 (highway scenario).



(b) HPL evaluation result of dataset 2011\_09\_26\_drive\_0029 (urban scenario).

Figure 5.16: Protection Levels (PL) comparison

the proposed method to assess integrity of perception sensors in localizing vehicles with required accuracy for urban and highway navigation.

Recalling the integrity events presented in Section 1.3.1, when PL is greater than the AL, a localization system is deemed unavailable. However, when it comes to information from perception systems, a greater PL than AL signifies the lack of quality of information. A meta-analysis of PL computation from several datasets are given in Table 5.2. It is observed that our *LatPL* computation are under the *Lat.AL* requirement in majority of situations in the considered datasets.

## 5.4 Conclusion

This work presents a framework for integrity monitoring of sources used in localization of autonomous vehicles. The limitations of common geometrical models in representing multi-modal data sources are identified in this work. To overcome these issues, a semantic Feature Grid model is proposed, that can geometrically represent different features using labels. A function for consistency evaluation between Feature Grids is formalized to

KITTI Datasets (2011_09_)	No. of Frames	Actual Scenario	Availability (in % of total no. of frames)				
			Highway		Urban		All
			Lat.	Lon.	Lat.	Lon.	Hori.
26_drive_0005	154	urban	98.04	58.82	87.58	18.30	94.77
26_drive_0011	233	urban	100	25	100	0	87.07
26_drive_0015	297	highway	100	0	100	0	62.16
26_drive_0017	114	junction	100	65.49	96.46	0	100
26_drive_0027	187	highway	100	0	100	0	63.10
26_drive_0028	429	highway	100	0	100	0	67.83
26_drive_0029	429	urban	99.77	32.87	97.20	1.86	74.13
26_drive_0032	389	highway	100	5.66	99.48	0	71.21
26_drive_0060	77	junction	100	6.49	98.70	0	83.12
26_drive_0070	419	highway	100	31.03	98.09	7.63	83.77
26_drive_0101	935	mixed	100	19.57	99.57	0	79.89
28_drive_0002	375	urban	100	77.87	99.73	0	92.27
29_drive_0026	157	urban	100	84.08	100	0	100

Table 5.2: Meta-Analysis of availability of data from different datasets according to urban and highway alert limits defined by Reid et al. 2019

iteratively optimize the localization as well as to assess the integrity of data sources. The framework is tested using different scenarios from datasets and the results show the versatility of the proposed model, which is able to provide reliable and consistent integrity estimation in highway as well as semi-urban and urban environments. This method is proven robust against inconsistencies in feature detections such as partial detections, occlusions and poor map rendering. The method presented claims scalability, since it can be implemented with any number of sensors and digital map sources. The only requirement for the applicability of this framework is the ability to detect common features from all the data sources and represent them geometrically in the proposed Feature Grid representations. This work also illustrates how classical integrity markers like Protection Levels can be transposed for perception data sources used in autonomous vehicles.



---

# Chapter 6

## Conclusions and Perspectives

### Conclusions

This thesis is aimed to address the integrity monitoring problem of data sources used in multimodal localization of intelligent vehicles. The necessity of this task is clearly identified and formalized through extensive research on the existing approaches in this field. The integrity monitoring of localization using GNSS positioning is found out to be a well-developed. However, a lack of integrity monitoring considerations is observed for other perception data sources such as vision, LiDAR, radar, digital maps etc., that are used in modern localization solutions capable of handling complex situations where GNSS localization is limited. Hence, our research mainly focused on proposing an integrity monitoring framework for the data produced at the sources-level of localization systems.

To this end, our first contribution is the formulation of an integrity protocol by transposing the data integrity concepts from the field of information systems to the data sources used in intelligent vehicles. We identified the core attributes of data integrity and their enablers considered for information systems and established the relevant attributes and enablers for integrity of data sources according to the existing requirements in intelligent vehicle localization process. Based on this protocol, an integrity monitoring framework was developed, which is capable of providing integrity estimation for different perception data sources simultaneously. The framework was developed based on three key factors:

1. All the considered data sources can provide information about at least one common

feature from the environment around the vehicle.

2. A common frame of reference is always identifiable and the transformation of all data from sources to this frame is computable.
3. There exists a common model or representation for all data sources in the common reference frame.

In our work, we have shown that all three factors are achievable for most of the commonly used perception data sources. Two methods have been proposed based on this framework, which use different sensor combinations and strategies for integrity estimation.

The first method is based on a quadratic polynomial model (QPM), and it considered integrity monitoring of three sources - GPS, digital map and vision - in highway or semi-urban scenarios. GPS trajectory of the vehicle, nodes from the map and lane marking detections from vision were represented in the ego-frame of the vehicle and the data from them were modeled using QPM. The transformation of data from map and GPS to ego-frame is achieved by iteratively maximizing the consistency of their models to the model from vision. The proposed integrity analysis checked the consistency of their models using a cross-consistency estimation technique. This analysis generated integrity weights (IW) for each source under consideration. We have also proposed alerts called Fault and Feasibility predictors (FP), based on the availability and accuracy of integrity monitoring method under certain critical situations.

In the second method, the scope of the work was extended to complex urban scenarios and LiDAR was added to the data sources. We used three categories of semantic features detected from each source in this method - markings on the road, drivable road regions and vegetation/grass-patches in the surroundings. Information from these detections were represented using a proposed spatial semantic grid representation called feature grids (FG). The ego-frame of the vehicle was used for these data representations, and a particle filter based transformation optimization algorithm was proposed to estimate and ensure reliable transformation of map data into ego-frame. Similar to the polynomial-based method, these FG representations are analyzed for their consistency to estimate integrity weights for each sources. Using the confidence of the process of transformation optimization, we proposed a method to calculate protection levels for the information provided by the data sources.

The experiments on both of these methods have been performed using publicly available datasets containing various driving scenarios. The analysis of the performance of the polynomial-based method was done by manually introducing errors of known properties to each data source and verifying the effect of errors in estimated integrity weights. For the FG-based method, the calculated protection levels from our proposed method is compared to quantified integrity requirements available in the state-of-the art methods.

## Perspectives and Future work

The research presented in this work explored different aspects of a challenging yet crucial problem of integrity monitoring of data sources and proposed novel framework to address this issue. The experiments performed and the evaluation of results provided us with some key perspectives on this particular problem and clues for further research.

The main perspective drawn from this work is the need of analysis and quantification of the effect of integrity monitoring of data sources in localization and sensor-fusion applications. Currently, the proposed framework only focuses on identifying and inferring the problems in the data sources, so that any inconsistent results obtained from a system that processed these data sources can be explained. However, in order to take full advantage of the integrity of data sources, the successive processes have to be modified or developed to be able to incorporate integrity weights in their input and use them. Hence, one important future work in this area is to propose a scheme that allows effortless integration of data source integrity weights to commonly used localization and fusion schemes such as Kalman filters, particle filters, etc. This will allow us to quantify the improvement provided by the proposed data source integrity solution.

Another perspective obtained from the experiments conducted on FG-based method is the importance of accurate map information from GISs. The rule-based map rendering technique used in this method is observed to be contributing several inconsistencies, which makes it difficult to isolate map rendering errors from GPS positioning errors. These errors have to be mitigated by using advanced map-rendering rules or the usage of High Definition maps which are enriched with globally localized lane-level information. The lack of accurate map information can be a bottleneck in integrity monitoring process when attempting to expand the set of features, to accommodate complex driving scenarios.

Finally, more experiments have to be conducted using this integrity monitoring framework with custom-made datasets as publicly available datasets are limited in the scenarios they contain. For example, we could not find any datasets generated in urban canyons or long underground tunnels. Likewise, specific highway scenarios, such as roundabouts of different sizes, roads with hairpin curves are not available in publicly available datasets. The examination of integrity of sources in such challenging scenarios would be of high importance and will definitely contribute to the usability of intelligent vehicles in general.

---

# Bibliography

- Ahmad, Khairol Amali Bin, Mohamed Sahnoudi, and Christophe Macabiau (2014). “Characterization of GNSS receiver position errors for user integrity monitoring in urban environments”. In: *European Navigation Conference-GNSS*, pp. 1152–1157.
- Armesto, Leopoldo, Josep Tornero, and Markus Vincze (2007). “Fast ego-motion estimation with multi-rate fusion of inertial and vision”. In: *The International Journal of Robotics Research* 26.6, pp. 577–589.
- Backman, J. et al. (2010). “Simulation Environment for Testing Guidance Algorithms with Realistic GPS Noise Model”. In: *IFAC Proceedings Volumes* 43.26, pp. 139–144.
- Bailey, James E. and Sammy W. Pearson (1983). “Development of a Tool for Measuring and Analyzing Computer User Satisfaction”. In: *Management Science* 29.5, pp. 530–545.
- Ballardini, A. L., D. Cattaneo, and D. G. Sorrenti (2019). “Visual Localization at Intersections with Digital Maps”. In: *2019 International Conference on Robotics and Automation*. Montreal, Canada, pp. 6651–6657.
- Ballardini, A. L. et al. (2016). “Leveraging the OSM building data to enhance the localization of an urban vehicle”. In: *2016 IEEE 19th International Conference on Intelligent Transportation Systems*. Rio de Janeiro, Brazil, pp. 622–628.
- Ballatore, Andrea, Michela Bertolotto, and David C Wilson (2013). “Geographic knowledge extraction and semantic similarity in OpenStreetMap”. In: *Knowledge and Information Systems* 37.1, pp. 61–81.
- Basnayake, Chaminda et al. (2010). “Can GNSS Drive V2X?” In: *GPS World* 21.10, pp. 35–43.
- Bentley, J. and H. Maurer (1980). “Efficient worst-case data structures for range searching”. In: *Acta Informatica* 13, pp. 155–168.

- Bloesch, Michael et al. (2015). “Robust visual inertial odometry using a direct EKF-based approach”. In: *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on*. IEEE, pp. 298–304.
- Boritz, J Efrim (2004). *Managing enterprise information integrity: security, control, and audit issues*. Isaca.
- Boritz, J. Efrim (2005). “IS practitioners’ views on core concepts of information integrity”. In: *International Journal of Accounting Information Systems* 6.4, pp. 260 –279. ISSN: 1467-0895.
- Bovee, Matthew, Rajendra P. Srivastava, and Brenda Mak (2003). “A conceptual framework and belief-function approach to assessing overall information quality”. In: *International Journal of Intelligent Systems* 18.1, pp. 51–74.
- Braga, Adriana and Robert K. Logan (2018). “Encyclopedia of Information Science and Technology”. In: IGI Global. Chap. Communication, Information, and Pragmatics, pp. 1238–1247.
- Brown, R Grover (1992). “A baseline GPS RAIM scheme and a note on the equivalence of three RAIM methods”. In: *Navigation* 39.3, pp. 301–316.
- Cadena, Cesar et al. (2016). “Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age”. In: *IEEE Transactions on Robotics* 32.6, pp. 1309–1332.
- Callmer, Jonas et al. (2011). “Radar SLAM using visual features”. In: *EURASIP Journal on Advances in Signal Processing* 2011.1, pp. 1–11.
- Cao, Yulong et al. (2019). “Adversarial sensor attack on lidar-based perception in autonomous driving”. In: *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, pp. 2267–2281.
- Capellier, Edouard et al. (2018). “Evidential grid mapping, from asynchronous LIDAR scans and RGB images, for autonomous driving”. In: *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*. IEEE, pp. 2595–2602.
- Cen, Sarah H and Paul Newman (2018). “Precise ego-motion estimation with millimeter-wave radar under diverse and challenging conditions”. In: *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, pp. 1–8.

- Chao, Pingfu et al. (2020). “A Survey on Map-Matching Algorithms”. In: *Databases Theory and Applications*. Ed. by Renata Borovica-Gajic, Jianzhong Qi, and Weiqing Wang. Cham: Springer International Publishing, pp. 121–133.
- CobiT, Cobit (2002). “Control Objectives for Information and related Technology”. In: *IT Governance Institute www.isaca.org*.
- Comport, Andrew I, Ezio Malis, and Patrick Rives (2007). “Accurate quadrifocal tracking for robust 3d visual odometry”. In: *Robotics and Automation, 2007 IEEE International Conference on*. IEEE, pp. 40–45.
- Concha, Alejo and Javier Civera (2014). “Using superpixels in monocular SLAM”. In: *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, pp. 365–372.
- Corke, Peter, Dennis Strelow, and Sanjiv Singh (2004). “Omnidirectional visual odometry for a planetary rover”. In: *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*. Vol. 4. IEEE, pp. 4007–4012.
- Cosmen-Schortmann, Joaquín et al. (2008). “Integrity in urban and road environments and its use in liability critical applications”. In: *2008 IEEE/ION Position, Location and Navigation Symposium*. IEEE, pp. 972–983.
- Davidson, Pavel, Jussi Collin, and Jarmo Takala (Oct. 2011). “Application of particle filters to map-matching algorithm”. In: *17th Saint Petersburg International Conference on Integrated Navigation Systems, ICINS 2010 - Proceedings 2*, pp. 285–292.
- DeCleene, Bruce (2000). “Defining pseudorange integrity-overbounding”. In: *Proceedings of the 13th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GPS 2000)*, pp. 1916–1924.
- Diel, David D, Paul DeBitetto, and Seth Teller (2005). “Epipolar constraints for vision-aided inertial navigation”. In: *Application of Computer Vision, 2005. WACV/MOTIONS’05 Volume 1. Seventh IEEE Workshops on*. Vol. 2. IEEE, pp. 221–228.
- Ekstrom, Arne D and Eve A Isham (2017). “Human spatial navigation: representations across dimensions and scales”. In: *Current Opinion in Behavioral Sciences* 17. Memory in time and space, pp. 84 –89. ISSN: 2352-1546.
- Engel, Jakob, Thomas Schöps, and Daniel Cremers (2014). “LSD-SLAM: Large-scale direct monocular SLAM”. In: *European Conference on Computer Vision*. Springer, pp. 834–849.

- Eskandarian, Azim (2012). *Handbook of intelligent vehicles*. Vol. 2. Springer.
- Ester, Martin et al. (1996). “A density-based algorithm for discovering clusters in large spatial databases with noise.” In: *Kdd*. Vol. 96. 34, pp. 226–231.
- Feng, R. et al. (2017). “Registration of multitemporal GF-1 remote sensing images with weighting perspective transformation model”. In: *2017 IEEE International Conference on Image Processing*. Beijing, China, pp. 2264–2268.
- Filliat, David and Jean-Arcady Meyer (Jan. 2003). “Map-based navigation in mobile robots:: I. A review of localization strategies”. In: *Cognitive Systems Research*, pp. 243–282.
- Flowerday, S. and R. V. Solms (2007). “What constitutes information integrity”. In: *SA Journal of Information Management* 9, pp. 1–19.
- Flowerday, Stephen and Rossouw von Solms (2005). “Real-time information integrity=system integrity+data integrity+continuous assurances”. In: 24.8, pp. 604 –613. ISSN: 0167-4048.
- Forster, Christian, Matia Pizzoli, and Davide Scaramuzza (2014). “SVO: Fast semi-direct monocular visual odometry”. In: *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. IEEE, pp. 15–22.
- Forster, Christian et al. (2017). “On-Manifold Preintegration for Real-Time Visual-Inertial Odometry”. In: *IEEE Transactions on Robotics* 33.1, pp. 1–21.
- Geiger, A et al. (2013). “Vision meets robotics: The KITTI dataset”. In: *The International Journal of Robotics Research* 32.11, pp. 1231–1237.
- Geiger, Andreas, Julius Ziegler, and Christoph Stiller (2011). “Stereoscan: Dense 3d reconstruction in real-time”. In: *2011 IEEE intelligent vehicles symposium (IV)*. Ieee, pp. 963–968.
- Ghabcheloo, Reza and Shadman Siddiqui (2018). “Complete odometry estimation of a vehicle using single automotive radar and a gyroscope”. In: *2018 26th Mediterranean Conference on Control and Automation (MED)*. IEEE, pp. 855–860.
- Grisetti, G. et al. (2010). “A Tutorial on Graph-Based SLAM”. In: *IEEE Intelligent Transportation Systems Magazine* 2.4, pp. 31–43.
- Groves, Paul D. et al. (2013). “A Portfolio Approach to NLOS and Multipath Mitigation in Dense Urban Areas”. In:

- Guerrero, José Jesús, Ruben Martinez-Cantin, and Carlos Sagüés (2005). “Visual map-less navigation based on homographies”. In: *Journal of Field Robotics* 22.10, pp. 569–581.
- Horn, Joachim and Gunther Schmidt (1995). “Continuous localization of a mobile robot based on 3D-laser-range-data, predicted sensor images, and dead-reckoning”. In: *Robotics and Autonomous Systems* 14.2-3, pp. 99–118.
- Howard, Andrew (2008). “Real-time stereo visual odometry for autonomous ground vehicles”. In: *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*. IEEE, pp. 3946–3952.
- Hsu, Chih-Ming and Chung-Wei Shiu (2019). “3D LiDAR-Based Precision Vehicle Localization with Movable Region Constraints”. In: *Sensors* 19.4, p. 942.
- Hu, Jwu-Sheng and Ming-Yuan Chen (2014). “A sliding-window visual-imu odometer based on tri-focal tensor geometry”. In: *Robotics and automation (ICRA), 2014 IEEE international conference on*. IEEE, pp. 3963–3968.
- Hwang, Inseok et al. (2009). “A survey of fault detection, isolation, and reconfiguration methods”. In: *IEEE transactions on control systems technology* 18.3, pp. 636–653.
- ICAO, July (1973). “International standards and recommended practices”. In: *Aeronautical Telecommunications, Annex 10 to the Convention of International Civil Aviation*. Vol. 4.
- International, SAE (2018). *Taxonomy and Definitions for Terms Related to Driving Automation Systems for On-Road Motor Vehicles*.
- Irani, M. and P. Anandan (2000). “About Direct Methods”. In: *Vision Algorithms: Theory and Practice: International Workshop on Vision Algorithms Corfu, Greece, September 21–22, 1999 Proceedings*. Ed. by Bill Triggs, Andrew Zisserman, and Richard Szeliski. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 267–277. ISBN: 978-3-540-44480-0.
- Isermann, Rolf (1997). “Supervision, fault-detection and fault-diagnosis methods -an introduction”. In: *Control engineering practice* 5.5, pp. 639–652.
- ISO, ISO (2011). “26262-1: 2011 Road Vehicles–Functional Safety–Part 1: Vocabulary”. In: *Berlin: ISO*.
- Jabbour, M., Philippe Bonnifait, and Véronique Cherfaoui (Oct. 2006). “Enhanced Local Maps in a GIS for a Precise Localisation in Urban Areas”. In: pp. 468–473.
- Jazwinski, Andrew H (2007). *Stochastic processes and filtering theory*. Courier Corporation.

- Jeon, Soo Jung et al. (2015). “GPS waypoint fitting and tracking using model predictive control”. In: *2015 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, pp. 298–303.
- Jones, Eagle S and Stefano Soatto (2011). “Visual-inertial navigation, mapping and localization: A scalable real-time causal approach”. In: *The International Journal of Robotics Research* 30.4, pp. 407–430.
- Kang, Jeong Min et al. (2020). “Lane-Level Map-Matching Method for Vehicle Localization Using GPS and Camera on a High-Definition Map”. In: *Sensors* 20.8, pp. 2166–2188. ISSN: 1424-8220.
- Kaplan, Elliott and Christopher Hegarty (2005). *Understanding GPS: principles and applications*. Artech house.
- Keivan, Nima, Alonso Patron-Perez, and Gabe Sibley (2016). “Asynchronous adaptive conditioning for visual-inertial SLAM”. In: *Experimental Robotics*. Springer, pp. 309–321.
- Kelly, Jonathan and Gaurav S Sukhatme (2011). “Visual-inertial sensor fusion: Localization, mapping and sensor-to-sensor self-calibration”. In: *The International Journal of Robotics Research* 30.1, pp. 56–79.
- Kerl, Christian, Jurgen Sturm, and Daniel Cremers (2013). “Dense visual SLAM for RGB-D cameras”. In: *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*. IEEE, pp. 2100–2106.
- Klein, Georg and David Murray (2008). “Improving the agility of keyframe-based SLAM”. In: *Computer Vision-ECCV 2008*, pp. 802–815.
- Kong, Xianglong et al. (2015). “Tightly-coupled stereo visual-inertial navigation using point and line features”. In: *Sensors* 15.6, pp. 12816–12833.
- Konolige, Kurt, Motilal Agrawal, and Joan Sola (2010). “Large-scale visual odometry for rough terrain”. In: *Robotics research*. Springer, pp. 201–212.
- Kühne, Thomas (2005). “What is a Model?” In: *Language Engineering for Model-Driven Software Development*. Dagstuhl Seminar Proceedings 04101.
- Latif, Yasir, Cesar Cadena, and José Neira (2013). “Robust loop closing over time”. In: *Proc. Robotics: Science Systems*, pp. 233–240.
- Le Marchand, Olivier et al. (2009). “Automotive localization integrity using proprioceptive and pseudo-ranges measurements”. In: *Accurate Localization for Land Transportation*. Vol. 125. Les Collections de l’INRETS, pp. 7–12.

- Lee, Young C et al. (1986). “Analysis of range and position comparison methods as a means to provide GPS integrity in the user receiver”. In: *Proceedings of the 42nd Annual Meeting of the Institute of Navigation*. Citeseer, pp. 1–4.
- Leutenegger, Stefan et al. (Mar. 2015). “Keyframe-based Visual-inertial Odometry Using Nonlinear Optimization”. In: *Int. J. Rob. Res.* 34.3, pp. 314–334. ISSN: 0278-3649.
- Lhuillier, Maxime (2005). “Automatic structure and motion using a catadioptric camera”. In: *Proceedings of the 6th Workshop on Omnidirectional Vision, Camera Networks and Non-Classical Cameras*.
- Li, Liang, Mohammed Quddus, and Lin Zhao (2013). “High accuracy tightly-coupled integrity monitoring algorithm for map-matching”. In: *Transportation Research Part C: Emerging Technologies* 36, pp. 13–26.
- Li, Lin et al. (2017). “Extraction of Road Intersections from GPS Traces Based on the Dominant Orientations of Roads”. In: *ISPRS Int. J. Geo Inf.* 6, p. 403.
- Li, You and Yassine Ruichek (June 2014). “Occupancy Grid Mapping in Urban Environments from a Moving On-Board Stereo-Vision System”. In: *Sensors (Basel, Switzerland)* 14, pp. 10454–10478.
- Liang, Wang, Yihuan Zhang, and Jun Wang (2017). “Map-Based Localization Method for Autonomous Vehicles Using 3D-LIDAR”. In: *IFAC-PapersOnLine* 50, pp. 276–281.
- Liu, Hang et al. (2019). “A Precise and Robust Segmentation-Based Lidar Localization System for Automated Urban Driving”. In: *Remote Sensing* 11.11, pp. 1348–1366.
- Longuet-Higgins, H Christopher (1987). “A computer algorithm for reconstructing a scene from two projections”. In: *Readings in Computer Vision: Issues, Problems, Principles, and Paradigms*, MA Fischler and O. Firschein, eds, pp. 61–62.
- Lu, Wenjie et al. (Dec. 2014). “Lane Marking Based Vehicle Localization Using Particle Filter and Multi-Kernel Estimation”. In: *The 13th International Conference on Control, Automation, Robotics and Vision, ICARCV 2014*. Marina Bay Sands, Singapore.
- Mammeri, Abdelhamid, Azzedine Boukerche, and Zongzhi Tang (2016). “A real-time lane marking localization, tracking and communication system”. In: *Computer Communications* 73, pp. 132–143.
- Matthies, Larry (1989). “Dynamic stereo vision”. PhD thesis. Carnegie Mellon University.
- Matthies, Larry and STEVENA Shafer (1987). “Error modeling in stereo navigation”. In: *IEEE Journal on Robotics and Automation* 3.3, pp. 239–248.

- Mencik, Jaroslav (2016). In: *Concise Reliability for Engineers*. Ed. by Jaroslav Mencik. Rijeka: IntechOpen. Chap. 5.
- Mohamed, Sherif AS et al. (2019). “A survey on odometry for autonomous navigation systems”. In: *IEEE Access* 7, pp. 97466–97486.
- Montemerlo, Michael et al. (2002). “FastSLAM: A Factored Solution to the Simultaneous Localization and Mapping Problem”. In: *In Proceedings of the AAAI National Conference on Artificial Intelligence*. AAAI, pp. 593–598.
- Moras, Julien, Véronique Cherfaoui, and Philippe Bonnifait (May 2011). “Credibilist Occupancy Grids for Vehicle Perception in Dynamic Environments”. In: pp. 84–89.
- Moravec, Hans P (1980). *Obstacle avoidance and navigation in the real world by a seeing robot rover*. Tech. rep. DTIC Document.
- Mostafa, Mostafa et al. (2018). “Radar and visual odometry integrated system aided navigation for UAVS in GNSS denied environment”. In: *Sensors* 18.9, p. 2776.
- Mourikis, Anastasios I and Stergios I Roumeliotis (2007). “A multi-state constraint Kalman filter for vision-aided inertial navigation”. In: *Robotics and automation, 2007 IEEE international conference on*. IEEE, pp. 3565–3572.
- Mueller, Andre et al. (Oct. 2011). “GIS-based topological robot localization through LIDAR crossroad detection”. In: Blacksburg, Washington DC, USA, pp. 2001–2008. ISBN: 978-1-4577-2198-4.
- Munguia, R. (2014). “A GPS-aided inertial navigation system in direct configuration”. In: *Journal of Applied Research and Technology* 12.4, pp. 803–814. ISSN: 1665-6423.
- Mur-Artal, Raul, Jose Maria Martinez Montiel, and Juan D Tardos (2015). “ORB-SLAM: a versatile and accurate monocular SLAM system”. In: *IEEE Transactions on Robotics* 31.5, pp. 1147–1163.
- Naumann, Felix and Johann Christoph Freytag (2005). *Completeness of information sources*. Humboldt-Universität zu Berlin, Mathematisch-Naturwissenschaftliche Fakultät.
- Nedevschi, S. et al. (2013). “Accurate Ego-Vehicle Global Localization at Intersections Through Alignment of Visual Data With Digital Map”. In: *IEEE Transactions on Intelligent Transportation Systems* 14.2, pp. 673–687.
- Newcombe, Richard A, Steven J Lovegrove, and Andrew J Davison (2011). “DTAM: Dense tracking and mapping in real-time”. In: *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, pp. 2320–2327.

- Nistér, David (2004). “An efficient solution to the five-point relative pose problem”. In: *IEEE transactions on pattern analysis and machine intelligence* 26.6, pp. 756–770.
- Nistér, David, Oleg Naroditsky, and James Bergen (2004). “Visual odometry”. In: *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on*. Vol. 1. Ieee, pp. I–I.
- Olson, Clark F et al. (2000). “Robust stereo ego-motion for long distance navigation”. In: *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*. Vol. 2. IEEE, pp. 453–458.
- Olson, Clark F et al. (2003). “Rover navigation using stereo ego-motion”. In: *Robotics and Autonomous Systems* 43.4, pp. 215–229.
- Papernot, Nicolas et al. (2017). “Practical black-box attacks against machine learning”. In: *Proceedings of the 2017 ACM on Asia conference on computer and communications security*, pp. 506–519.
- Peynot, Thierry (2012). *Sensor Data Integrity and Mitigation of Perceptual Failures*. Tech. rep. SYDNEY UNIV (AUSTRALIA) AUSTRALIAN CENTER FOR FIELD ROBOTICS.
- Pretto, Alberto, Emanuele Menegatti, and Enrico Pagello (2011). “Omnidirectional dense large-scale mapping and navigation based on meaningful triangulation”. In: *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, pp. 3289–3296.
- Pusztai, Zoltán, Ivan Eichhardt, and Levente Hajder (July 2018). “Accurate Calibration of Multi-LiDAR-Multi-Camera Systems”. In: *Sensors* 18, pp. 2139–2161.
- Quddus, Mohammed A. and Nagendra R. Velaga (2012). “Enhancing Vehicle Positioning Data Through Map-Matching”. In: *Handbook of Intelligent Vehicles*. Ed. by Azim Eskandarian. London: Springer London, pp. 343–364.
- Quist, Eric B and Randal W Beard (2016). “Radar odometry on fixed-wing small unmanned aircraft”. In: *IEEE Transactions on Aerospace and Electronic Systems* 52.1, pp. 396–410.
- Quist, Eric B, Peter C Niedfeldt, and Randal W Beard (2016). “Radar odometry with recursive-RANSAC”. In: *IEEE Transactions on Aerospace and Electronic Systems* 52.4, pp. 1618–1630.

- Rabe, J. et al. (2017). “Ego-lane estimation for downtown lane-level navigation”. In: *2017 IEEE Intelligent Vehicles Symposium (IV)*, pp. 1152–1157.
- Reid, Tyler G.R. et al. (2019). “Localization Requirements for Autonomous Vehicles”. In: *SAE International Journal of Connected and Automated Vehicles 2.3*, pp. 2574–2590. ISSN: 2574-075X.
- Roysdon, P. F. and J. A. Farrell (2017). “GPS-INS outlier detection elimination using a sliding window filter”. In: *2017 American Control Conference*. Seattle, WA, USA, pp. 1244–1249.
- Saito, Rui, Keigo Watanabe, and Isaku Nagai (2015). “Laser odometry taking account of the tilt on the laser sensor”. In: *2015 10th Asian Control Conference (ASCC)*. IEEE, pp. 1–4.
- Sandhu, R., S. Dambreville, and A. Tannenbaum (2008). “Particle filtering for registration of 2D and 3D point sets with stochastic dynamics”. In: *2008 IEEE Conference on Computer Vision and Pattern Recognition*. Anchorage, Alaska, USA, pp. 1–8.
- Sanz Subirana, Jaume et al. (2008). “The User Domain Integrity Assessment Technique”. In: *4th ESA Workshop on Satellite Navigation User Equipment Technologies, NAVITEC 2008*.
- Scaramuzza, Davide, Friedrich Fraundorfer, and Roland Siegwart (2009). “Real-time monocular visual odometry for on-road vehicles with 1-point ransac”. In: *Robotics and Automation, 2009. ICRA '09. IEEE International Conference on*. Ieee, pp. 4293–4299.
- Scaramuzza, Davide et al. (2009). “Absolute scale in structure from motion from a single vehicle mounted camera by exploiting nonholonomic constraints”. In: *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, pp. 1413–1419.
- Schuster, Frank et al. (2016). “Robust localization based on radar signal clustering”. In: *2016 IEEE Intelligent Vehicles Symposium (IV)*. IEEE, pp. 839–844.
- Segal, Aleksandr, Dirk Haehnel, and Sebastian Thrun (2009). “Generalized-icp.” In: *Robotics science and systems*. Vol. 2. 4. Seattle, WA, p. 435.
- Sethian, J (1998). “Fast marching methods and level set methods for propagating interfaces”. In: *Computational Fluid Dynamics, Annual Lecture Series, 29 th, Rhode-Saint-Genese, Belgium*.

- Sharath, M.N., Nagendra R. Velaga, and Mohammed A. Quddus (2019). “A dynamic two-dimensional (D2D) weight-based map-matching algorithm”. In: *Transportation Research Part C: Emerging Technologies* 98, pp. 409–432.
- Shytermeja, E., A. Garcia-Pena, and O. Julien (2014). “Proposed architecture for integrity monitoring of a GNSS/MEMS system with a Fisheye camera in urban environment”. In: *International Conference on Localization and GNSS 2014 (ICL-GNSS 2014)*, pp. 1–6.
- Sirtkaya, S., B. Seymen, and A. A. Alatan (2013). “Loosely coupled Kalman filtering for fusion of Visual Odometry and inertial navigation”. In: *Proceedings of the 16th International Conference on Information Fusion*, pp. 219–226.
- Speidel, Jan et al. (2013). “Integrity for aviation: Comparing future concepts”. In: *Inside GNSS* 8.4, pp. 54–64.
- Stachniss, Cyrill, John J Leonard, and Sebastian Thrun (2016). “Simultaneous localization and mapping”. In: *Springer Handbook of Robotics*. Springer, pp. 1153–1176.
- Staff, Merriam-Webster (2004). *Merriam-Webster’s collegiate dictionary*. Vol. 2. Merriam-Webster.
- Stephenson, Scott (2016). “Automotive applications of high precision GNSS”. PhD thesis. University of Nottingham.
- Strasdat, Hauke, JMM Montiel, and Andrew J Davison (2010). “Scale drift-aware large scale monocular SLAM”. In: *Robotics- Science and Systems VI*.
- Sünderhauf, Niko et al. (2006). “Visual odometry using sparse bundle adjustment on an autonomous outdoor vehicle”. In: *Autonome Mobile Systeme 2005*. Springer, pp. 157–163.
- Szegedy, Christian et al. (2013). “Intriguing properties of neural networks”. In: *arXiv preprint arXiv:1312.6199*.
- Tao, Zui et al. (2013). “Mapping and localization using GPS, lane markings and proprioceptive sensors”. In: *2013 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, pp. 406–412.
- Tardif, Jean-Philippe, Yanis Pavlidis, and Kostas Daniilidis (2008). “Monocular visual odometry in urban environments using an omnidirectional camera”. In: *Intelligent Robots and Systems, 2008. IROS 2008. IEEE/RSJ International Conference on*. IEEE, pp. 2531–2538.

- Tardif, Jean-Philippe et al. (2010). “A new approach to vision-aided inertial navigation”. In: *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE, pp. 4161–4168.
- Toledo-Moreo, R., D. Betaille, and F. Peyret (2010). “Lane-Level Integrity Provision for Navigation and Map Matching With GNSS, Dead Reckoning, and Enhanced Maps”. In: *IEEE Transactions on Intelligent Transportation Systems* 11.1, pp. 100–112.
- Toledo-Moreo, R. et al. (2007). “High-Integrity IMM-EKF-Based Road Vehicle Navigation With Low-Cost GPS/SBAS/INS”. In: *IEEE Transactions on Intelligent Transportation Systems* 8.3, pp. 491–511. ISSN: 1524-9050.
- Toledo-Moreo, Rafael et al. (2018). “Chapter 4 - Positioning and Digital Maps”. In: *Intelligent Vehicles*. Butterworth-Heinemann, pp. 141–174.
- Tossaint, M et al. (2007). “The Stanford–ESA Integrity Diagram: A New Tool for The User Domain SBAS Integrity Assessment”. In: *Navigation* 54.2, pp. 153–162.
- Toth, Charles et al. (2017). “Positioning Slow-Moving Platforms by UWB Technology in GPS-Challenged Areas”. In: *Journal of Surveying Engineering* 143, p. 04017011.
- Triggs, Bill et al. (1999). “Bundle adjustment—A modern synthesis”. In: *International workshop on vision algorithms*. Springer, pp. 298–372.
- Velaga, Nagendra R et al. (2012). “Map-aided integrity monitoring of a land vehicle navigation system”. In: *IEEE Transactions on Intelligent Transportation Systems* 13.2, pp. 848–858.
- Viandier, N. et al. (2008). “Gnss Performance Enhancement in Urban Environment Based on Pseudo-range Error Model”. In: *2008 IEEE/ION Position, Location and Navigation Symposium*, pp. 377–382.
- Viandier, N. et al. (2010). “GNSS pseudorange error density tracking using Dirichlet Process Mixture”. In: *2010 13th International Conference on Information Fusion*, pp. 1–7.
- Vishen, Rahul, Marius C Silaghi, and Joerg Denzinger (2015). “GPS Data Interpolation: Bezier Vs. Biarcs for Tracing Vehicle Trajectory”. In: *International Conference on Computational Science and Its Applications*. Springer, pp. 197–208.
- Vivet, Damien, Paul Checchin, and Roland Chapuis (2013). “Localization and mapping using only a rotating FMCW radar sensor”. In: *Sensors* 13.4, pp. 4527–4552.

- Wang, Richard Y., M.P. Reddy, and Henry B. Kon (1995). “Toward quality data: An attribute-based approach”. In: *Decision Support Systems* 13.3. Information technologies and systems, pp. 349–372. ISSN: 0167-9236.
- Wang, Richard Y. and Diane M. Strong (1996). “Beyond Accuracy: What Data Quality Means to Data Consumers”. In: *Journal of Management Information Systems* 12.4, pp. 5–33. ISSN: 07421222.
- Weiss, S. et al. (2012). “Real-time onboard visual-inertial state estimation and self-calibration of MAVs in unknown environments”. In: *2012 IEEE International Conference on Robotics and Automation*, pp. 957–964.
- White, Christopher E, David Bernstein, and Alain L Kornhauser (2000). “Some map matching algorithms for personal navigation assistants”. In: *Transportation Research Part C: Emerging Technologies* 8.1, pp. 91–108.
- Worner, M. et al. (2016). “Integrity for autonomous driving: A survey”. In: *2016 IEEE/ION Position, Location and Navigation Symposium*. Savannah, Georgia, USA, pp. 666–671.
- Wu, Kai et al. (2014). “Enhanced monocular visual odometry integrated with laser distance meter for astronaut navigation”. In: *Sensors* 14.3, pp. 4981–5003.
- [www.egnos-user-support.essp-sas.eu](http://www.egnos-user-support.essp-sas.eu). [https://egnos-user-support.essp-sas.eu/new\\_egnos\\_ops/protection\\_levels](https://egnos-user-support.essp-sas.eu/new_egnos_ops/protection_levels). (Accessed: 2020-07-16).
- Xu, Fenglei et al. (2019). “A real-time road detection method based on reorganized lidar data”. In: *PLOS ONE* 14.4, pp. 1–17.
- Yan, Min et al. (2017). “Loose coupling visual-lidar odometry by combining VISO2 and LOAM”. In: *2017 36th Chinese Control Conference (CCC)*. IEEE, pp. 6841–6846.
- Yu, Chunlei, Véronique Cherfaoui, and Philippe Bonnifait (June 2015). “Evidential occupancy grid mapping with stereo-vision”. In: pp. 712–717.
- Zhang, Ji and Sanjiv Singh (2014). “LOAM: Lidar Odometry and Mapping in Real-time.” In: *Robotics: Science and Systems*. Vol. 2. 9.
- Zheng, Sheng et al. (2008). “Robust smooth fitting method for LIDAR data using weighted adaptive mapping LS-SVM”. In: *Proc SPIE* 7144.21, pp. 5669–5683.
- Zhou, Dingfu, Y. Dai, and Hongdong Li (2016). “Reliable scale estimation and correction for monocular Visual Odometry”. In: *2016 IEEE Intelligent Vehicles Symposium (IV)*, pp. 490–495.

- Zhu, Ni et al. (2018). “GNSS Position Integrity in Urban Environments: A Review of Literature”. In: *IEEE Transactions on Intelligent Transportation Systems* 19.9, pp. 2762–2778.
- Zhuang, Yan et al. (2011). “3D-laser-based visual odometry for autonomous mobile robot in outdoor environments”. In: *2011 3rd International Conference on Awareness Science and Technology (iCAST)*. IEEE, pp. 133–138.
- Zinoune, Clément, Philippe Bonnifait, and Javier Ibañez-Guzmán (2014). “Integrity monitoring of navigation systems using repetitive journeys”. In: *2014 IEEE Intelligent Vehicles Symposium Proceedings*. Ypsilanti, Michigan, USA, pp. 274–280.
- (2016). “Sequential FDIA for autonomous integrity monitoring of navigation maps on board vehicles”. In: *IEEE Transactions on Intelligent Transportation Systems* 17.1, pp. 143–155.



**Titre:** Contributions à l'analyse de l'intégrité des sources de données d'un système de localisation multi-modale

**Mots clés:** Véhicules autonomes, source de données multimodale, évaluation de l'intégrité, localisation, modélisation d'erreur, véhicules intelligents

**Résumé:** Les véhicules intelligents sont un élément clé pour des systèmes de transport plus sûrs, efficaces et accessibles à travers le monde. En raison de la multitude de sources de données et de processus associés aux véhicules intelligents, la fiabilité de l'ensemble du système dépend fortement de la possibilité d'erreurs ou de mauvaises performances observées dans ses composants. Dans notre travail, nous nous intéressons à la tâche critique de localisation des véhicules intelligents et relevons les défis de la surveillance de l'intégrité des sources de données utilisées dans la localisation. La contribution clé de notre recherche est la proposition d'un nouveau protocole d'intégrité en combinant les concepts d'intégrité des systèmes d'information et les concepts d'intégrité existants dans les systèmes de transport intelligents (STI). Un cadre de surveillance de l'intégrité basé sur le protocole d'intégrité proposé qui peut gérer les problèmes de localisation multimodale est développé. Des techniques d'estimation de l'intégrité basées sur la cohérence pour les sources de données sont développées et le cadre est testé sur les autoroutes ainsi que dans des scénarios urbains et semi-urbains complexes. Ces itérations démontrent que le cadre proposé est capable de fournir des estimations d'intégrité continues de différents types de sources de données utilisées dans la localisation intelligente des véhicules.

**Title:** Integrity Analysis of Data Sources in Multimodal Localization System

**Keywords:** Autonomous vehicles, multimodal data sources, integrity assessment, localization, error modeling, intelligent vehicles

**Abstract:** Intelligent vehicles are a key component in humanity's vision for safer, efficient, and accessible transportation systems across the world. Due to the multitude of data sources and processes associated with Intelligent vehicles, the reliability of the total system is greatly dependent on the possibility of errors or poor performances observed in its components. In our work, we focus on the critical task of localization of intelligent vehicles and address the challenges in monitoring the integrity of data sources used in localization. The primary contribution of our research is the proposition of a novel protocol for integrity by combining integrity concepts from information systems. An integrity monitoring framework based on the theorized integrity protocol that can handle multimodal localization problems is developed. Consistency-based integrity estimation techniques for data sources are developed and the framework is tested in highways as well as complex urban and semi-urban scenarios. These iterations demonstrate that the proposed framework is capable of providing continuous integrity estimates of different types of data sources used in intelligent vehicle localization.

**Université Paris-Saclay**

Espace Technologique / Immeuble Discovery

Route de l'Orme aux Merisiers RD 128 / 91190 Saint-Aubin, France