

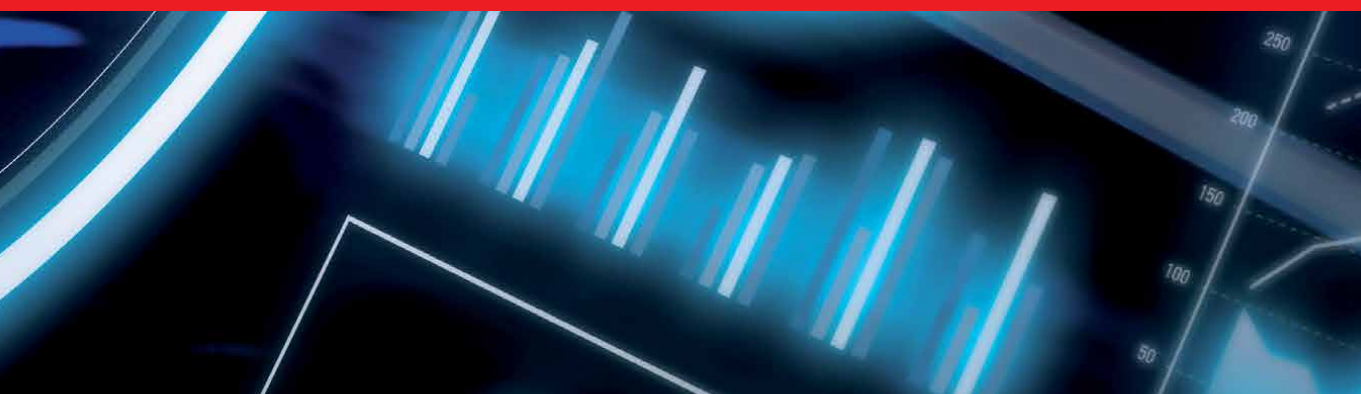


**IntechOpen**

IntechOpen Book Series  
Artificial Intelligence, Volume 6

# Self-Driving Vehicles and Enabling Technologies

*Edited by Marian Găiceanu*





---

# Self-Driving Vehicles and Enabling Technologies

*Edited by Marian Găiceanu*

Published in London, United Kingdom

---



## IntechOpen





*Supporting open minds since 2005*



Self-Driving Vehicles and Enabling Technologies  
<http://dx.doi.org/10.5772/intechopen.87678>  
Edited by Marian Găiceanu

Part of IntechOpen Book Series: Artificial Intelligence, Volume 6  
Book Series Editor: Marco Antonio Aceves-Fernandez

#### Contributors

Christian Wolmar, Prem Chand Jain, Michela Longo, Wahiba Yaïci, Federica Foiadelli, Denis Kotarski, Petar Piljek, Josip Kasać, Said Easa, Yang Ma, Ashraf Elshorbagy, Ahmed Shaker, Songnian Li, Shrinivas Arkatkar, Witold Bużantowicz, Piotr Turek, Stanisław Grzywiński, Richard A. Anthony Guinee, Marian Găiceanu

© The Editor(s) and the Author(s) 2021

The rights of the editor(s) and the author(s) have been asserted in accordance with the Copyright, Designs and Patents Act 1988. All rights to the book as a whole are reserved by INTECHOPEN LIMITED. The book as a whole (compilation) cannot be reproduced, distributed or used for commercial or non-commercial purposes without INTECHOPEN LIMITED's written permission. Enquiries concerning the use of the book should be directed to INTECHOPEN LIMITED rights and permissions department ([permissions@intechopen.com](mailto:permissions@intechopen.com)).

Violations are liable to prosecution under the governing Copyright Law.



Individual chapters of this publication are distributed under the terms of the Creative Commons Attribution 3.0 Unported License which permits commercial use, distribution and reproduction of the individual chapters, provided the original author(s) and source publication are appropriately acknowledged. If so indicated, certain images may not be included under the Creative Commons license. In such cases users will need to obtain permission from the license holder to reproduce the material. More details and guidelines concerning content reuse and adaptation can be found at <http://www.intechopen.com/copyright-policy.html>.

#### Notice

Statements and opinions expressed in the chapters are these of the individual contributors and not necessarily those of the editors or publisher. No responsibility is accepted for the accuracy of information contained in the published chapters. The publisher assumes no responsibility for any damage or injury to persons or property arising out of the use of any materials, instructions, methods or ideas contained in the book.

First published in London, United Kingdom, 2021 by IntechOpen

IntechOpen is the global imprint of INTECHOPEN LIMITED, registered in England and Wales, registration number: 11086078, 5 Princes Gate Court, London, SW7 2QJ, United Kingdom  
Printed in Croatia

British Library Cataloguing-in-Publication Data

A catalogue record for this book is available from the British Library

Additional hard and PDF copies can be obtained from [orders@intechopen.com](mailto:orders@intechopen.com)

Self-Driving Vehicles and Enabling Technologies

Edited by Marian Găiceanu

p. cm.

Print ISBN 978-1-83881-067-2

Online ISBN 978-1-83881-072-6

eBook (PDF) ISBN 978-1-83881-076-4

ISSN 2633-1403

# We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

**5,400+**

Open access books available

**134,000+**

International authors and editors

**165M+**

Downloads

**156**

Countries delivered to

Our authors are among the  
**Top 1%**

most cited scientists

**12.2%**

Contributors from top 500 universities



**WEB OF SCIENCE™**

Selection of our books indexed in the Book Citation Index  
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?  
Contact [book.department@intechopen.com](mailto:book.department@intechopen.com)

Numbers displayed above are based on latest data collected.  
For more information visit [www.intechopen.com](http://www.intechopen.com)







# IntechOpen Book Series

# Artificial Intelligence

## Volume 6



Professor Marian Găiceanu graduated from the Naval and Electrical Engineering Faculty of “Dunarea de Jos University” of Galati in 1997. He received a Ph.D. (Magna Cum Laude) in Electrical Engineering in 2002. Since 2017, Dr. Găiceanu has served as a Ph.D. supervisor in Electrical Engineering. Since 1996, he has been employed at “Dunarea de Jos University” of Galati where he is currently a professor. He is also a member of the *National Council for Attesting Titles, Diplomas and Certificates; an expert of the Executive Agency for Higher Education, Research Funding, and member of the Senate of Dunarea de Jos University*. Since March 2016, Dr. Găiceanu has been head of the Integrated Energy Conversion Systems and Advanced Control of Complex Processes Research Center, Romania. Dr. Găiceanu has conducted several projects in the area of power converter systems for electrical drives, power quality, PEM and SOFC fuel cell power converters for utilities, electric vehicles, and marine applications with the Department of Regulation and Control, SIEI S.p.A. (2002–2004) and Polytechnic of Turin, Italy (2002–2004, 2006–2007). He is an IEEE member and cofounder-member of IEEE Power Electronics Romanian Chapter. He is a guest editor at *Energies*, academic editor for IntechOpen, and member of the *Journal of Electrical Engineering, Electronics, Control and Computer Science* and *Sustainability* editorial boards. He has also served as general chairman of the IEEE International Symposium on Electrical and Electronics Engineering, ISEEE, in the last five editions.

### **Editor of Volume 6:** **Marian Găiceanu**

Department of Automatic Control and Electrical Engineering, “Dunărea de Jos”  
University of Galati, Galati, Romania

**Book Series Editor: Marco A. Aceves-Fernandez**  
Universidad Autónoma de Querétaro  
Faculty of Engineering, Querétaro, México

## Scope of the Series

Artificial Intelligence (AI) is a rapidly developing multidisciplinary research area that aims to solve increasingly complex problems. In today’s highly integrated world, AI promises to become a robust and powerful mean for obtaining solutions to previously unsolvable problems. This book series is intended for researchers and students alike, as well as all those interested in this fascinating field and its applications, in particular in areas related to the topics on which it is focused.



# Contents

<b>Preface</b>	<b>XIII</b>
<b>Section 1</b> Towards Self-Driving Vehicles	<b>1</b>
<b>Chapter 1</b> Introductory Chapter: European Union towards Self-Driving Car Pathway. M2M Era <i>by Marian Găiceanu</i>	<b>3</b>
<b>Chapter 2</b> The Long Journey of the Driverless Car <i>by Christian Wolmar</i>	<b>11</b>
<b>Section 2</b> Design Issues	<b>19</b>
<b>Chapter 3</b> Design Considerations for Autonomous Cargo Transportation Multirotor UAVs <i>by Denis Kotarski, Petar Piljek and Josip Kasac</i>	<b>21</b>
<b>Chapter 4</b> Novel Application of Fast Simulated Annealing Method in Brushless Motor Drive (BLMD) Dynamical Parameter Identification for Electric Vehicle Propulsion <i>by Richard A. Guinee</i>	<b>39</b>
<b>Section 3</b> Enabling Technologies	<b>99</b>
<b>Chapter 5</b> Visibility-Based Technologies and Methodologies for Autonomous Driving <i>by Said Easa, Yang Ma, Ashraf Elshorbagy, Ahmed Shaker, Songnian Li and Shriniwas Arkatkar</i>	<b>101</b>
<b>Chapter 6</b> Selected Issues and Constraints of Image Matching in Terrain-Aided Navigation: A Comparative Study <i>by Piotr Turek, Stanisław Grzywiński and Witold Bużantowicz</i>	<b>129</b>

<b>Section 4</b>	
Future Mobility	<b>151</b>
<b>Chapter 7</b>	<b>153</b>
Future Mobility Advances and Trends	
<i>by Michela Longo, Wahiba Yaïci and Federica Foiadelli</i>	
<b>Chapter 8</b>	<b>173</b>
Trends in Next Generation Intelligent Transportation Systems	
<i>by Prem Chand Jain</i>	

# Preface

More than 90 percent of transportation accidents are caused by drivers and pedestrians. To increase safety on the road and decrease environmental pollution, Electrical Autonomous Driving (EAD) has become a potential solution. The Vision Zero project is a plan from the European Union that aims to achieve a highway system with no fatalities or serious injuries involving road traffic by 2050. One aspect of this plan involves the use of Autonomous Driving (AD). Introducing AD will eliminate car insurance, technical inspections, and maintenance expenses for individuals. Therefore, as road assurance and parking spaces increase, care costs and pollution will decrease.

The book creates a bridge between companies in the AD field and people by eliminating possible risks through technological development. People fear AD, making social acceptance of this technology difficult to achieve. Examples of AD include the recent collaboration of the Computer Science and Artificial Intelligence Lab (CSAIL), Massachusetts Institute of Technology, and the Toyota Research Institute (TRI) to create the MapLite AD navigation system for rural and suburban roads. The use of intelligent control and sensor fusion algorithms for driverless control makes AD possible and safe in cloudy, foggy, snowy, rainy, or sleet conditions. In Boston, WaveSense makes radar for AD to “see” underground via non-invasive penetration technology. As such, innovative road infrastructure can be used to avoid inappropriate surface conditions.

The challenges of AD include finding viable solutions to barriers to all levels of autonomy, supporting AD development, and distributing control algorithms in a comprehensive way. AD includes both inland and ocean water transport, and can refer to autonomous cars, flying cars, unmanned aerial systems, autonomous underwater vehicles, unmanned surface vehicles, and more. It is important to note that the network communication in AD is vulnerable to cyberattacks. Therefore robust and safe management of sensing technologies is integral. Also essential is the adoption of new standards for mass-producing driverless vehicles.

The book is organized into four sections containing two chapters each: “Towards Self-Driving Vehicles”; “Design Issues”; “Enabling Technologies”; and “Future Mobility”. The book examines and discusses the applications of AD in transportation, logistics, space, agriculture, and industrial and home automation.

The editor gives special thanks to the staff at IntechOpen, including Author Service Managers, Ms. Marijana Francetic and Ms. Rebekah Pribetic, and Commissioning Editors, Sandra Maljavac, and Anja Filipovic, for their contributions to the editorial process.

**Marian Găiceanu**  
Department of Automatic Control and Electrical Engineering,  
“Dunărea de Jos” University of Galati,  
Galati, Romania



---

Section 1

# Towards Self-Driving Vehicles

---





# Introductory Chapter: European Union towards Self-Driving Car Pathway. M2M Era

*Marian Găiceanu*

## 1. Introduction

Europa, follows the Wim van de CAMP Euro-parliamentary initiative [1, 2], adopting *On the road to automated mobility: An EU strategy for mobility of the future*, Brussels, 2018. The communication includes the automated and connected mobility for all kinds of transport fields: underwater, water surface, on ground and underground, or by air. The main advantages of adopting Driverless vehicles are as follows: reducing transport costs, increased mobility access (by elder people or with disabilities), sharing mobility, increased safety transport and comfort, more efficient urban planning. For long term, the EU Vision Zero (2050) concept was introduced, which means no road fatalities. The EU regulatory framework has in view the deployment of interoperable Cooperative Intelligent Transport Systems.

The ethical side in EU is equal in worth with the automated mobility, as in worldwide. The driverless vehicle should be safety, socially responsible (to respect the freedom of human choice and the human dignity), efficiently, and environmentally friendly. All these aspects are coordinated and investigated by the European Artificial Intelligence (AI) Alliance. The European AI Alliance born in 2018 with the main target: AI implementation in Europe taking into account the ethics rules in Science and New Technologies.

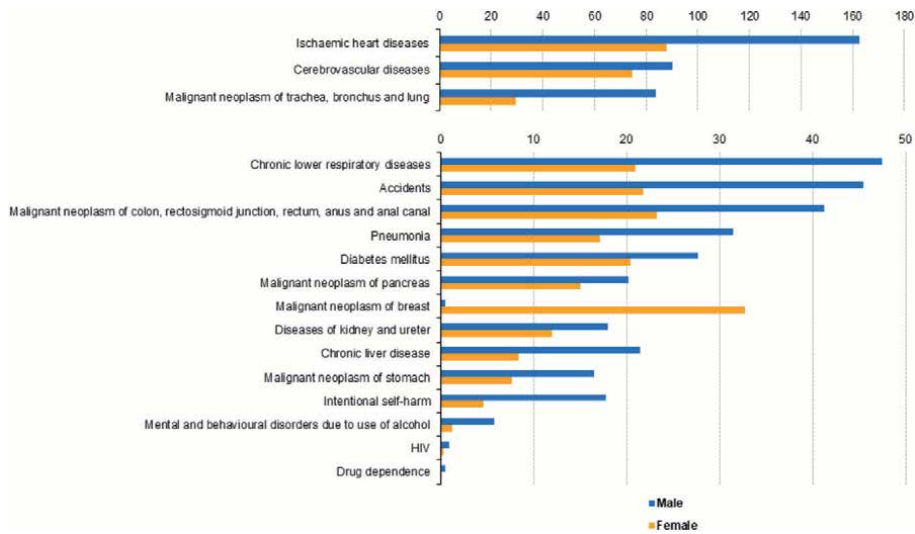
## 2. Motivation

One sight over the causes of death statistics in the European Union (EU) shows that the accident occurrences take the second place (most were male population). Human error [3] is the most cause (95%) in road accidents (**Figure 1**).

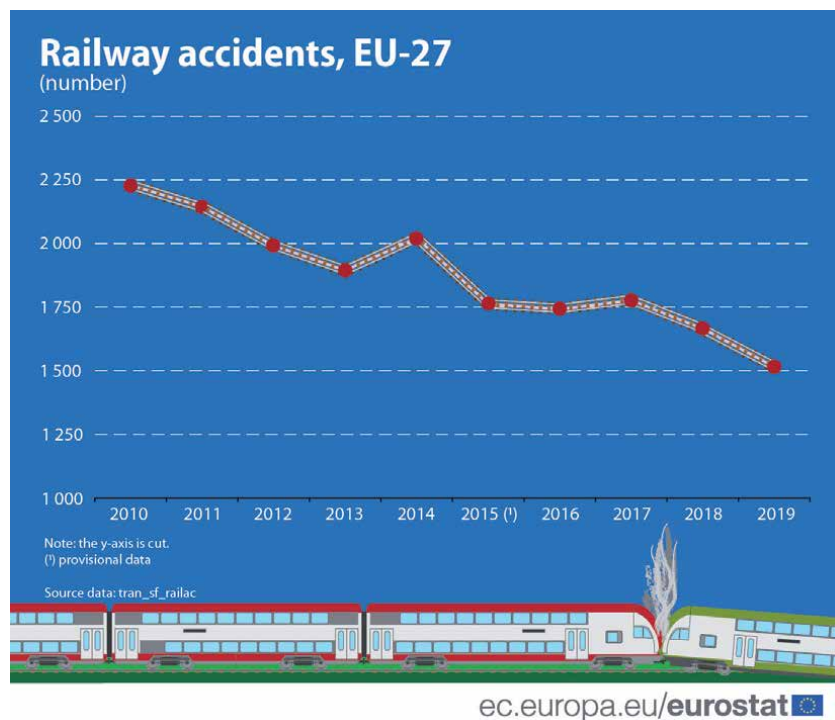
As it is depicted in **Figure 2**, the number of railway accidents in EU27 decreases from 2249 (2010) to 1506 (2019). These accidents cause a total of 802 fatalities, and significant injuries for the other 600.

The number of road accident fatalities per million inhabitants was 50.5 in the EU in 2016. Sweden and the United Kingdom recorded the lowest rate, Bulgaria (99) and Romania (97) the highest rate.

Taking into consideration the statistics result from **Figures 2** and **3**, one dramatic remark can be concluded that the accident fatalities on the road, railway, or by extrapolating on sea/undersea, or by air could be avoided if the human factor can be replaced. Therefore, the European Parliament [7] takes emergent measures [1, 2] to avoid in the near future the accidents in the transport area.



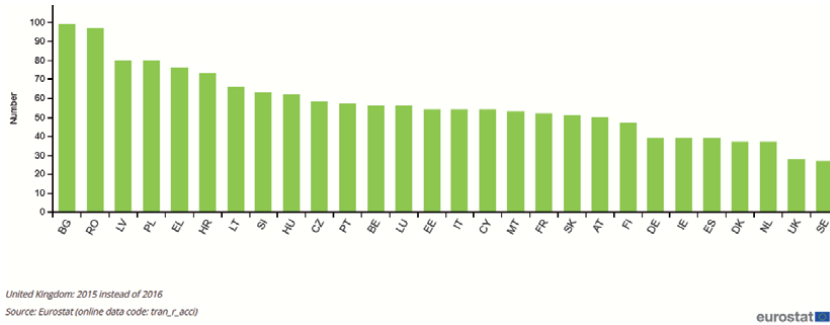
**Figure 1.** Causes of death — Standardised death rate, EU-27, 2016, (per 100 000 inhabitants). Source: Eurostat [4].



**Figure 2.** Statistics of major railway accidents between 2010 and 2019, in EU27 [5].

The exponentially rise of the self-driving market is expected [8]. The main benefits of this trends is new jobs creation, involving the economic growth with safer roads, increased comfort, and more accessibility [9]. Moreover, the non-polluted self-driving cars implies a protection of the environment [9].

The challenges of introducing self-driving cars are in the field of legislation, ethics, cyber-security, maps creation, weather conditions, infrastructure technology, artificial intelligence adoption strategy by each EU country.

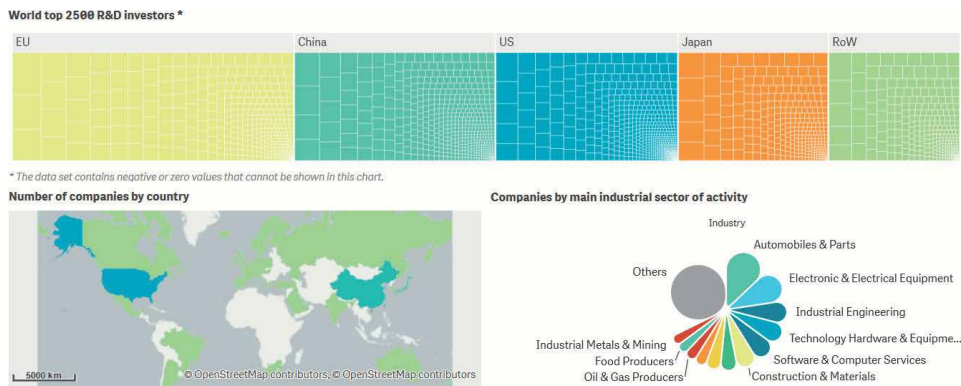


**Figure 3.**  
 The accidents fatalities in road traffic per million inhabitants in 2016 [6].

The main investment of the EU27 Companies in research and development (R&D) is in transport (automobiles) field [10]. The second investment sector is in ICT, followed by health sector (**Figure 4**).

The first two sectors are strong related to the green vehicles development. The self-driving cars development is facilitated by both telecommunication infrastructure development and adopting of artificial intelligence strategy by the most countries. Cybersecurity, safety, and ethical rules should be inside of the regulatory frameworks of self-driving car developing.

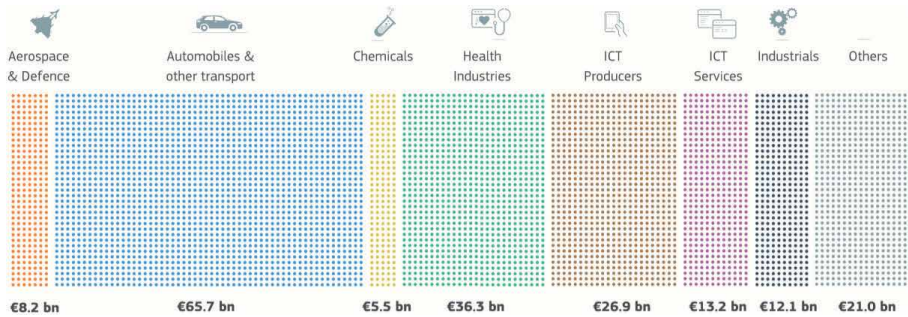
One key factor of developing is the investment in R&D. In Europe, by benchmarking EU Industry Innovation Performance, there is a step forward in the last years, the 5,6% of R&D rising only in one year, 2019, while in the rest of the world by 8,9% [11]. This demonstrates the high rate of investment in R&D in Europe (**Figure 5**).



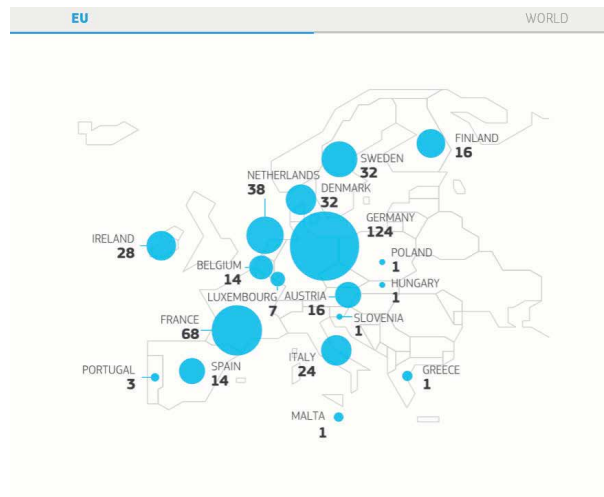
**Figure 4.**  
 Top: World top 2500 R&D investors. Down: Left number of companies by country; right: Companies by main industrial sector of activity [10, 11].



**Figure 5.**  
 R&D investment rate in Europe (2019) [12].



**Figure 6.**  
The sector hierarchy in Europe of the R&D investment [12].



**Figure 7.**  
The country hierarchy in Europe of the R&D investment [12].

By analysing the above-mentioned **Figures 5–7**, the main conclusion is the fast-accelerating rate by the R&D investors in vehicle development sector in EU countries.

### 3. Enabled actions and advanced technologies

The grow-up of the driverless-vehicles is enabled by the on-going research project on the ecosystem (smart cities) and in the transport field.

The H2020 pilot research projects are the main catalysators towards Self-driving car Pathway. AUTOMated driving Progressed by Internet Of Things (AUTOPILOT) [13] is an open source platform for IoT vehicle (containing three main ingredients: vehicles, road infrastructure, and the IoT objects). The main objective includes the self-driving safety. Additional services are investigated (self-parking, terrain mapping or vehicle-sharing).

At the same time, the already open source IoT experimental platforms represent one important tool to test the new IoT technologies before integrating in the real world. Different M2M architectures, cyber physical systems (IoT-A), standards (for example oneM2M) and open platforms (UniversAAL, FIWARE) can be exploited by the pilot projects [14].

The pioneer project 5GCAR: Fifth Generation Communication Automotive Research and innovation [15–17] has in view connectivity and testing of V2X network by an 5G-based system architecture and new V3X services.

The future generation communication infrastructure and connectivity standards will be handled by the 5G Public Private Partnership (5G PPP) [18], in the framework of The European Union Program for Research and Innovation, Horizon 2020 program. The Public Private Partnership is formed by the European Commission and Infrastructure Association (5G IA), the private side [19].

There are ambitious objectives to take care of 5G infrastructure development, and Key Performance Indicators (KPI) to be reach, including 1000 times higher wireless data volume than of 2010, large access (10 to 100 times more connected devices) to low-cost applications and services with a high reliable (End-to-End latency of <1 ms; 10 times to 100 times higher typical user data rate), secure Internet with low energy consumption (10 times lower energy consumption) [20].

Enabling Safe Multi-Brand platooning (ENSEMBLE) for Europe large scale research project has in view the multi-brand truck platooning in real world of the Europe traffic to ensure a safety road, low fuel consumption [21]. One of the project outcomes will be the standardisation of the communication protocols between the trucks platoon, in which each truck is followed by the other by respecting the traffic rules (speed, distance, traffic signs) [22].

The safe distance is mentioned from the preceding vehicle to lead vehicle by using radar or/and laser sensor systems.

According to European Commission Decision 2008/671 (named ‘5.9 GHz ITS Decision’) there is specified for the safety-used communication technologies in Intelligent Transport Systems (ITS) the availability of the 5875–5905 MHz spectrum use for IS applications in the EU [23].

#### **4. Cooperative intelligent transport systems. Towards M2M**

In 2016 European Commission adopted European Strategy on Cooperative Intelligent Transport Systems (C-ITS), rising cooperative, connected and automated mobility (CCAM) [24]. In order to develop C-ITS services, the European Commission has in view to strength of the relationship between the investments and regulatory frameworks. The outcomes of the large pilot projects in intelligent transport will enable the specific standards within C-ITS. In USA, Society of Automotive Engineers (SAE) introduce the level of driving automation in 2014 [25], and updated in 2019 [26].

One year later, on 21st of May 2020, SAE International comes with the new standard, SAE J3216: Taxonomy and Definitions for Terms Related to Cooperative Driving Automation for On-Road Motor Vehicles [27].

The cooperative driving of the self-driving vehicles is the next level of full automation. The new called Cooperative Driving Automation standard was introduced by the SAE on May 2020 through the SAEJ3216 standard. This standard is based to the predecessor SAE J3016 which stipulates the six levels of driving technologies (from 0 to 5).

This new standard, J3216, defines the progress of cooperative control, from A to D classes, enabling machine to machine (M2M) communication systems.

The self-driving vehicles should include both e-safety systems and vehicle safety communications. The e-safety systems comprise at least: Advanced driver-assistance systems (ADAS), automatic emergency braking systems, forward and reverse collision warning (FCW, RCW) system, adaptive cruise control, lane-keeping technology [28–30].



**Figure 8.**  
*The 3 trucks platoon formation [21].*

The safety of the vehicle communications becomes a critical safety point. This is a complex task due to the diversity aspects: ethical, legislative framework, technology, governance. The solutions to this task are offered by the cybersecurity and data protection rules [31].

Practically, the J3216 standard improves the level of full automation by facilitating the car road traffic in platoon formation, as in **Figure 8**. This concept is named by SAE as Cooperative Driving Automation (CDA).

Connected eco-driving concept is based on the V2I real time communication to advice the drivers about the traffic congestion or other infrastructure conditions [32, 33].

## 5. Conclusion

The content of this chapter is structured within 4 Sections. The first Section, there is an Introduction in the European initiative of self-driving technology. The Second Section includes the main motivation of enabling such technologies, the Vision Zero accident fatalities. The third Section, Enabled Actions and Advanced Technologies, includes the main catalysators of implementing self-driving car: the H2020 pilot research projects. The last Section, “*Cooperative Intelligent Transport Systems. Towards M2M*”, presents both the promoted EU tools, i.e., cooperative, connected and automated mobility (CCAM), and USA tools, SAEJ3216 standard, to enable self-driving car pathway towards M2M level.

## Acknowledgements


The project leading to this application has received funding from the Research Fund for Coal and Steel under grant agreement No 899469.

## Author details

Marian Găiceanu  
Integrated Energy Conversion Systems and Advanced Control of Complex  
Processes Research Center, Dunarea de Jos University of Galati, Romania

\*Address all correspondence to: marian.gaiceanu@ugal.ro

## IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] [https://www.europarl.europa.eu/meps/en/96754/WIM\\_VAN+DE+CAMP/history/8](https://www.europarl.europa.eu/meps/en/96754/WIM_VAN+DE+CAMP/history/8) [Accessed: 2021-02-18]
- [2] <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52018DC0283&from=EN> [Accessed: 2021-02-18]
- [3] <https://www.europarl.europa.eu/news/en/headlines/economy/20190110STO23102/self-driving-cars-in-the-eu-from-science-fiction-to-reality> [Accessed: 2021-02-24]
- [4] [https://ec.europa.eu/eurostat/statistics-explained/index.php/Causes\\_of\\_death\\_statistics](https://ec.europa.eu/eurostat/statistics-explained/index.php/Causes_of_death_statistics) [Accessed: 2021-03-11]
- [5] [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Railway\\_accidents\\_2019.png](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=File:Railway_accidents_2019.png) [Accessed: 2021-03-11]
- [6] [https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Road\\_safety\\_statistics\\_-\\_characteristics\\_at\\_national\\_and\\_regional\\_level](https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Road_safety_statistics_-_characteristics_at_national_and_regional_level) [Accessed: 2021-03-11]
- [7] <https://www.europarl.europa.eu> [Accessed: 2021-03-11]
- [8] <https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/analysis-possible-socio-economic-effects-cooperative-connected-and-automated-mobility-ccam> [Accessed: 2021-03-11]
- [9] <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52016DC0787&from=EN> [Accessed: 2021-03-17].
- [10] [https://ec.europa.eu/commission/presscorner/detail/en/IP\\_20\\_2458](https://ec.europa.eu/commission/presscorner/detail/en/IP_20_2458) [Accessed: 2021-03-17]
- [11] <https://iri.jrc.ec.europa.eu/scoreboard/2020-eu-industrial-rd-investment-scoreboard#dialog-node-5706> [Accessed: 2021-03-17]
- [12] [https://iri.jrc.ec.europa.eu/rd\\_monitoring](https://iri.jrc.ec.europa.eu/rd_monitoring) [Accessed: 2021-03-17]
- [13] <https://cordis.europa.eu/project/id/731993> [Accessed: 2021-03-17]
- [14] [https://cordis.europa.eu/programme/id/H2020\\_IoT-01-2016](https://cordis.europa.eu/programme/id/H2020_IoT-01-2016) [Accessed: 2021-03-17]
- [15] <https://cordis.europa.eu/project/id/731993> [Accessed: 2021-03-18]
- [16] <https://cordis.europa.eu/project/id/761510/reporting> [Accessed: 2021-03-18]
- [17] <https://5gcar.eu> [Accessed: 2021-03-18]
- [18] <https://5g-ppp.eu> [Accessed: 2021-03-18]
- [19] <https://5g-ia.eu> [Accessed: 2021-03-18]
- [20] <https://5g-ppp.eu/kpis/> [Accessed: 2021-03-18]
- [21] <https://platooningensemble.eu/> [Accessed: 2021-03-18]
- [22] <https://ec.europa.eu/inea/en/horizon-2020/projects/h2020-transport/automated-road-transport/ensemble> [Accessed: 2021-03-18]
- [23] <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32008D0671> [Accessed: 2021-03-19]
- [24] [https://ec.europa.eu/transport/themes/its/c-its\\_en](https://ec.europa.eu/transport/themes/its/c-its_en) [Accessed: 2021-03-20]
- [25] [https://www.sae.org/standards/content/j3016\\_201401/](https://www.sae.org/standards/content/j3016_201401/) [Accessed: 2021-03-20]

[26] <https://www.sae.org/news/2019/01/sae-updates-j3016-automated-driving-graphic/> [Accessed: 2021-03-20]

[27] [https://www.sae.org/standards/content/j3216\\_202005/](https://www.sae.org/standards/content/j3216_202005/) [Accessed: 2021-03-20]

[28] [https://ec.europa.eu/transport/themes/its/c-its\\_en](https://ec.europa.eu/transport/themes/its/c-its_en) [Accessed: 2021-03-20]

[29] Collision avoidance systems, [https://ec.europa.eu/transport/road\\_safety/specialist/knowledge/esave/esafety\\_measures\\_unknown\\_safety\\_effects/collision\\_avoidance\\_systems\\_en](https://ec.europa.eu/transport/road_safety/specialist/knowledge/esave/esafety_measures_unknown_safety_effects/collision_avoidance_systems_en) [Accessed: 2021-03-20]

[30] Road safety: Commission welcomes agreement on new EU rules to help save lives, [https://ec.europa.eu/commission/presscorner/detail/en/IP\\_19\\_1793](https://ec.europa.eu/commission/presscorner/detail/en/IP_19_1793) [Accessed: 2021-03-20]

[31] [https://www.eca.europa.eu/Lists/ECADocuments/BRP\\_CYBERSECURITY/BRP\\_CYBERSECURITY\\_EN.pdf](https://www.eca.europa.eu/Lists/ECADocuments/BRP_CYBERSECURITY/BRP_CYBERSECURITY_EN.pdf) [Accessed: 2021-03-20]

[32] <https://ziranw.github.io> [Accessed: 2021-03-20]

[33] Automated vehicles in the EU, 07-01-2016, [https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS\\_BRI\(2016\)573902](https://www.europarl.europa.eu/thinktank/en/document.html?reference=EPRS_BRI(2016)573902) [Accessed: 2021-03-20]



# The Long Journey of the Driverless Car

*Christian Wolmar*

## Abstract

There has been considerable hype about the expectations around driverless cars but tests and trials have shown that the concept is far more difficult to bring to fruition than expected. Since around 2010, there have been predictions of the imminent arrival of driverless cars. All these predictions have proved to be over optimistic and none of the goals have been achieved. Companies like Waymo, who are most advanced in the field, are beginning to admit that the task they faced is far more difficult than originally envisaged. This chapter will examine the obstacles to the achievement of the driverless car concept and assess whether the models of shared use driverless vehicle posited by the auto manufacturers and tech companies are realistic.

**Keywords:** autonomous vehicles, driverless cars

## 1. Introduction

For more than a decade, a group of auto manufacturers (OEMs, or Original Equipment Manufacturers) and technology companies have been working on the development of autonomous vehicles. There had, in fact, been work on the concept stretching back to the 1930s but it was only towards the end of the first decade of the 21st century that there was widespread interest in the concept. This seems to have been stimulated by the need for the tech companies, which had generated huge surpluses, to find projects in which to invest their money combined with the desperate fear of motor manufacturers that autonomy would be an essential part of the offer within a few years.

Throughout this period, the claims for this new technology have been ambitious. One of the earliest presentations by a senior motor industry figure was at the Shanghai Expo held in 2010. With a backdrop of a film showing a blind girl being raced through canyons of Shanghai's tower blocks in her driverless pod and a pregnant mother being rushed to hospital in an autonomous ambulance, Kevin Wale, the then boss of General Motors set out his prediction for 2030: 'Our vision for the future is free from petroleum, free from emissions, free from accidents, free from congestion and at the same time fun and fashionable' [1]. The key was for cars to be autonomous. This would ensure, he said, that there would be no traffic jams, no accidents and no emissions since all the vehicles would be electric.

This optimistic view set the tone for much of the subsequent coverage. The presentation of the concept of autonomy has been relentlessly positive emphasising a series of potential advantages. The aspect that is stressed most often is safety.

Protagonists of the new technology point to the fact that about 1.25 million people are killed on roads annually, including around 40,000 in the US. Since more than 90 per cent of these are the result of human error, the claim is that this number could be dramatically reduced. Take out the drivers, and the errors will go with them. Autonomous cars do not get drunk or fall asleep at the wheel, so the argument goes, and therefore they are will undoubtedly be safer. The National Highway Traffic Safety Administration suggested that ‘automated vehicles’ potential to save lives and reduce injuries is rooted in one critical and tragic fact: 94 per cent of serious crashes are due to human error. Automated vehicles have the potential to remove human error from the crash equation, which will help protect drivers and passengers, as well as bicyclists and pedestrians’ [2].

A second key argument is convenience. Regular daily commutes of an hour or more in each direction are commonplace and that time will become available again to the drivers who can use it to answer emails, make calls or even just read a book. There is too, the potential for the technology to enable many more people to use cars, such as blind people or dementia sufferers. This idea was a key part of the presentation by Google’s then head of its autonomous car project at to a congressional committee, Chris Urmson who cited the example of ‘Justin Harford, a man who is legally blind’ who had told him ‘what this is really about is who gets to access transportation and commerce and who doesn’t’ [3]. These comments were met with great enthusiasm by campaigners for people with disabilities such as, for example, Parnell Diggs, the Director of Government Affairs for the National Federation of the Blind, who told the committee ‘we anxiously anticipate the day that all blind people will have the opportunity to driver independently, and we believe that autonomous vehicles will make this day possible’.

A third major advantage claimed for the technology is that people will no longer need to own their own cars. The idea is that vehicles will be shared use, ready to be called up at a moment’s notice through an app. This, in turn, will enable vast swathes of parking areas to be repurposed since once at their destination people will be able to despatch their vehicle to its next user or to distant car parks.

With reduced car ownership, there will be more road space available as cars will no longer be parked on kerbs. Moreover, because autonomous cars will be driven in a controlled way, without the vagaries of human control, there will be a much more efficient use of highways as the well-known wave effect will be eliminated. Zenzic, the organisation which coordinates the UK’s research programme claimed in a press release that connected autonomous vehicles could reduce transport emissions by between 5 to 20 per cent by reducing congestion and was ‘the key to becoming climate neutral’ [4].

There have been numerous attempts to quantify all the gains from the introduction of autonomous vehicles. A study [5] by KPMG, for example, suggested that British drivers would save £5bn per year in reduced insurance, car parking and running costs. A report by Rand [6] argued that the increase in lane capacity on highways might amount to 500 per cent and that autonomous cars would lead to an improvement in fuel use of between 4 to 10 per cent. Ohio University’s Future of Driving report [7] stated that harmful emissions would be reduced by 60 per cent by the introduction of autonomous vehicles. Zenzic claims that the industry will be worth £52 billion in the UK and £907 billion worldwide by 2035 [8].

All this, however, is rather mundane and to make it more exciting the promoters of the technology is that they wrap up these ideas with language that represents a radical and exciting vision for the future such as in the speech by Kevin Wale of GM. There is talk of ‘life-changing’ experience, of ‘freeing up large amounts of time’, of clean air and ‘emptier roads’.

While the various manufacturers and tech companies have different conceptions of what this new driverless world may look like, the long term vision converges around a triple revolution: in the future vehicles will be driverless, electric and shared used. This is the Holy Grail for the industry as in this scenario driverless vehicles would dominated the transport landscape, taking over not just the existing privately driven car market but also making deep inroads into public transit and expanding the use of cars by enabling, as mentioned above, many people with disabilities or without a licence to 'drive'.

## **2. The triple revolution**

This is based on a variety of assumptions around very profound and radical societal changes. Yet, neither the breadth of these changes nor the huge number of obstacles that need to be overcome before this vision can become a reality are examined by those putting out this vision. Quite apart from the depiction of a transport world completely different from the one in which we live today, the very long period during which there would be a mix of driverless and conventionally-driven vehicles is given little attention.

Indeed, the idea that a totally driverless world is possible stretches credulity. The very example given at the Shanghai Expo of an ambulance carrying a pregnant woman is an unlikely scenario for a driverless vehicle as emergency vehicles are allowed to break the rules precisely because they must have priority. Even in a near driverless world, emergency vehicles, VIP limos, other urgent transport and various other types of vehicle are likely to remain conventionally driven.

The scenario presented by the concept's enthusiasts is, in fact, three separate revolutions bundled into one. The least innovatory and radical of these assumptions is that vehicles will increasingly be electric. That is highly likely but upscaling the production and sale of electric cars beyond the current minority market has proved difficult because of the high initial cost, the short range (or more pertinently fears about the range) and the slow rate of development of new models. Currently sales represent around 2.6 per cent of the global market [9]. This is growing but only slowly and there are concerns that the biggest constraint will be the production of sufficient batteries to support a rapid expansion in the electric and hybrid share of the market. The availability of charging points, the difficulties many flat dwellers would have in charging their vehicles overnight and the various issues around the sustainability of batteries all point to a relatively slow take-up of electric vehicles.

Setting these difficulties aside, the second assumption is an even bigger obstacle. The notion that drivers will happily dispense with their own cars once driverless models become widely available and rely on Uber type services to call up vehicles when they are needed has very little evidence to support it. There are indeed a minority of Millennials living in urban areas who are happy to dispense with car ownership. For people who at the moment live in a city served by good private hire and taxi services including Uber, the option of not owning a car is perfectly feasible. However, once they move to the suburbs, or have children, they tend to purchase their own vehicles. For the past century or so, people have bought their own cars, despite the high cost, for a whole host of reasons: convenience, choice of type of vehicle, accessibility, enjoyment and, for many, keeping up or bettering the Joneses. In fact, driving is still considered by many to be a pleasure. The idea that suddenly this will all be abandoned because vehicles will no longer be driven but will be autonomous has little logic and no research to back it up. Indeed, on the contrary, the providers of shared use vehicles accept that 'car clubs are not for everyone and

there are many who still aspire to car ownership, even Millennials. I don't see a time when all vehicles will be shared [10]. People like the convenience of having, say, baby seats, golf clubs or tools in the car and moreover, the guarantee that the car is outside the home for immediate use. Relying on a shared use vehicle accessed through an app when they have to get to work at a particular time or take the kids to school will never be able to replace that flexibility.

There is another practical objection to the model here. At the moment services such as Uber and Lyft are available principally in large urban conurbations. If the world were genuinely to become dominated by share use vehicles, they would have to serve small towns and even villages. There is simply no feasible business model in which such areas would have access to a pool of shared use vehicles at short notice. Even if the shared use model might be widely accepted in central urban areas, it is difficult to envisage it taking off in more sparsely populated suburbs let alone small towns, villages or rural areas. The provision of sufficient cars would simply not be cost effective as no supplier would take the financial risk.

The recent pandemic leads to other difficulties. Who would guarantee that these cars were clean and not full of the previous occupants' litter or, worse, germs? There are a myriad other reasons why this scenario is implausible such as the lack of any business model: the costs of maintaining a service as these vehicles would need supervision and a back-up service; the initial investment required to set up such a business given the cost of the technology; and the reluctance of the public to part with their own vehicles and effectively replace them with an app. This model is, on the face of it, a very strange basis for the massive investment programmes by the tech and auto manufacturers given the lack of evidence that people are prepared to buy into this model. So why has this shared use concept become so important for the autonomous car protagonists?

The reason, in fact, points to their Achilles Heel and demonstrates that the extent to which this triple revolution is an impossible dream that more sensible advocates now see as being 'decades away' [11]. The supporters of autonomous cars have been forced to put forward this shared use scenario because of their fear of the criticism that the advent of driverless cars will lead to an increase in cars on the road and consequently greater congestion. They argue that since cars are in use for only around 5 per cent of their time, having autonomous cars which are shared will lead to a massive reduction in the number of vehicles on the road. There are obvious logical objections to this. Most people want their cars at peak times in the morning and evening, and very few use them at 3 am in the morning. Therefore the parc of vehicles would have to be far higher than the 5 per cent figure which this scenario implies even if all were shared use and driverless. Moreover, no clear business model has been set out for how such a massive business as providing vehicles for, literally, millions of people in a city would work. The practicalities of essentially making available hundreds of thousands of vehicles that would need to be centrally owned by a single entity (competition would add another layer of complexity) has never been set out. This is not an evolutionary process but a revolutionary one. In reality, the prompt for this scenario is the auto manufacturers' concern about understandable concerns that mass autonomy would lead to an increase, not a reduction, in congestion. There is much logic in that argument. If autonomy makes it easier for people to access individual cars rather than public transport, then it is highly likely there will be an increase in demand. Moreover, in a world dominated by autonomous vehicles, there would be considerable mileage undertaken by completely empty cars travelling between users. Uber presently has an average passenger occupancy rate of 0.6 (in addition to the driver) which means their vehicles are been driven for nearly half the time without a passenger. This emphasis on shared use is therefore borne of the necessity to argue that the spread of autonomy will lead

to a reduction in congestion when the opposite has much more logic. It is a defence to a criticism, not a presentation of a realistic scenario.

The third element of this triple revolution, the widespread use of cars that are entirely capable of driving without human intervention is an even tougher obstacle to overcome. At present, the technology is at what has been called level 3. Cars can perform routine driving tasks such as on highways, even selecting routes and not requiring human input for steering but they still require constant attention from the driver. There have been countless tests and trials, and millions of miles have been driven by vehicles that have many features that allow them to be computer-controlled but despite the investment of an estimated \$100bn [12], the technology is nowhere close to delivering a car that can be driven anywhere in any weather conditions with complete safety which is defined at Level 5.

Waymo's 'robo taxi' service in Phoenix, Arizona, and Silicon Valley (for employees only) started operating in December 2018 but has been beset with problems. In fact, all the cars still have safety drivers, except for a minority which are 'geo-fenced' and all are monitored - and sometimes controlled - remotely. Passengers have complained of being dropped off in the wrong place, experiencing unexplained stops sometimes so sudden that they have caused whiplash and near collisions with cyclists: 'In about 2.5 per cent of Phoenix rides and 6.5 per cent of Silicon Valley rides, Waymo vehicles stood still for a long period of time before either the human driver took over or a Waymo representative monitoring the vehicle from a remote location helped the car figure out how to start moving again. One Waymo rider *The Information* that during three trips in one week this summer, the Waymo vehicle got stuck each time' [13].

Most of the testing in the US has been carried out by cars monitored by an operator who is supposed to intervene when things are about to go wrong - something that clearly did not happen when the unfortunate woman wheeling a bike which had bags on its handlebars in Arizona was killed because the car failed to recognise her as human. It identified her initially as a plastic bag and then as a cyclist who was not on a collision course and only too late as a human being. This accident, which caused the death of Elaine Herzberg in Tempe Arizona in March 2018 was a key demonstration of the inability of even the most sophisticated computers to recognise 'outlier' situations. The fact that Herzberg was pushing a bicycle which had bags on its handlebars clearly was not a situation that the on board computer had been programmed to recognise. This is proving to be the biggest single obstacle to progress in the development of the autonomy aspect of these vehicles. However many millions of miles have been covered on the road, they will never be sufficient for the vehicles to learn about all eventualities and therefore the ability to reach full driverlessness must be in doubt. Indeed, despite the large amount of testing that has already taken place, most of the cars still cannot operate in heavy rain, snow or off road.

### **3. Public acceptance**

All of this has helped increase scepticism about the concept. Almost half of Americans say they would not get in a self-driving taxi, according to a poll by the advocacy group Partners for Automated Vehicle Education [14]. The poll, carried out at the beginning of 2020, found that 48 per cent of the 1200 adults surveyed would 'never get in a taxi or ride-share vehicle that was being driven autonomously', while a further 21 per cent said they were unsure about doing so. While a fifth of respondents said that autonomous vehicles would never be safe, another fifth stated, incorrectly, that it is possible 'to own a completely driverless vehicle today',

highlighting the confusion that still remains over how far the technology has already developed. On the other side of the coin, people want to continue driving. A post-pandemic lockdown survey in *Le Monde* [15] found that half of all car owners actually missed driving while they were unable to travel.

*The fact that so many people believe that the driverless car is already a reality is the product of the tremendous hype that has accompanies the investment. entities. In an article for the online academic magazine Transportation Research Interdisciplinary Perspectives [16], Liza Dixon argues that much of the material put out by the companies developing autonomous vehicles is misleading as it fails to distinguish between autonomy and driving aids. She defines autonowashing as making unverified or misleading claims that misrepresent the appropriate level of human supervision required by a partially or semi-autonomous product, service or technology. This is, in fact, a characteristic of much of the PR output of the industry.*

*She cites the use of vague language and the failure to prove claims as being characteristics of autonowashing, and she highlights the media's culpability in relation to its 'utopian' reporting and exaggeration of the level of autonomy. Indeed, there are numerous examples of articles whose headlines suggest they are about 'driverless' vehicles but that go on to reveal that there is a safety driver at the wheel.*

Dixon points out, the phenomenon is somewhat self-defeating for the industry, which depends on building trust among potential users. By exaggerating claims and failing to consider disadvantages, the industry is weakening its own case. She writes: 'Autonowashing leads to overtrust, which leads to misuse. If a driver management system is unable to assist the user in error prevention, accident, injury or death may occur. This results in negative media coverage which can then stir public distrust in vehicle automation, threatening the return on investment.'

The extraordinary level of hype is, in fact, a key part of the current business model which appear to more about attracting investment funds than actually developing a fully autonomous vehicle. Given the clear and obvious obstacles facing the industry, the reasons to justify the vast level of investment are surprisingly unclear. Yet it continues unabated. Waymo managed to raise \$3bn in the market in the Spring of 2020 while survey of the top thirty companies in the field published in *The Information* [17] revealed that \$16 billion was spent on autonomous vehicle R&D in 2019: 'Just three companies spent half of that money – Alphabet's Waymo, GM's Cruise and Uber... Four other companies, including Apple, Baidu, Ford and Toyota, spent most of the rest.' According to a Fortune magazine article of 7 January 2020, while Waymo remains the market leader after eleven years of research, the company 'remains an expensive science project in search of a business.'

The benefits of removing the driver from cars have been heavily promoted by these companies but as we have seen do not stand up to close scrutiny. Since clearly the model of the triple revolution is unlikely ever to be realised, which means that the notion of blind or infirm people being able to regain autonomous mobility is a myth, what of the other purported benefits of a move to driverless vehicles?

The safety benefits are far less marked than suggested by the industry. The Insurance Institution for Highway Safety has calculated [18] that just a third of accidents would be prevented by the use of autonomous vehicles. This is because only accidents that are what the researchers call 'sensing and perception' errors, such as driver distraction or failure to spot a hazard, will be prevented. The technology cannot prevent the majority of accidents, which the IIHS believes are caused by 'prediction errors', such as misjudging the speed of other vehicles, excessive speed

when road conditions are treacherous, and mistaken driver efforts to avoid a crash. One example is when a cyclist swerves into the path of an autonomous car. The vehicle may have seen the cyclist but it cannot manoeuvre quickly enough to avoid hitting them.

These doubts make the motivation of those seeking to promote this technology unclear. There seems to be no short or medium term prospect of making a return on this capital. One driver of the high levels of investment is the assumption that the first to develop full autonomy will make super profits by establishing a monopoly. However even Waymo is now suggesting that the full driverless model is not achievable. An article on CNET [19] in November 2018 quoted the CEO of Waymo, John Krafcik, as expressing doubts over whether autonomous cars would ever become ubiquitous:

‘It’ll be decades before autonomous cars are widespread on the roads – and even then, they won’t be able to drive themselves in certain conditions. Autonomy always will have some constraints’.

While this suggests that there is a need for a model that is very different from the ones previously proposed, there is no sign at this stage of what it is.

Christian Wolmar, author of 20 books principally on transport matters including *Driverless Cars: on a road to nowhere?*, London Partnership Publishing, 2020. [www.christianwolmar.co.uk](http://www.christianwolmar.co.uk)


## Author details

Christian Wolmar  
Independent Researcher, United Kingdom

\*Address all correspondence to: [christian.wolmar@gmail.com](mailto:christian.wolmar@gmail.com)

## IntechOpen

---

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] B. Ji, GM: glimpse into auto future, China Daily, 1 May 2010.
- [2] <https://www.nhtsa.gov/technology-innovation/automated-vehicles-safety>
- [3] Statement by Chris Urmson to the Committee on Commerce, Science and Transportation of the United States Senate, March 15 2016.
- [4] <http://www.pes.eu.com/renewable-news/connected-vehicles-key-to-achieving-net-zero-2050-says-uk-self-driving-experts/>
- [5] <https://home.kpmg/content/dam/kpmg/pdf/2015/04/connected-and-autonomous-vehicles.pdf>
- [6] [https://www.rand.org/content/dam/rand/pubs/research\\_reports/RR400/RR443-2/RAND\\_RR443-2.pdf](https://www.rand.org/content/dam/rand/pubs/research_reports/RR400/RR443-2/RAND_RR443-2.pdf)
- [7] <https://onlinemasters.ohio.edu/blog/the-future-of-driving/>
- [8] This figure is taken from a 2017 Market Forecast report by the Connected Places Catapult, a government research organisation supported by Zenzic.
- [9] <https://www.iea.org/reports/global-ev-outlook-2020>
- [10] Jonathan Hampson, head of Zipcar, London, conversation with the author.
- [11] [https://www.businessinsider.com/self-driving-cars-fully-autonomous-vehicles-future-prediction-timeline-2019-8?amp;r=US&IR=T&r=DE&IR=T&\\_\\_twitter\\_impression=true](https://www.businessinsider.com/self-driving-cars-fully-autonomous-vehicles-future-prediction-timeline-2019-8?amp;r=US&IR=T&r=DE&IR=T&__twitter_impression=true)
- [12] <https://www.ibtimes.com/look-investment-self-driving-cars-who-has-spent-most-2848289>
- [13] <https://www.theinformation.com/articles/waymo-riders-describe-experiences-on-the-road>
- [14] <https://www.cnet.com/roadshow/news/pave-autonomous-vehicle-survey/>
- [15] May 25 2020
- [16] <https://www.journals.elsevier.com/transportation-research-interdisciplinary-perspectives/>
- [17] <https://medium.com/unikie/the-most-interesting-self-driving-car-companies-9df5e15c3cac>
- [18] <https://www.roadtraffic-technology.com/news/self-driving-cars-one-third-crashes-study/>
- [19] <https://www.cnet.com/news/alphabet-google-waymo-ceo-john-krafcik-autonomous-cars-wont-ever-be-able-to-drive-in-all-conditions/>



---

Section 2

# Design Issues

---



# Design Considerations for Autonomous Cargo Transportation Multirotor UAVs

*Denis Kotarski, Petar Piljek and Josip Kasac*

## Abstract

Unmanned aerial vehicles (UAVs) have proven to be an advanced tool for a variety of applications in the civilian and military sectors. Different categories of UAVs are used in various missions and are also the subject of numerous researches. Due to their characteristics and potential in specific conditions, multirotor UAVs imposes itself as a solution for many tasks, including transport. This chapter presents a conceptual solution of autonomous cargo transportation where the primary research objective is the design of a heavy lift multirotor UAV system. The process of designing a multirotor UAV that can carry heavy lift cargo is quite challenging due to many parameters and constraints. Five selected series of electric propulsion systems are analyzed, with different multirotor configurations, and results are graphically displayed for payloads from 10 kg up to 100 kg.

**Keywords:** multirotor UAV, autonomous cargo transportation, heavy lift transport, electric propulsion system, configuration parameters analysis

## 1. Introduction

In recent decades, technology has enabled the development of components and systems that have numerous capabilities in the field of autonomous vehicles and robots. From the aspect of the control system, developed control components with higher processing speed and integrated MEMS sensors allow a certain degree of autonomy of the vehicle. On the other hand, the development of the propulsion system and batteries has enabled a wide range of applications in various missions on the ground, in water, or in the air. Such as autonomous cars, unmanned ground vehicles (UGVs) can use existing infrastructure which is vulnerable to failure, and congestion that can be caused by other vehicles. These problems can potentially be overcome by using unmanned aerial vehicles (UAVs) as autonomous vehicles or/and robots.

There are different categories of UAVs that are used in various missions such as construction management [1, 2], agriculture [3], surveillance [4], search and rescue [5, 6], firefighting [7, 8], transport (delivery) [9] and many others. Aircraft can generally be divided according to the lifting mechanism into fixed-wing [10], rotary-wing [11, 12], and hybrid aircraft [13]. Because of its ability to vertically take-off and land (VTOL), rotary-wing (rotorcraft) UAVs do not need a launchpad or runway so the degree of autonomy can be higher and the cost of supporting infrastructure lower. Rotary-wing UAVs can be further divided into aircraft with variable-pitch

propellers where the typical representative is a helicopter, and multirotor (multi-copter) aircraft with fixed-pitch propellers. The advantages of the multirotor UAVs over other categories of aircraft are agility and maneuverability which is important in missions that include interaction with the environment and precise movements. Therefore this type of UAV is increasingly being considered as an alternative to UGVs in delivery and transport. The design of an autonomous heavy lift cargo transportation multirotor UAV is a quite challenging process since this type of aircraft is characterized by high energy consumption. Various conventional aircraft configurations such as quadrotor [14], hexarotor [15], or octocopter [16], allow a wide range of applications. Numerous research has recently been conducted with the aim of the design and development of new configurations with improved performance [17, 18].

There are several groups of researchers and companies engaged in the research and development of multirotor UAVs for the transportation of heavy loads. Brar et al. in their technical report [19] addressed several aspects of UAVs for deliveries such as the current market, available technology, regulation, and the impact on society. In general, multirotor cargo transport can be achieved with two basic transport strategies. The load can be attached to the multirotor body, or suspended through cables [20]. Ong et al. presented design methodology for heavy-lift UAVs with coaxial rotors [21]. Several companies deal with the production of aircraft for the transportation of heavy cargo [22, 23]. The motivation to design an autonomous system stems from the fact that such a system can improve some aspects of society. Life on the islands is specific given the needs of the population and the infrastructure and institutions that exist on certain islands. On the larger islands, there are schools, ambulance, post offices. All other needs of the island's inhabitants are met through the connections that exist with the mainland and are mostly met by sea. Transport of passengers, goods, transport vehicles, and others takes place by ferries and ships. The existing type of transport is characterized by limited line frequency and high transportation costs.

In this chapter, a conceptual solution of on-demand autonomous heavy lift cargo transportation is presented which can reduce operational costs and carbon emissions. The overall system consists of a network of multirotor UAVs and docking facilities with the purpose of transportation from the mainland to the islands and vice versa. Based on the distance analysis and the existing infrastructure on the considered central Adriatic islands in Croatia, the topology (network) of the system was proposed whose endpoints are Zadar ports (mainland). The main focus of this research is on the design of heavy lift multirotor UAVs which can carry loads from 10 kg up to 100 kg. The multirotor UAV system divided into four key subsystems allows a methodological approach to aircraft design. The performance of the multirotor UAV is determined by the parameters and components of the propulsion and energy subsystem. The parameter analysis of conventional configurations for five selected setups of electric propulsion units was performed and presented. Based on the analysis, it is possible to select the aircraft parameters and components for a particular cargo and planned flight route. Furthermore, the parametric design of the aircraft is presented and preliminary simulations are performed.

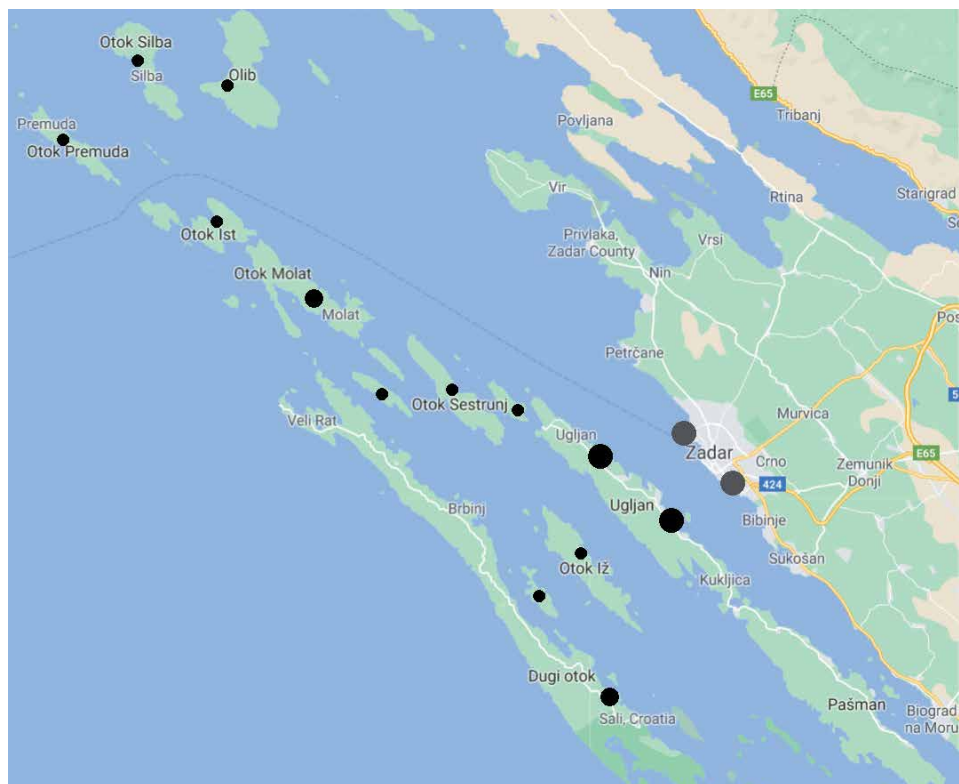
## **2. Autonomous cargo transportation system concept**

In this research, the aim is to show the benefits of the autonomous cargo transportation system (ACTS). The implementation of on-demand ACTS using multirotor UAVs can potentially reduce transport costs and increase the frequency and speed of transportation. This is of particular importance for the inhabitants of the islands, for whom such a system would enable better communication with the mainland, and thus would improve the standard of living on the islands. The concept of the system

allows for a multi-purpose character and could be used for missions involving the delivery of postal packages, the transport of goods such as fresh fish and fruit, and such a system could even be used for fire prevention purposes. The system consists of docking facilities on certain islands and a fleet of multirotor UAVs deployed on an on-demand basis that connect islands facilities to the overall system.

A case study was considered for the central Adriatic islands, which administratively belong to the city of Zadar, which is located on the coast of the Croatian mainland (**Figure 1**). The archipelago consists of ten smaller and three larger inhabited islands, two of which are connected by a bridge. Common to all islands is that there is the infrastructure to accommodate smaller and/or larger ships and ferries. On these islands, among other facilities, there are 21 post offices as shown in **Table 1**, which is important from the aspect of the delivery and distribution of postal packages. The ACTS can consist of several to several dozen multirotor UAVs, depending on the needs and scope of purposes that such a system can perform. A fleet of aircraft is considered, consisting of five series of aircraft that can carry from 10 kg up to 100 kg of cargo. This also enables the modular character of the overall system, and the use of different aircraft series depending on the cargo they need to carry can potentially reduce the energy consumption. The idea is that the smaller islands are connected to the larger islands with the possibility of using particular islands as an intermediate docking, and the aircraft with the largest payload are provided for connection to the mainland.

Docking facilities have several functions, among others, they need to perform a user-friendly interface that makes it simple to use for the residents of the island and the services bear in mind that this is an on-demand system. Taking into account the current state of available components and technologies, the docking facility could consist

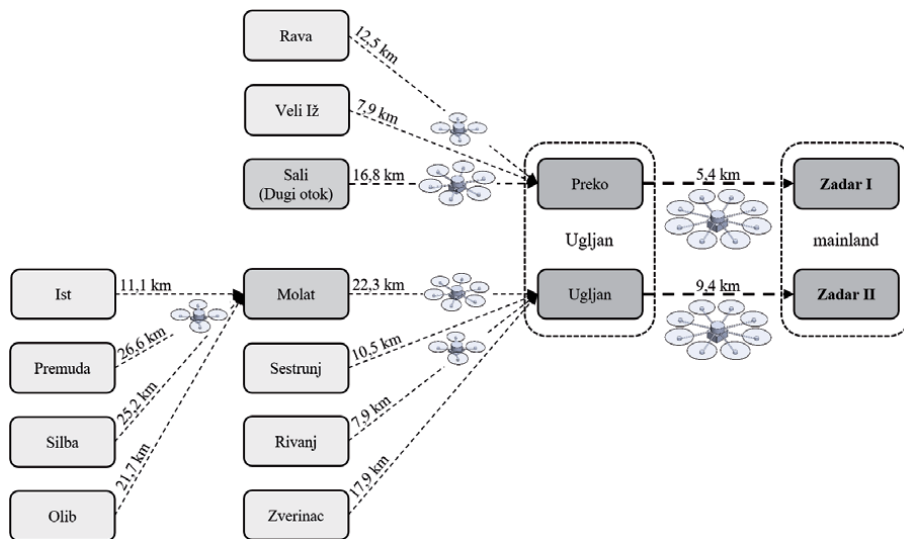


**Figure 1.**  
*Considered archipelago of central Adriatic islands in Croatia.*

Pašman	23,262	Ugljan	23,275	Veli Rat	23,287
Ždrelac	23,263	Sali	23,281	Sestrunj	23,291
Nevidane	23,264	Žman	23,282	Molat	23,292
Kukljica	23,271	Rava	23,283	Ist	23,293
Kali	23,272	Veli Iž	23,284	Premuda	23,294
Preko	23,273	Brbinj	23,285	Silba	23,295
Lukoran	23,274	Božava	23,286	Olib	23,296

**Table 1.**  
Post offices located on the central Adriatic islands.

of a multirotor UAV docking assembly and storage for depositing or retrieving cargo. It can additionally be contained with other features such as a mini solar power plant or a battery charging module for electric cars or UGVs. The multirotor UAV docking assembly is connected to the aircraft via sensors and telemetry and have to be designed to allow take-off and landing of the aircraft. It consists of a module to recharge the multirotor UAV batteries or additionally replace the batteries when the aircraft needs to be used urgently. The storage for depositing or retrieving cargo is associated with the user interface and is connected to the docking assembly to allow the exchange of cargo between the user and the aircraft. **Figure 2** schematically shows the possible topology of the overall ACTS where the most distant islands are connected to the mainland by three connections. It is important to note that the marked distances are for planned routes where flight over settlements and infrastructure is avoided.



**Figure 2.**  
Considered topology of the overall ACTS system.

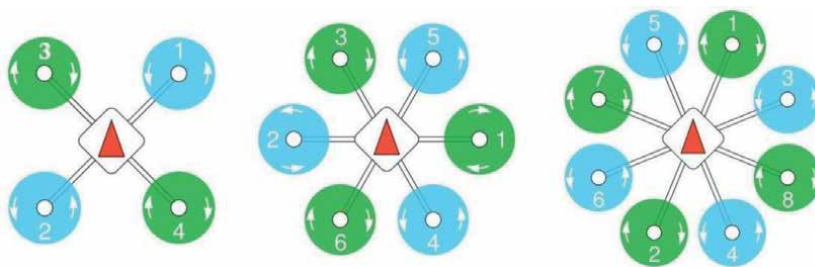
### 3. Multirotor UAV system description

Multirotor UAVs are mechanical systems represented as rigid bodies with six degrees of freedom. Common to all designs is that they consist of N rotors (propulsion units) whose geometric arrangement also determines the configuration of the

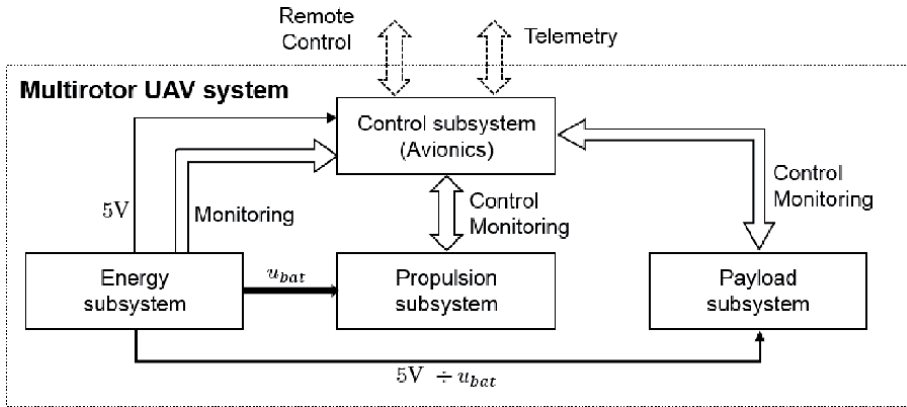
aircraft. They are mathematically described by a dynamic model of a rigid body with six second-order differential equations, twelve state variables, and  $N$  input variables, making them a multivariable system. Conventional configurations are characterized by a planar arrangement of the rotors, where the typical and most common configurations with four (quadrotor), six (hexarotor), and eight (octorotor) rotors are shown in **Figure 3**. The propulsion units mainly consist of an electric motor, suitable motor driver, and fixed-pitch propeller. It follows that the rotor's angular velocities are the only variables that have a direct impact on flight dynamics since propellers by their rotation create aerodynamic forces and moments.

The development and design of multirotor UAVs are significantly limited by both their size and energy consumption. For simpler analysis, design, and construction of this type of UAV, the aircraft system is divided into four key subsystems regardless of the configuration or purpose of the aircraft. The propulsion subsystem consists of  $N$  rotors that generate the necessary forces and moments for the movement of the aircraft in 3D space. The energy subsystem consists of one or more lithium-polymer batteries with joined components that need to deliver a large amount of energy essential to achieve the desired performance. The control subsystem takes care of UAV navigation and the functioning of the overall system by managing and monitoring other subsystems. Another task for the control subsystem is to be the interface between the aircraft and the base station (docking facility). The multirotor payload subsystem includes all equipment and cargo required to perform a particular mission whereby, this paper discusses the missions of heavy lift cargo transportation. Generally, it can be said that the performance of the multirotor UAV is determined by the parameters and components of the propulsion and energy subsystem. **Figure 4** schematically shows the multirotor UAV system which consists of four key subsystems.

From a control point of view, multirotor UAVs are inherently unstable and highly nonlinear systems. The inherent instability stems from the fact that this type of UAV cannot return to the equilibrium point on its own if it loses the functionality of the control loops. Furthermore, multirotor UAVs are highly nonlinear systems since the propulsion aerodynamic forces and moments are proportional to the square of the rotor angular velocity, and the transformations of the coordinate systems involve trigonometric functions. The basic task of the control subsystem is to navigate the multirotor UAV according to the given mission. The considered control subsystem is based on PX4 Autopilot. Generated control signals by the PX4 Autopilot, are sent to the electric propulsion units in order to achieve the desired movement in 3D space, ie to perform the mission. Orientation sensors are integrated into the PX4 Autopilot, while the position is estimated using compatible peripheral sensors. In considered concept, a global positioning system (GPS) and sensors for precise docking are incorporated in the control system. Given that the cargo mass in transport missions is unpredictable and there are real external disturbances such as wind gusts, it is necessary to consider robust control algorithms.



**Figure 3.**  
*Conventional multirotor UAV configurations [24].*



**Figure 4.** Schematic representation of an electric multirotor UAV system.

### 3.1 Electric propulsion subsystem

The propulsion subsystem provides the aircraft system with the necessary power to move in 3D space. The choice of propulsion subsystem configuration and propulsion unit type affects flight performance and it is a key step in designing a multirotor type of UAV. Electric motor based propulsion unit enables precise and fast control of forces and moments which directly affect the position and orientation of the aircraft. The reliability of electrical systems reduces the possibility of aircraft crash due to motor failure. The required performance of the aircraft, which depends on the type and profile of the mission, determines the choice of propulsion configuration and components. The electric propulsion unit consists of a control unit and a mechanical assembly of the motor on whose rotor a fixed-pitch propeller is mounted, which creates forces and moments by its rotation. Propulsion units with brushless DC (BLDC) motors coupled with electronic speed controllers (ESCs) are suitable for a wide range of tasks, including missions with heavy lift transportation (**Figure 5**).

In this case study, seven setups of electric propulsion units were considered which will be paired with the high voltage (HV) setup of the energy subsystem. Based on the component manufacturer's specification, it is possible to characterize the propulsion units which is the first step in designing the aircraft propulsion subsystem for heavy lift multirotor system. **Table 2** shows the considered electric propulsion components for which characterization is presented. Motor velocity constant (back EMF constant)  $K_v$  of motors intended for heavy payloads is typically small ( $K_v < 200$ ), resulting in lower speeds and higher torques. The propeller designations in the table describe its geometry where the first two numbers indicate the diameter of the propeller, and the other two the propeller pitch, both in inch.



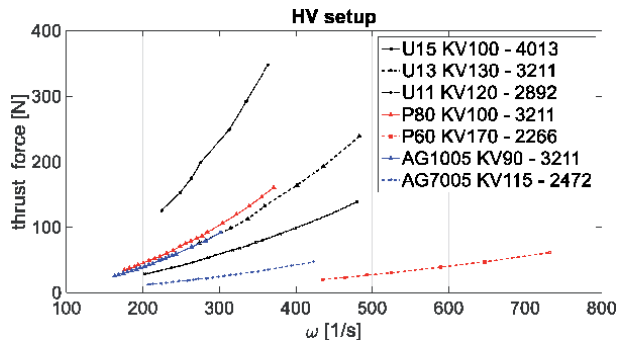
**Figure 5.** Multirotor UAV electric propulsion unit.



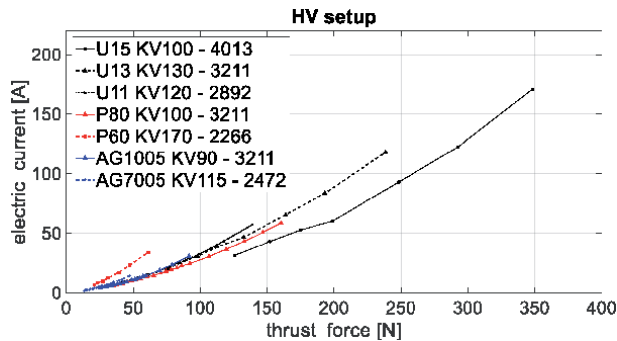
Characterization is necessary for the appropriate selection of propulsion components and analysis of the electric propulsion system which further allows system optimization. **Figure 6** shows thrust force as a function of rotor angular velocity for seven considered setups. As expected, setups consisting of motors with a lower Kv achieve lower angular velocities, and propellers with larger diameter achieve higher thrust forces. **Figure 7** shows the electric current as a function of the required thrust force, which is very important from the aspect of estimating the flight time. **Figure 8** shows the overall efficiency of the propulsion unit which is expressed by the ratio of thrust and electric power. In this case, efficiency is represented as a function of electric power. From the graph, it can be concluded that the increase in efficiency can generally be achieved by choosing propulsion units consisting of a larger diameter propeller combined with a suitable motor.

BLDC motor	Kv	ESC	Propeller
U15 II	100	FLAME 180A HV	4013 (40 × 13.1)
U13 II	130	ALPHA 120A HV	3211 (32 × 11)
U11 II	120	ALPHA 80A HV	2892 (28 × 9.2)
P80	100	ALPHA 80A HV	3211 (32 × 11)
P60	170	FLAME 60A HV	2266 (22x6.6)
Antigravity 1005	90	FLAME 60A HV	3211 (32 × 11)
Antigravity 7005	115	FLAME 60A HV	2472 (24 × 7.2)

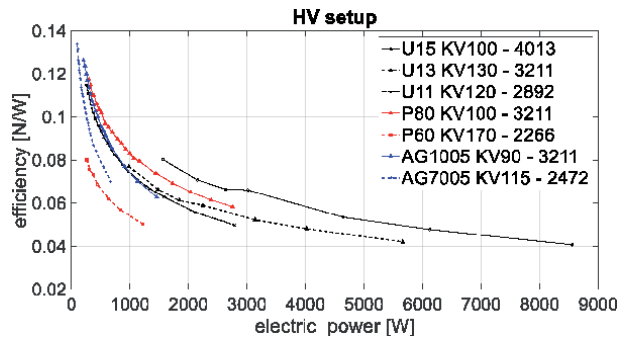
**Table 2.**  
 Considered electric propulsion unit setups [25].



**Figure 6.**  
 Thrust force with respect to rotor angular velocity.



**Figure 7.**  
 Electric current with respect to the thrust force.



**Figure 8.**  
The overall efficiency of the electric propulsion unit with respect to electric power.

### 3.2 Electric energy subsystem

The energy subsystem must provide sufficient energy for multirotor UAV system in order to perform the intended missions. Multirotor UAVs are characterized by high energy consumption as they consist of a minimum of four propulsion units. When choosing an energy subsystem, it is necessary to take into account several parameters, the most important of which is the type of propulsion subsystem. Electric propulsion units based on BLDC motors are combined with an energy subsystem consisting of one or more lithium-polymer (LiPo) batteries. Each LiPo battery contains one or more electrochemical cells to ensure a continuous flow of energy to power the propulsion and other subsystems. A very important feature of a LiPo battery is high energy density. Compared to other types of batteries such as nickel-metal hydride (NiMH), LiPo batteries have a higher discharge rate, which allows more power and consistent energy flow to the propulsion. LiPo batteries are defined by the capacity, discharge rate (C), and the number of cells that determine the operating voltage (S). The nominal voltage of a single battery cell is 3.7 V, and the voltage of a fully charged cell is 4.2 V.

When selecting batteries, the energy requirements of the propulsion subsystem must be taken into account, which in turn depends on the mass and size of the aircraft and the number of propulsion units. It follows that when designing a system, the relationship between mass and battery capacity is one of the



**Figure 9.**  
LiPo battery Tattu plus 12S, 22,000 mAh, 25C.



**Figure 10.**  
*Power hub MAUCH power cube 4.*

key data. For the characterized propulsion units, commercial high voltage (12S) LiPo batteries of capacity 8000 mAh, 16,000 mAh, and 22,000 mAh, as shown in **Figure 9**, were selected for further analysis [26]. Since the mass of the batteries in relation to other system masses dominantly affects the aircraft dynamics, it is desirable to place the energy subsystem centrally, as close as possible to the aircraft center of gravity. Energy subsystems of large aircraft, such as in this research, have more sophisticated energy distribution circuits (**Figure 10**) that provide different voltage levels and also have the function of measuring the battery's electrical parameters.

#### **4. Design considerations for heavy lift multirotor UAV**

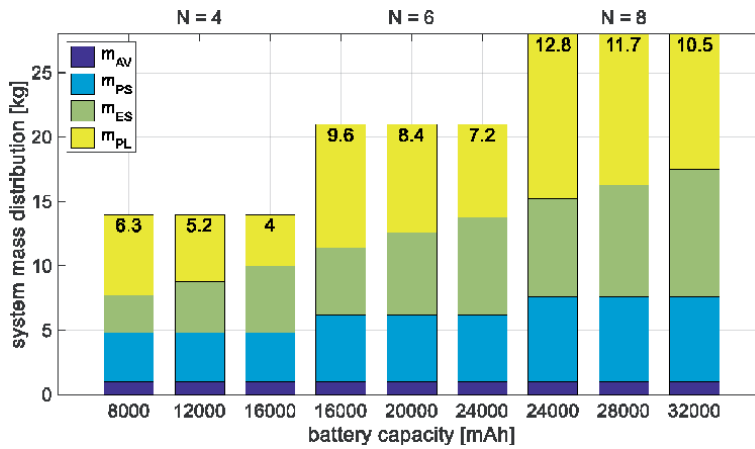
To ensure overall flight performance, it is necessary to determine the required thrust-to-weight ratio (TWR). As a rule, aircraft are designed with approximately twice the thrust force in comparison with aircraft weight, from which it can be concluded that the mass of the aircraft is a key parameter in the design of the system. The division of the multirotor UAV system into four key subsystems, where the subsystems are determined by masses as the basic parameters, represents the first step in the design of the aircraft. The equipment subsystem is directly determined by the type of mission which in this research is heavy lift transportation. The mass of this subsystem directly affects the selection of the propulsion and energy subsystem. These two subsystems are interdependent and when choosing components it is necessary to maintain a balance with the existing constraints defined by the mission. Parameters which significantly influence the dynamics and duration of the flight were analyzed for five selected aircraft series. The design of the propulsion subsystem affects the performance of the aircraft. Given the interdependence of the propulsion and energy subsystems, the parameters of LiPo batteries are included in the analysis.

##### **4.1 Mass distribution of the aircraft system**

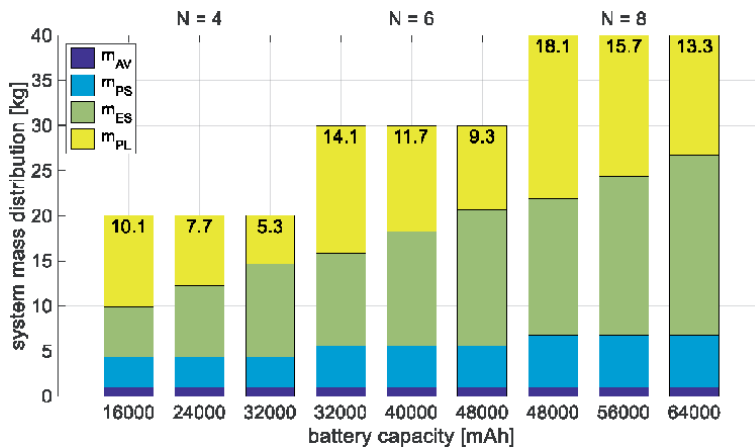
Based on the division of the aircraft system into four key subsystems, the mass distributions of conventional multirotor configurations for five selected aircraft series (S, M, L, XL, XXL) are graphically shown. To ensure basic flight performance, the required TWR was determined according to the propulsion manufacturer's recommendations and is approximately 1.8 for all five series. **Figures 11–15** show the mass distributions of payload mass ( $m_{PL}$ ), energy subsystem mass ( $m_{ES}$ ),

propulsion subsystem mass ( $m_{PS}$ ), and control subsystem (avionics) mass ( $m_{AV}$ ) for five generic series of aircraft whose propulsion subsystem consists of four, six, and eight rotors. For the selected battery capacities and the number of rotors, the maximum masses of the payload subsystems for the assumed TWR are expressed in kilograms. The first aircraft series – S (**Figure 11**) whose propulsion subsystem is based on propulsion units with P60 BLDC motors, is considered for payloads up to 10 kg. The second series – M (**Figure 12**) based on propulsion units with Antigravity 1005 motors, is considered for payloads from 10 kg up to 15 kg. The third series – L (**Figure 13**) based on propulsion units with P80 motors, is considered for payloads from 15 kg up to 25 kg. The fourth series – XL (**Figure 14**) based on propulsion units with U13 motors, is considered for payloads from 25 kg up to 50 kg. And lastly, the fifth series – XXL (**Figure 15**) based on propulsion units with U15 motors, is considered for payloads from 50 kg up to 100 kg.

The analysis of the multirotor UAV system mass distribution and the graphical representation, shown in **Figures 11–15**, was performed using a script written in the Matlab software package.



**Figure 11.** Mass distribution of the S aircraft series system (TWR = 1.8).



**Figure 12.** Mass distribution of the M aircraft series system (TWR = 1.8).

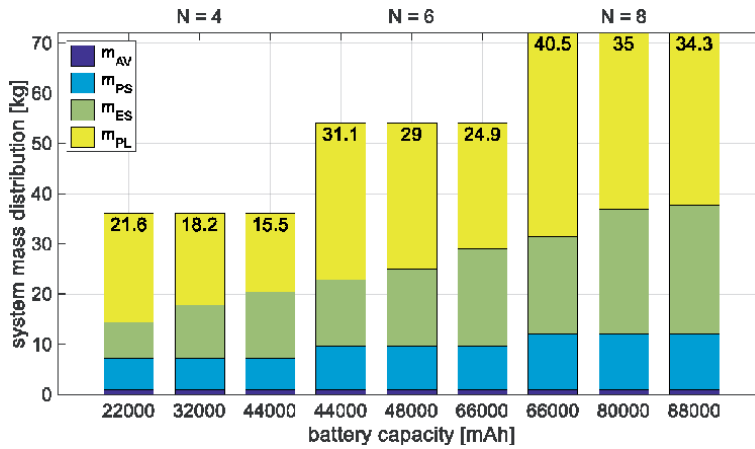


Figure 13.  
 Mass distribution of the L aircraft series system ( $TWR = 1.8$ ).

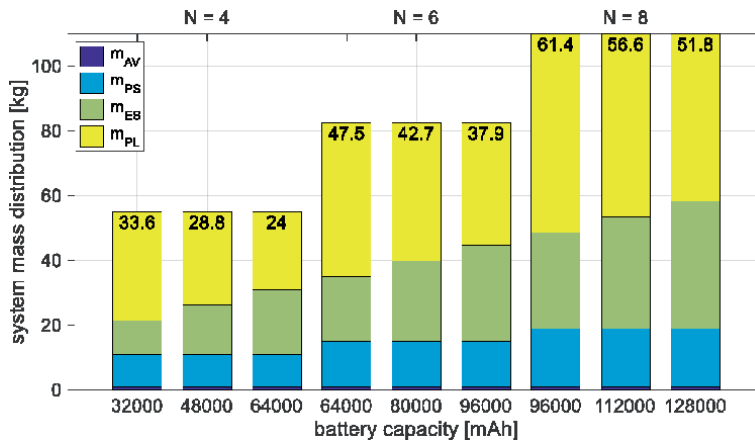


Figure 14.  
 Mass distribution of the XL aircraft series system ( $TWR = 1.8$ ).

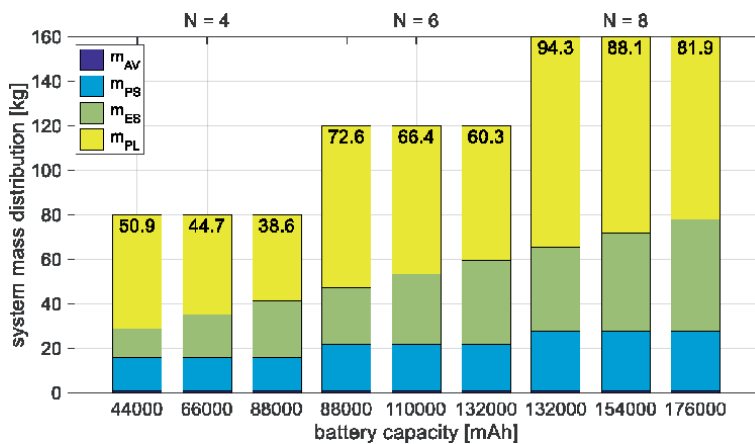
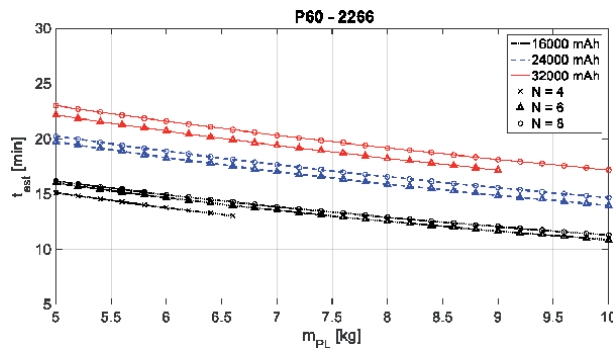


Figure 15.  
 Mass distribution of the XXL aircraft series system ( $TWR = 1.8$ ).

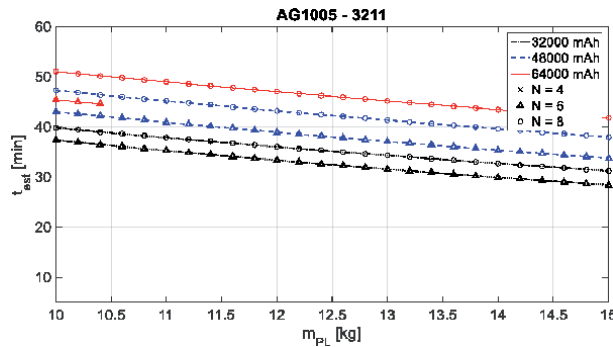
### 4.2 Numerical estimation of stationary flight time

Very important information is the total time that the aircraft can be in the air, which depends on the mission itself, ie on the required flight performance and the cargo that the aircraft carries. Based on the characteristics of the propulsion units, an estimate of the flight time was performed for selected series of aircraft defined by the parameters of the propulsion and energy subsystem. The basic case is considered, a stationary flight of conventional configurations, assuming that the drop in battery voltage and the power consumption by the control subsystem (possibly also the payload subsystem) are ignored. The estimated flight time is calculated based on the available battery capacity and the electric current required to generate adequate thrust force. It is also important to note that the complete battery capacities were used in the calculation, which is not possible in practice since the batteries must not be completely discharged. The required thrust force to reach the steady state of the aircraft depends on the mass of the system.

**Figure 16** shows the estimated stationary flight time for possible nine configurations of the first aircraft series – S. Configurations for each series are determined by the number of rotors ( $N$ ) which is a parameter of the propulsion subsystem, and by battery capacity ie the number of considered batteries which is a parameter of the energy subsystem. The first series (S), as mentioned in the last subsection, is considered for payloads up to 10 kg. Furthermore, **Figures 16–19** show the estimated stationary flight time for other aircraft series (M, L, XL, and XXL) which were considered and analyzed in the last subsection.

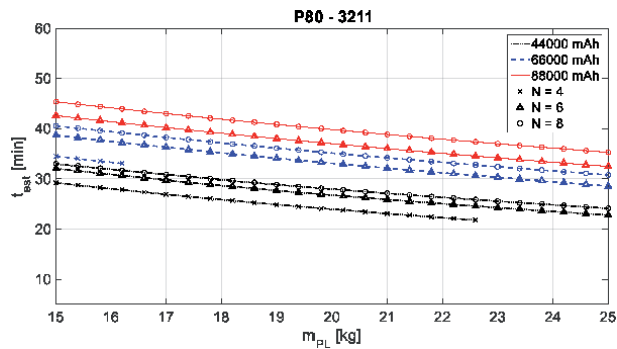


**Figure 16.**  
*Estimated stationary flight time of the S aircraft series.*

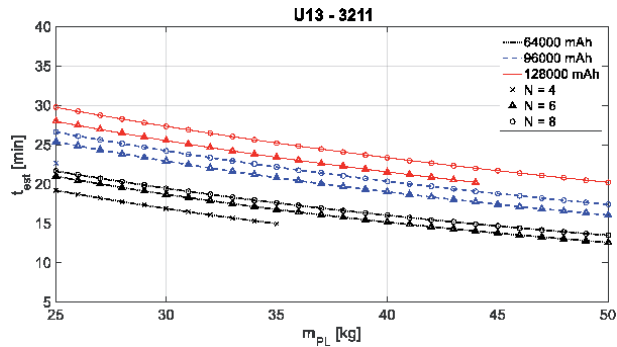


**Figure 17.**  
*Estimated stationary flight time of the M aircraft series.*

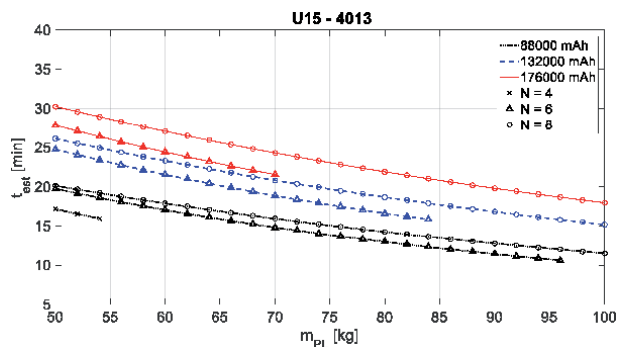
The estimation of the multirotor UAV system stationary flight time and the graphical representation, shown in **Figures 16–20**, based on the parameters of the propulsion and energy subsystems, was performed using a script written in the Matlab software package.



**Figure 18.**  
*Estimated stationary flight time of the L aircraft series.*



**Figure 19.**  
*Estimated stationary flight time of the XL aircraft series.*



**Figure 20.**  
*Estimated stationary flight time of the XXL aircraft series.*

### 4.3 Parametric design of a multirotor UAV

Based on the analysis, it is possible to select the aircraft parameters and components for a particular cargo and planned flight route. In considered design, the propulsion

unit (rotor) is defined by the propeller diameter and the propulsion subsystem by the number of rotors. The size of the aircraft, which is defined by the aircraft diameter, derives from these two parameters. For given propulsion units defined with propeller diameter ( $d$ ) in inch, and with the number of rotors ( $N$ ), **Table 3** shows the propulsion subsystem diameter ( $D$ ) in mm, which actually defines the construction parameters of the aircraft. In order to reduce the cost of prototyping and the potential production of system parts, it is necessary to achieve a certain degree of modularity. The idea is to turn predefined aircraft subsystems into modules that can be easily connected to each other. Of particular importance is the modularity of the propulsion subsystem. One of the ways to achieve this goal is by parameterizing the propulsion construction that connects the propulsion components with other subsystems.

	$d = 22''$	$d = 24''$	$d = 28''$	$d = 32''$	$d = 40''$
$N = 4$	1100	1200	1400	1600	2000
$N = 6$	1400	1600	1800	2000	2500
$N = 8$	1800	2000	2200	2500	3000

**Table 3.**  
Conventional multirotor configuration sizes ( $D$ ).

## 5. Conclusion

In this chapter, the concept of on-demand autonomous cargo transportation is presented by employing multirotor UAVs. The topology of a network of docking facilities and a multirotor UAV fleet was proposed for the considered autonomous transport within the central Adriatic islands in Croatia. Currently available technologies allow relatively simple implementation of the proposed concept, however, regulations applicable in a particular area should also be considered in the future. The market of commercially available components for propulsion and energy subsystem was investigated and it was found that it is possible to develop multirotor UAVs that can carry up to 100 kg of cargo. Based on the considered setups of electric propulsion units, the analysis of the Multirotor UAV propulsion and energy subsystems parameters were performed. In the multirotor UAV design process, the mass of five series of aircraft is considered and the flight time with respect to the payload mass is approximated and shown. The proposed concept of an autonomous cargo transportation system has great potential for future development and implementation since it can reduce transport costs, increase the frequency and speed of transport, and reduce carbon emissions.

## Acknowledgements

This research was funded by European Regional Development Fund, Operational programme competitiveness and cohesion 2014-2020, as part of the call for proposals entitled “Investing in science and innovation – first call”, grant number KK.01.1.1.04.0092.

## Appendices and Nomenclature

UAV	unmanned aerial vehicle
MEMS	micro-electromechanical systems



UGV	unmanned ground vehicle
VTOL	vertically take-off and land
ACTS	autonomous cargo transportation system
GPS	global positioning system
BLDC	brushless direct current
ESC	electronic speed controller
HV	high voltage
EMF	electromotive force
LiPo	lithium-polymer
NiMh	nickel-metal hydride
TWR	thrust-to-weight ratio

## Author details

Denis Kotarski<sup>1\*</sup>, Petar Piljek<sup>2</sup> and Josip Kasac<sup>2</sup>

1 Karlovac University of Applied Sciences, Karlovac, Croatia

2 Faculty of Mechanical Engineering and Naval Architecture, University of Zagreb, Zagreb, Croatia

\*Address all correspondence to: [denis.kotarski@vuka.hr](mailto:denis.kotarski@vuka.hr)

## IntechOpen

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Palacios AT, Cordero JM, Bello MR, Palacios ET, González JL. New Applications of 3D SLAM on Risk Management Using Unmanned Aerial Vehicles in the Construction Industry. In: Dekoulis G, editors. *Drones-Applications*. London: IntechOpen; 2018. p. 97-118. DOI: 10.5772/intechopen.73325
- [2] Li Y, Liu C. Applications of multirotor drone technologies in construction management. *International Journal of Construction Management*. 2018;19(5):401-412. DOI: 10.1080/15623599.2018.1452101
- [3] Yinka-Banjo C, Ajayi O. Sky-Farmers: Applications of Unmanned Aerial Vehicles (UAV) in Agriculture. In: Dekoulis G, editors. *Autonomous Vehicles*. London: IntechOpen; 2020. DOI: 10.5772/intechopen.89488
- [4] Kanistras K, Martins G, Rutherford MJ, Valavanis KP. Survey of Unmanned Aerial Vehicles (UAVs) for Traffic Monitoring. In: Valavanis K, Vachtsevanos G, editors. *Handbook of Unmanned Aerial Vehicles*. Dordrecht: Springer; 2015. p. 2643-2666. DOI: 10.1007/978-90-481-9707-1\_122
- [5] Półka M, Ptak S, Kuziora Ł, Kuczyńska A. The Use of Unmanned Aerial Vehicles by Urban Search and Rescue Groups. In: Dekoulis G, editors. *Drones-Applications*. London: IntechOpen; 2018. p. 83-96. DOI: 10.5772/intechopen.73325
- [6] Briod A, Kornatowski P, Zufferey JC, Floreano D. A Collision-resilient Flying Robot. *Journal of Field Robotics*. 2014;31(4):496-509. DOI: 10.1002/rob.21495
- [7] Yuan C, Liu Z, Zhang Y. Aerial Images-Based Forest Fire Detection for Firefighting Using Optical Remote Sensing Techniques and Unmanned Aerial Vehicles. *Journal of Intelligent & Robotic Systems*. 2017;88:635-654. DOI: 10.1007/s10846-016-0464-7
- [8] Imdoukh A, Shaker A, Al-Toukhy A, Kablaoui D, El-Abd M. Semi-autonomous indoor firefighting UAV. In: 18th International Conference on Advanced Robotics (ICAR '17); 10-12 July 2017; Hong Kong, China: IEEE; 2017. p. 310-315. DOI: 10.1109/ICAR.2017.8023625
- [9] Otero Arenzana A, Escribano Macias JJ, Angeloudis P. Design of Hospital Delivery Networks Using Unmanned Aerial Vehicles. *Transportation Research Record*. 2020;2674(5):405-418. DOI: 10.1177/0361198120915891
- [10] Escobar-Ruiz AG, Lopez-Botello O, Reyes-Osorio L, Zambrano-Robledo P, Amezcua-Brooks L, Garcia-Salazar O. Conceptual Design of an Unmanned Fixed-Wing Aerial Vehicle Based on Alternative Energy. *International Journal of Aerospace Engineering*. 2019;2019:1-13. DOI: 10.1155/2019/8104927
- [11] Rotaru C, Todorov M. Helicopter Flight Physics. In: Volkov K, editors. *Flight Physics*. London: IntechOpen; 2018. p. 19-48. DOI: 10.5772/intechopen.71516
- [12] Kotarski D, Kasać J. Generalized Control Allocation Scheme for Multirotor Type of UAVs. In: Dekoulis G, editors. *Drones-Applications*. London: IntechOpen; 2018. p. 43-58. DOI: 10.5772/intechopen.73006
- [13] Öznalbant Z, Kavsaoğlu MŞ. Flight control and flight experiments of a tilt-propeller VTOL UAV. *Transactions of the Institute of Measurement and*

Control. 2018;40(8):2454-2465. DOI: 10.1177/0142331218754618

[14] Benić Z, Piljek P, Kotarski D. Mathematical modelling of unmanned aerial vehicles with four rotors. *Interdisciplinary Description of Complex Systems*. 2016;14(1):88-100. DOI: 10.7906/indecs.14.1.9

[15] Jannoura R, Brinkmann K, Uteau D, Bruns C, Joergensen RG. Monitoring of crop biomass using true colour aerial photographs taken from a remote controlled hexacopter. *Biosystems Engineering*. 2015;129:341-351. DOI: 10.1016/j.biosystemseng.2014.11.007

[16] Ikeda T, Yasui S, Fujihara M, Ohara K, Ashizawa S, Ichikawa A, Okino A, Oomichi T, Fukuda T. Wall contact by Octo-rotor UAV with one DoF manipulator for bridge inspection. In: *IEEE International Conference on Intelligent Robots and Systems (IROS '17)*; 24-28 September 2017; Vancouver. Canada: IEEE; 2017. p. 5122-5127. DOI: 10.1109/IROS.2017.8206398

[17] Driessens S, Pounds PE. The Triangular Quadrotor: A More Efficient Quadrotor Configuration. *IEEE Transactions on Robotics*. 2015;31(6):1517-1526. DOI: 10.1109/TRO.2015.2479877

[18] Ryll M, Muscio G, Pierrri F, Cataldi E, Antonelli G, Caccavale F, Franchi A. 6D Physical Interaction with a Fully Actuated Aerial Robot. In: *IEEE International Conference on Robotics and Automation (ICRA '17)*; 29 May-3 June 2017; Singapore. USA: IEEE; 2017. p. 5190-5195. DOI: 10.1109/ICRA.2017.7989608

[19] Brar S, Rabbat R, Raithatha V, Runcie G, Yu A: Drones for Deliveries. Technical Report. 2015. Available from: <http://scet.berkeley.edu/wp-content/uploads/ConnCarProjectReport-1.pdf> [Accessed: 2020-10-16]

[20] Villa DKD, Brandão AS, Sarcinelli-Filho M. A Survey on Load Transportation Using Multirotor UAVs. *Journal of Intelligent & Robotic Systems*. 2020;98:267-296. DOI: 10.1007/s10846-019-01088-w

[21] Ong W, Srigrarom S, Hesse H. Design Methodology for Heavy-Lift Unmanned Aerial Vehicles with Coaxial Rotors. In *AIAA SciTech Forum*; 7-11 January 2019; San Diego. USA: American Institute of Aeronautics and Astronautics; 2019. p. 7-11. DOI: 10.2514/6.2019-2095

[22] DRONE VOLT. Hercules 20 Heavy lift drone [Internet]. Available from: <https://www.dronevolt.com/en/expert-solutions/hercules-20/> [Accessed: 2020-10-20]

[23] Haris Areal. Carrier HX8 Sprayer Drone [Internet]. Available from: <https://www.harrisaerial.com/carrier-hx8-sprayer/> [Accessed: 2020-10-20]

[24] PX4 Autopilot User Guide. Airframes Reference [Internet]. Available from: [https://docs.px4.io/v1.9.0/en/airframes/airframe\\_reference.html](https://docs.px4.io/v1.9.0/en/airframes/airframe_reference.html) [Accessed: 2020-10-26]

[25] Electric Propulsion Components. Available from: <https://store-en.tmotor.com/category.php?id=11> [Accessed: 2020-10-17]

[26] Tattu Lipo Battery Pack. Available from: <https://www.gensace.de/tattu-plus> [Accessed: 2020-10-19]



# Novel Application of Fast Simulated Annealing Method in Brushless Motor Drive (BLMD) Dynamical Parameter Identification for Electric Vehicle Propulsion

*Richard A. Guinee*

## Abstract

Permanent magnet brushless motor drives (BLMD) are extensively used in electric vehicle (EV) propulsion systems because of their high power and torque to weight ratio, virtually maintenance free operation with precision control of torque, speed and position. An accurate dynamical parameter identification strategy is an essential feature in the adaptive control of such BLMD-EV systems where sensorless current feedback is employed for reliable torque control, with multi-modal penalty cost surfaces, in EV high performance tracking and target ranging. Application of the classical Powell Conjugate Direction optimization method is first discussed and its inaccuracy in dynamical parameter identification is illustrated for multimodal cost surfaces. This is used for comparison with the more accurate Fast Simulated Annealing/Diffusion (FSD) method, presented here, in terms of the returned parameter estimates. Details of the FSD development and application to the BLMD parameter estimation problem based on the minimum quantized parameter step sizes from noise considerations are provided. The accuracy of global parameter convergence estimates returned, cost function evaluation and the algorithm run time are presented. Validation of the FSD identification strategy is provided by excellent correlation of BLMD model simulation trace coherence with experimental test data at the optimal estimates and from cost surface simulation.

**Keywords:** electric vehicle propulsion systems, Permanent magnet brushless motor drives, dynamical parameter identification, sensorless current feedback, multimodal cost surfaces, Powell Conjugate Direction optimization, Simulated Annealing, Fast Simulated Diffusion, quantized parameters, cost surface noise, parameter estimation accuracy

## 1. Introduction

High performance permanent magnet brushless motor drive (BLMD) systems are now widely used in electric vehicle (EV) propulsion [1–3], because of their

higher power factor and efficiency, and are central to modern industrial automation [4] in such scenarios as aerospace systems control and maneuverability, numerical control (NC) machine tools and robotics. The benefits accruing [1, 3, 5, 6] from the application of such servodrives over other electric motor systems are higher power and better torque to weight ratio, a considerable saving of energy and higher precision control of torque, speed and position which promote better electric vehicle propulsion performance and optimal EV target ranging along with automation and control. This is due largely to the high torque-to-weight ratio and compactness of permanent magnet (PM) drives, lower heat dissipation and the virtually maintenance free operation of brushless motors in inaccessible locations when compared to conventional DC & AC electric motors. The controllers of these machine drives, which incorporate wide bandwidth speed and torque control loops, are adaptively tuned to meet the essential requirements of system robustness, high tracking performance and EV range acquisition without overstressing the hardware components [3, 7, 8]. An essential feature of the adaptive optimal control procedure in such BLMD-EV systems is the accurate dynamical parameter identification strategy used in conjunction with an accurate BLMD model [9–12], where sensorless feedback current control loops are employed for reliable torque control during EV propulsion, with multi-modal penalty cost surfaces.

This book chapter, which is divided into three separate but interrelated sections, concerns BLMD dynamical parameter identification in which the choice of target data used in the Minimum Squared Error (MSE) objective function formulation has a significant effect on cost surface topology and selectivity in the vicinity of the global minimum [13, 14]. The penalty cost function selection in the identification of the motor drive dynamics impacts directly on the type of parameter search strategy to be adopted where embedded cost surface multim minima are concerned, which is discussed below, in terms of the accuracy of the returned parameter estimates. A short review of classical identification techniques is provided initially with an explanation of global convergence failure due to local minimum trapping over a multim minima cost surface based on BLMD step response feedback current (FC) target data. This difficulty with classical optimization methods highlights the need for an effective search strategy of parameter space with an in-built adaptive jump mechanism which can facilitate escape from possible local minimum capture and guarantee eventual global convergence. Such identification search features are provided by statistical optimization methods based on a Simulated Annealing (SA) kernel [15–21] which has an adjustable pseudo-temperature parameter that controls the jump related magnitude of the inherent random fluctuations. The Fast Simulated Diffusion algorithm [22–24], which is a more efficient version of SA that exploits the use of the classical gradient search technique at low pseudo-temperatures, surmounts the above obstacles and can be deployed as an effective global parameter extraction tool for BLMD identification.

In the first section the classical application of the Powell Conjugate Direction (PCD) optimization technique of motor parameter extraction over a two dimensional shaft velocity cost surface, which has a parabolic wedge shaped topography with a ‘line minimum’ stationary region [13, 14], is discussed. This will be used later for comparison purposes with the more successful Fast Simulated Diffusion (FSD) method, which is presented in the second section, to show that the FSD method is better in terms of the accuracy of returned FSD parameter estimates. It will also be shown that the PCD method does not converge to the global minimum and that the returned parameter estimates are less than satisfactory in the approximation of the optimal parameter vector when contrasted with the FSD method deployed over a multim minima FC cost surface. The convergence inaccuracy of the PCD method lays the groundwork for the introduction of a more accurate and efficient parameter

identification search strategy based on the FSD method over a multim minima cost surface that is more selective at the global minimum.

In the second section the application of the FSD optimization search technique as an extension of the SA method is presented, with a multim minima objective function based on step response FC target data, for motor dynamical parameter extraction. Details of the FSD development and integration into the BLMD parameter estimation problem based on the minimum quantized step sizes from noise considerations are provided. The results of global parameter convergence estimates returned, which are very accurate, including the number of iterations, cost function evaluation and the algorithm run time are presented. Validation of the FSD identification strategy is provided by excellent correlation of BLMD model simulation trace coherence, at the optimal parameter estimates, with experimental test data and from cost surface simulation.

In the third section the development details of a novel modified form of the FSD algorithm are presented for fast accurate parameter identification [22, 23, 25–28]. This will be based on the heuristics of the FC sinc-like surface topography in an effort to reduce the computational cost and search time to acquire global optimality. The beneficial effect of the incorporation of an approximated gradient search, along with the gathered cost statistics, at the end of each anneal step in the cooling schedule during the high temperature random search phase is discussed. Furthermore the condition, pertaining to the occurrence of random search trapping within the capture cross section of the cost surface containing the global extremum, for elimination of the reheat and thermal condensation phases of the FSD method is explained. Details of a set of tests to establish modified FSD (MFSD) performance in terms of convergence accuracy accompanied with a substantial reduction in search time, for three known cases of motor shaft load inertia ( $J$ ) in the parameter identification process, are furnished with remote initialization in the vicinity of a potential local minimum trap. An error analysis of the returned parameter estimates is provided, which are almost identical to actual  $J$  values, and a comparison is made with alternative global estimates obtained from cost surface simulation to validate the modified FSD search technique. A theoretical analysis of the effect of FC data training record length on the probability of global minimum capture and capture cross sectional ‘area’ is provided along with a discussion of the impact of data record length on FC cost surface selectivity for SI purposes.

## 2. Motivation

The benefits accruing from the use of an embedded BLMD system in high performance adjustable speed drive applications has resulted in a need for an accurate physical model [12, 29] for the purpose of parameter identification of the drive dynamics. Allied to this need is a general requirement for an accurate and efficient search strategy of parameter space in the design of optimal drive controllers [30–32] where system identification is an implicit feature during online operation. This is necessary for PID auto-tuning of wide bandwidth current loops in torque control mode to speed up embedded BLMD commissioning and facilitate control optimization through regular retuning. The identification process is dependent on an accurate model of the nonlinear electromechanical system [33] which includes the pulse width modulated (PWM) inverter with power transistor turnon delay to avoid current shoot-through.

The extraction of the drive dynamical parameters generally relies on the minimization of some quantitative measure of error cost in terms of the goodness-of-fit, based on the MSE norm [13], between the observed motor drive output

experimental test data and its model equivalent. The presence of multim minima in the MSE penalty function, however, results in a large spread of parameter estimates about the global minimum with model accuracy and subsequent controller design performance very dependent on the minimization technique adopted and the initial search point chosen. The existence of a noisy cost function, resulting in 'false' local minima proliferation in the stationary region containing the global extremum [13, 28], depends on the numerical accuracy with which the PWM delayed inverter switching instants are resolved in the model simulation [13]. Furthermore the plurality of genuine local minima is governed by the choice of data training record used in the objective function formulation which in the case of step response feedback current (FC) has a sinc-like topography [13]. The use of a step input (i/p) as a test stimulus is motivated by the fact that in normal industrial applications in the online mode input command changes are generally sudden and step-like and are sufficient for persistent excitation of the BLMD system. The accompanying transients in the observed variables can then be used effectively for parameter identification of motor shaft viscous damping factor and inertia changes during normal operation.

Experimental data training sets, of the observed variables [13] for various known shaft load inertia, can be used during the parameter identification procedure with the corresponding BLMD model simulation runs to establish a parameter mean squares error cost surface as the objective function to be minimized. The deployment of step response winding current feedback as a target function is found to be particularly beneficial as it exhibits the frequency modulated (FM) characteristic of a constant amplitude swept frequency sinusoid. The FC response also has better overall cost surface selectivity, which improves with observed data record length, by comparison with shaft velocity target information [13]. The corresponding simulated step response, upon parameter convergence, will be shown to have both frequency and phase coherence which attests to the accuracy of the identification methodology. Furthermore the selection of an FC multim minima cost function as a suitable choice for investigation in motor parameter extraction is motivated by the scenario, as a secondary consideration, that winding current flow information will be the only feedback signal available for sensorless motor control [34] applications.

The step response FC cost function exhibits an apparently smooth continuous one dimensional surface when plotted against either inertia or the damping factor as the free parameter to be identified [13] for coarse step size simulation. The local minimisers appear to be symmetrically disposed about the global minimum in accordance with a sinc function profile. The attendant unimodal velocity cost function appears to be parabolic in the inertia and friction parameter variables [13]. When high resolution of the motor shaft inertia estimates is required, in simulation trials, during system identification both cost functions appear to possess a granulated response surface thus rendering uncertainty in the parameter extraction process [13]. However this difficulty in parameter convergence is somewhat tempered by the observation that the noise floor in the error response surface is eclipsed by the residual error magnitude at the global minimum, which ameliorates the parameter search process, and is alleviated by the adoption of parameter quantization as discussed in [13, 28]. The response surface 'noisiness' arises primarily from the delay nonlinearity in the PWM current loops of the BLMD model structure and the accuracy with which the PWM crossover times are determined with the subsequent timing of the delayed inverter firing signals [9, 10]. The presence of 'genuine' local minimisers in the FC response cost surface is manifested through interference by the relative phasing of the swept frequency motor current sinusoids with target data at different inertia values [13, 28].



### 3. Choice of parameter identification methods

Two methods of parameter identification, which are based on the PCD and FSD optimization techniques and linked with the shape of the error response surface, will be investigated in sections 4 and 7 in the extraction process of known motor shaft inertia from data training records. The first, rooted in classical optimization techniques [35] where derivative information is not required [36, 37], relies on the application of Powell's conjugate direction search to the concave velocity response surface with single parameter variation [13]. This classical parameter extraction technique is superior to and much more efficient [35], due to the orthogonality of its conjugate directions of search, than other direct search methods [38–40] such as the Simplex method [41] and the method of Hooke and Jeeves [42, 43] which are expensive in CPU time and slow to converge. Other more efficient classical optimization techniques [35, 44–46], such as the Polak-Ribiere conjugate gradient method or the Newton-like BFGS method [47, 48] or the hybrid Levenberg–Marquardt method [49–51] are not considered in this case because of cost surface noisiness and the need to calculate partial derivatives which is difficult with the presence of 'noise' related PWM computational uncertainty. These alternative methods are computationally expensive in BLMD simulation time, where the evaluation of the gradient vector and Hessian matrix [35] are concerned, and are difficult to apply in any case because of the model complexity of the complete motor drive system with PWM inverter delay operation without resorting to difference equation approximation of the derivatives [35, 52]. Furthermore partial derivative evaluation is suspect in the presence of computational 'noise', inherent in the cost function, for an infinitesimal change in the relevant parameters. This can result in an erratic hill-descent, associated with the conjugate gradient direction search, over the response surface and entrapment of the minimization procedure in a false minimum in the noise grained incline of the bowl shaped velocity cost function [13, 28]. All classical minimization procedures are well known to have difficulty with the FC cost surface topography illustrated in [13, 28] because they are easily trapped in one of the embedded local minima and thus fail to converge to the optimal parameter set. This problem is accentuated by initialization of the search process remote from the global minimum, due to uncertainty, in a region where there are local minima. Classical hill-descent methods that rely on following the cost gradient are easily captured at local minima in this instance.

The second identification method is based on the statistical search technique of simulated diffusion (SD) [24, 53], as an adjunct of simulated thermal annealing (SA) [15–17] used in combinatorial optimization [25], and has to be deployed to acquire the global minimizer of the FC multim minima cost surface. The SD technique in contrast to the PCD method has been successfully applied to the multioptimal problem of MOSFET parameter extraction [19, 24] and simulated annealing in circuit placement tasks [20, 21], with a cost function having a fractal landscape, in the VLSI domain. Both the PCD and FSD methods of parameter extraction presented compare favorably in terms of the returned shaft load inertia, for short data records with initialization close to the global minimum, which enhances confidence in the model of the BLMD servodrive system as well as in the performance of both identification strategies. However the sensitivity of the FC objective function to inertia parameter changes is an order of magnitude greater than that of its shaft velocity counterpart [13, 28]. This results in better selectivity and more accurate convergence in the case of the SD method whereas Powell's method is more susceptible to capture in a false minimum due to cost surface noise in the vicinity of the global minimum. Furthermore with longer data records the selectivity of the FC cost function increases to a maximum, when the motor has reached full speed,

accompanied with ‘genuine’ local minima proliferation. The reverse effect is manifested as a flattening of the response surface in the neighborhood of the global minimum, with the resulting minimization procedure susceptible to remote trapping in noisy local minima, for the shaft velocity cost function. In the former procedure relative phase information can be used effectively for estimating parameter variation at motor speed saturation while the benefits of the initial speed response transient are attenuated in the latter case. The FSD search technique is verified, with initialization far from the global minimum, for known motor shaft inertia. The accuracy of the returned FSD estimates is also checked against the cost surface simulation. Convergence details and comparisons of both identification methods in terms of iteration count, objective function evaluations or motor simulation runs and CPU time are presented.

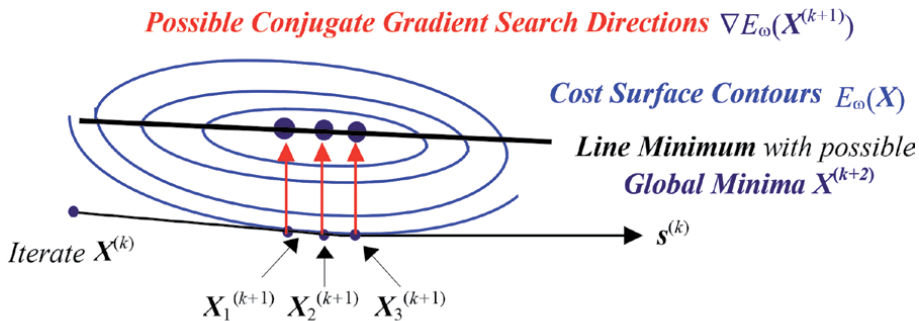
#### 4. Application of Powell’s identification method to velocity cost surface

The choice of Powell’s Conjugate Direction Set method [35, 49, 54, 55] of unconstrained optimization as a method of BLMD dynamical parameter identification within the indicated parameter tolerance bounds was motivated by topographical considerations of the shaft velocity objective function  $E_\omega(\mathbf{X})$  and ingrained cost surface noisiness. The penalty cost function concerned appears to have a ‘line minimum stationary region’ predominantly in the  $B$  parameter direction as elucidated in [13]. Consequently convergence difficulties can arise from application of any of the steepest decent conjugate gradient methods of parameter identification, such as the Polack-Ribiere technique [35], with this type of cost function. This problem results from the requirement that

$$\nabla E_\omega(\mathbf{X}^{(k+1)})^T \mathbf{s}^{(k)} < 0 \tag{1}$$

for hill decent and approaches zero as the global minimum is reached for the possible gradient search directions  $\nabla E_\omega(\mathbf{X}_j^{(k+1)})$  indicated in **Figure 1**.

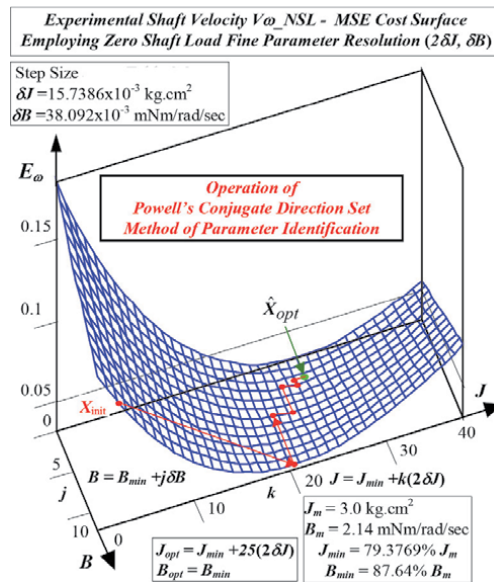
The component of the gradient  $\nabla E_\omega(\mathbf{X}^{(k+1)})$ , which is tangential to the cost contour in this instance, vanishes as shown along the  $k^{\text{th}}$  iterate search direction  $\mathbf{s}^{(k)}$  at the points  $\mathbf{X}_j^{(k+1)}$  for  $j = 1, 2$ , etc. This condition results in a multiplicity of search directions and possible global minimum values  $\mathbf{X}^{(k+2)}$  along the ‘line minimum’ of



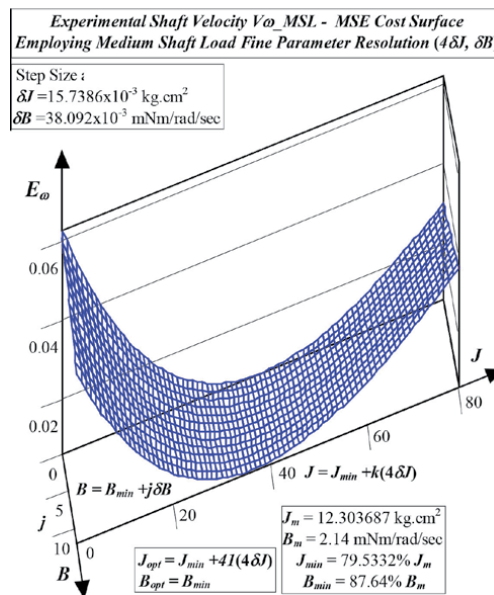
- Multiplicity  $j$  of Global Minima with Termination Points  $\mathbf{X}^{(k+2)}$
- Conjugate Direction  $\mathbf{s}^{(k)}$  is predominantly along  $B$ -Parameter Axis

**Figure 1.**  
Global minima multiplicity.

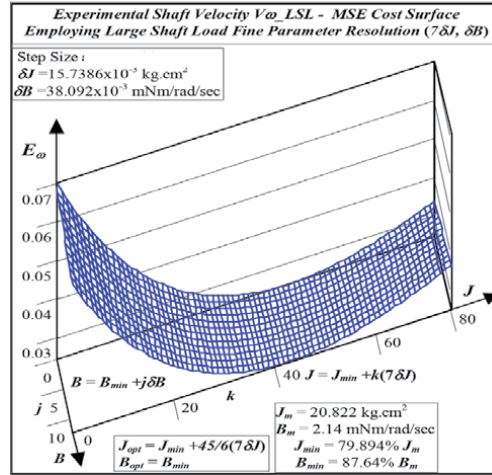
the wedge shaped syncline of the cost velocity surface as illustrated in **Figures 2–4**. This difficulty is partially borne out in [13] where the fitted response quadratic model parameters were re-evaluated at successive iteration points to improve the global convergence estimate. Furthermore the presence of point-like singularities due to cost surface noisiness, with ‘false’ local minima proliferation in the neighborhood of the global minimizer [13], results in a discontinuous cost function with consequent difficulties with derivative calculations. Also the derivative computation burden increases with expensive BLMD cost function evaluation where lengthy



**Figure 2.**  
 NSL velocity cost surface.



**Figure 3.**  
 MSL velocity cost surface.



**Figure 4.**  
LSL velocity cost surface.

simulation times are increased with a small time step  $\Delta t$  for accurate resolution of PWM edge transitions. The steepest decent gradient search technique can also suffer from oscillatory behavior and poor convergence results due to numerical round-off effects as reported by Fletcher [35].

The PCD method, which relies on cost function evaluations only without derivative information, is cyclically deployed instead as a line search algorithm in parameter space beginning at  $\mathbf{X}^{(1)}$ . This algorithm progresses along  $N$  mutually orthogonal directions conforming to the dimensionality of parameter space until sufficient accuracy of the global minimum estimate has been attained according to some convergence stopping criterion. With this method the iterate moves gradually towards the neighborhood of the global optimum  $\mathbf{X}_{opt}$  via inexact line searches initially and then rapidly converges to the stationary point itself. The iteration process is terminated upon some user supplied convergence test of the form [49].

$$2 \left\{ \left( E_{\omega}(\mathbf{X}^{(1)}) - E_{\omega}(\mathbf{X}^{(N+1)}) \right) \right\} / \left\{ E_{\omega}(\mathbf{X}^{(N+1)}) + E_{\omega}(\mathbf{X}^{(1)}) \right\} < \varepsilon \quad (2)$$

being satisfied with a gradual reduction in the cost  $E(\mathbf{X})$  towards an accumulation point  $\hat{\mathbf{X}}_{opt}$  which approximates the global minimum  $\mathbf{X}_{opt}$  within the specified error bound  $\varepsilon$ . Furthermore the application of a specific threshold step size  $\delta \mathbf{X}^L$  [13, 28] for parameter space quantization ameliorates the difficulty with response surface noisiness [13, 28]. This methodology results in inexpensive line searching during parameter extraction with a reasonable degree of convergence accuracy maintained. In the application of the PCD algorithm a quadratic model is used to approximate the two dimensional MSE objective function, pertaining to the shaft velocity in terms of the motor dynamical parameters [13], so that a prediction of the location of the local minimum can be made. This localized response surface modeling technique guarantees second order convergence [35] and is a very suitable candidate for the velocity cost function given the parabolic nature of its topography with respect to the inertia parameter which is the most likely to vary in high performance drive applications. The applied PCD method is based on the property of quadratic termination of the approximation model, with Hessian  $\hat{\mathbf{G}} > 0$ , at the global minimizer. This model strategy, which is similar to the normal form in [13, 28], admits to the existence of at most  $N$  line searches  $\{\mathbf{s}^{(1)}, \mathbf{s}^{(2)}, \dots, \mathbf{s}^{(N)}\}$  in

$N$ -dimensional parameter space  $\mathbf{X}$  for global convergence along independent mutually conjugate directions [35] such that

$$\mathbf{s}^{(i)T} \hat{\mathbf{G}} \mathbf{s}^{(j)} = \begin{cases} 1 & \forall i = j \\ 0 & \forall i \neq j \end{cases} \quad (3)$$

## 5. Application of PCD method to BLMD parameter extraction

The PCD method, which begins with a user supplied estimate  $\mathbf{X}^{(1)}$  and a set  $S$  of directions based on the column vectors of any orthogonal matrix such as the identity  $I$  with

$$S = \left\{ \mathbf{s}^{(k)} \mid \mathbf{s}^{(k)} = \mathbf{e}^{(k)} = [0, 0, \dots, 0, 1^{(k)}, 0, \dots, 0]^T, \forall k = 1, \dots, N \right\} \quad (4)$$

initially, proceeds according to a cyclic  $N$  dimensional search of parameter space as per the flowchart in **Figure 5** until some stopping criterion based on sufficient cost reduction is satisfied. The basic structure of the algorithm can be summarized as a line minimum search routine due to Brent [49, 56] for the  $k^{\text{th}}$  iterative search as

- Determine the search direction  $\mathbf{s}^{(k)}$  in discretized parameter space which is conjugate to all previous excursions into the quantized parameter lattice.
- Evaluate  $\alpha^{(k)}$  so as to minimize  $E_\omega(\mathbf{X}^{(k)} + \alpha^{(k)} \mathbf{s}^{(k)})$  with respect to  $\alpha$  in discrete space  $\mathbf{X}$  based on a quadratic model approximation of  $E_\omega(\mathbf{X})$  in this direction and proceed to set

$$\mathbf{X}^{(k+1)} = \mathbf{X}^{(k)} + \alpha^{(k)} \mathbf{s}^{(k)} \quad (5)$$

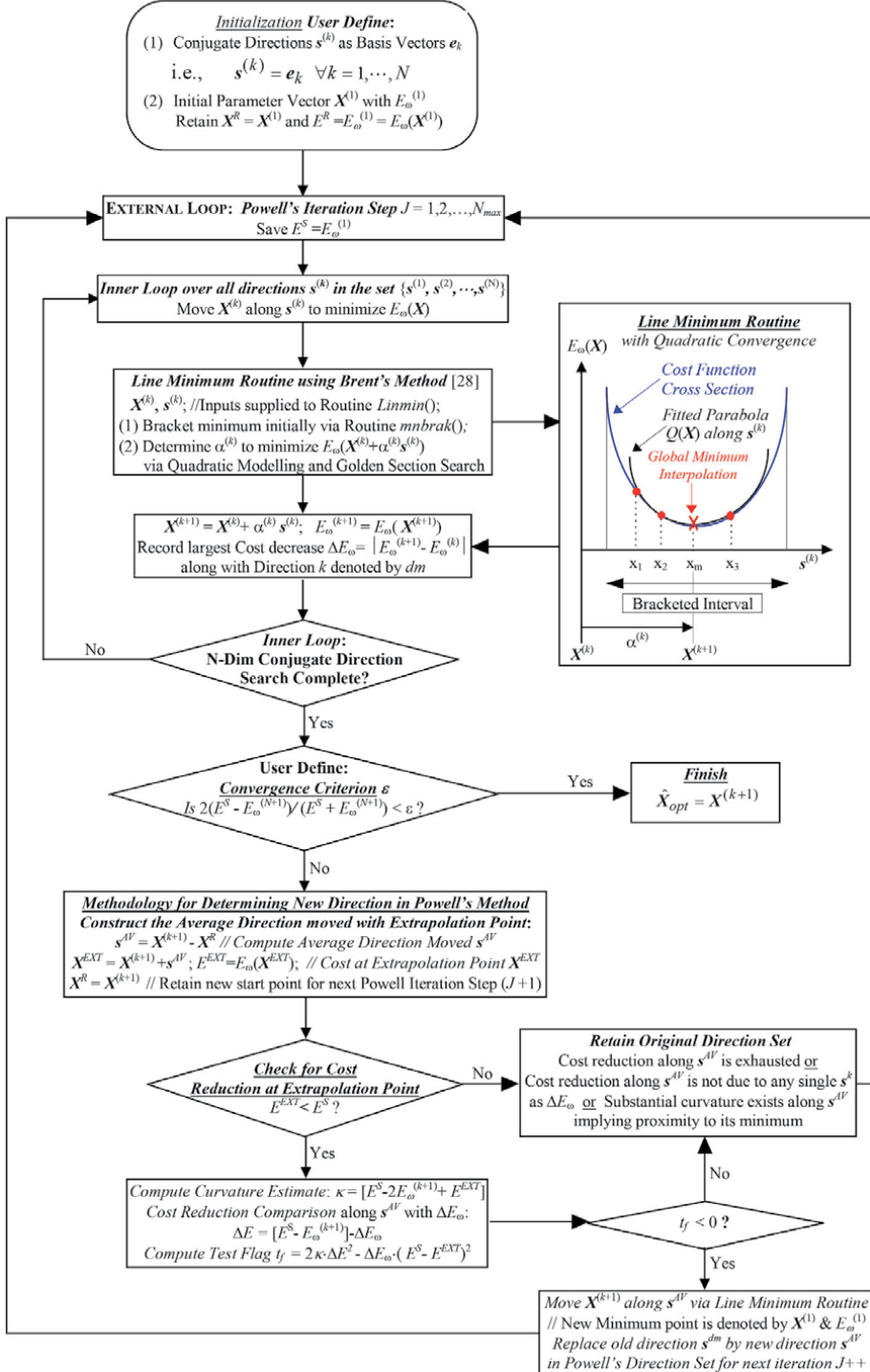
with resolved accuracy given by the worst case quantized step size  $\pm \delta \mathbf{X}^L$  in [13, 28]. In the above line minimum iterative search procedure along a particular direction  $\mathbf{s}^{(k)}$  the minimizer of the cost function is crudely bracketed initially by quadratic polynomial extrapolation. The cost function minimum estimate  $\mathbf{X}^{(k+1)}$  is then obtained by successive approximation within the bracketed interval via interpolation from the fitted quadratic model  $Q(\mathbf{X})$  to the triplet of points  $\{\mathbf{X}_1, \mathbf{X}_2, \mathbf{X}_3\}$ , where the response surface is concave, as

$$\begin{aligned} \mathbf{X}^{(k+1)} &= \mathbf{X}_m \\ &= \mathbf{X}_2 - \frac{1}{2} \cdot \frac{(\mathbf{X}_2 - \mathbf{X}_1)^2 [E_\omega(\mathbf{X}_2) - E_\omega(\mathbf{X}_3)] - (\mathbf{X}_2 - \mathbf{X}_3)^2 [E_\omega(\mathbf{X}_2) - E_\omega(\mathbf{X}_1)]}{(\mathbf{X}_2 - \mathbf{X}_1) [E_\omega(\mathbf{X}_2) - E_\omega(\mathbf{X}_3)] - (\mathbf{X}_2 - \mathbf{X}_3) [E_\omega(\mathbf{X}_2) - E_\omega(\mathbf{X}_1)]} \end{aligned} \quad (6)$$

This minimum estimation procedure is assisted, if iterative progress towards quadratic termination is stalled in a non convergent limit cycle, by conducting a golden section search [56] of the bounded interval through contraction in terms of the golden mean  $g_m$  ( $\sim 0.382$ ) as

$$\mathbf{X}_m = \begin{cases} (\mathbf{X}_3 - \mathbf{X}_2)g_m; & \text{If } \mathbf{X}_2 < (\mathbf{X}_1 + \mathbf{X}_3)/2 \\ (\mathbf{X}_2 - \mathbf{X}_1)g_m; & \text{If } \mathbf{X}_2 > (\mathbf{X}_1 + \mathbf{X}_3)/2 \end{cases} \quad (7)$$

to trap the least  $MSE$  cost  $E_\omega(\mathbf{X}^{(k+1)})$ . During a single cycle iteration of the PCD algorithm the single largest cost decrease  $\Delta E_\omega(\mathbf{X})$  along a particular search direction



**Figure 5.**  
Flowchart of Powell's conjugate direction set optimization technique.

$s^{(dm)}$  is monitored. This cost reduction metric is then used to determine whether or not the set  $S$  of conjugate directions needs to be updated before commencing the next Powell ( $J+1$ )<sup>th</sup> iteration cycle. The decision to include a new search direction, by

replacing that in  $\mathcal{S}$  along which the largest cost decrease  $\Delta E_\omega$  was observed, is made by first evaluating the cost at the extension point

$$\mathbf{X}^{EXT} = \mathbf{X}^{(k+1)} + \mathbf{s}^{AV} \quad (8)$$

along the proposed average direction

$$\mathbf{s}^{AV} = \mathbf{X}^{(k+1)} - \mathbf{X}^{(k)} \quad (9)$$

to avoid linear dependence windup and resultant loss of conjugacy. This cost is then compared with that at the initial search point  $\mathbf{X}^{(S)}$  along with the differential comparison

$$\Delta E = \left[ \left( E_\omega(\mathbf{X}^{(S)}) - E_\omega(\mathbf{X}^{(k+1)}) \right) \right] - \Delta E_\omega \quad (10)$$

to determine if any reduction is achievable along  $\mathbf{s}^{AV}$  for the  $J^{\text{th}}$  iteration, which is regarded as the average path traversed over all possible  $N$  directions in  $\mathcal{S}$ , other than that via  $\mathbf{s}^{(dm)}$ . Furthermore the curvature estimate at  $\mathbf{X}^{(k+1)}$  in the direction  $\mathbf{s}^{AV}$ , given by

$$\kappa = \left\{ E_\omega(\mathbf{X}^{(S)}) - 2E_\omega(\mathbf{X}^{(k+1)}) + E_\omega(\mathbf{X}^{EXT}) \right\} \quad (11)$$

is also checked for possible minimum convergence of the iterative search along with Eq. (10) using the flowchart test condition in **Figure 5** as

$$2\kappa \cdot \Delta E^2 \geq \Delta E_\omega \left\{ E_\omega(\mathbf{X}^{(S)}) - E_\omega(\mathbf{X}^{EXT}) \right\}^2. \quad (12)$$

If the test condition Eq. (12) is true then the old set of directions is retained for the next Powell iteration either because cost reduction is already exhausted or no significant reduction is observed in any particular direction or a substantial second derivative exists along  $\mathbf{s}^{AV}$  indicating minimum convergence of the search estimate. Conversely if Eq. (12) is false then the direction set  $\mathcal{S}$  is updated by replacing  $\mathbf{s}^{(dm)}$  with  $\mathbf{s}^{AV}$  before embarking on the next pass of the Powell algorithm which is usually restricted to some maximum user supplied iteration count  $N_{max}$  of typically 200.

The PCD method was applied to the identification of three known values of BLMD shaft load inertia with experimental shaft velocity target data deployed in the MSE objective function  $E_\omega(\mathbf{X})$  formulation displayed in **Figures 2–4**. An insight into the progress of this optimization method towards global optimality in the extraction of the dynamical  $J$  and  $B$  parameters with initialization at the tolerance band edge can be obtained, for example with zero shaft inertial load conditions, from **Figure 2** over the generated velocity cost surface. After one cycle of the Powell method the iterate has reached the ‘line minimum’ of the cost surface syncline by completing alternate line searches along the  $J$  and  $B$  parameter directions. Quadratic convergence of the PCD method thereafter is relatively ‘slow’ in that a further three iteration steps of the PCD algorithm are required to reach a limit point estimate  $\hat{\mathbf{X}}_{opt}$  of the global minimum. This is manifested as a zigzag search pattern over the stationary region of the response surface in the quantized  $J$  and  $B$  parameter directions along the ‘line minimum’, which is essentially in the  $B$  direction, where the cost function curvature is low with inadequate selectivity for global minimum reachability.

## 6. Parameter convergence results for the PCD method

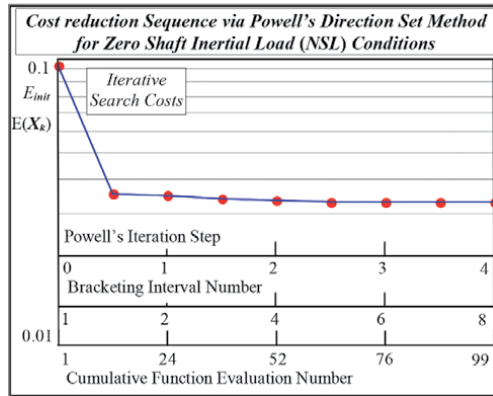
The optimal estimates of the BLMD dynamics returned by the PCD method including convergence details for three known cases of shaft load inertia are summarized in **Table 1**. The penalty cost reduction sequence associated with the application of the Powell algorithm is displayed in **Figures 6–8** for each of identified shaft inertial loads along with the cumulative number of bracketing intervals and cost function evaluations required at each iterative step. Substantial cost reduction, which can be attributed mostly to the adjustment of the inertial parameter, ceases after one iterative cycle of the PCD method in each case signifying the arrival of the parameter estimate in the stationary region along the ‘line minimum’ as in **Figure 2**.

This observation can be adduced from **Table 1** where the listed parameter estimates after the first iteration, in each test case, are reasonably close to the eventual global minimizer estimates. Further gains in MSE reduction are expensive in this zone of the cost surface, due to poor selectivity as explained in [13], with the bulk of the PCD effort devoted towards improving global minimum convergence estimate of the returned parameter vectors. The computational runtime of Powell’s algorithm increases in accordance with the iterative count and the accumulated number of BLMD simulation related MSE evaluations for each of the inertial test cases displayed in **Figures 9–11**. The corresponding averaged BLMD simulation time increases almost in proportion with the inertial loading for a fixed experimental shaft velocity data capture  $V_{\text{or}}$ , normalized to about 10 equivalent machine cycles of motor FC transient as in **Tables 2 and 3**, with time decimation to 4095 sample points.

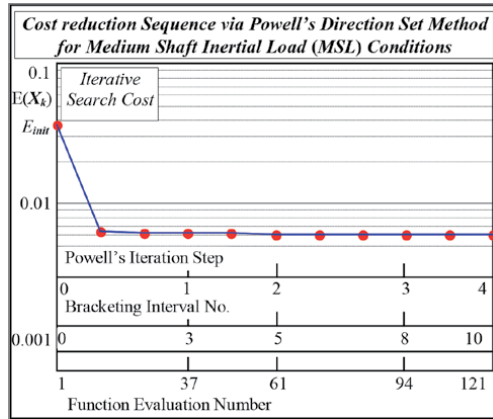
Total Shaft Load Inertia $J$ (kg.cm <sup>2</sup> )	Figure 2 - NSL No Shaft Load $J_m = 3.0$	Figure 3 - MSL Medium Inertia $J = 12.303687$	Figure 4 - LSL Large Inertia $J = 20.822$
Parameter Initialization: $\{J_{\text{init}} = 0.82 J, B_{\text{init}} = 1.09 B_m\}$ with $B_m = 2.14 \times 10^{-3}$ Nm/rad/sec			
Initial Cost $E_{\text{init}}$	$10.2481 \times 10^{-2}$	$36.5942 \times 10^{-3}$	$54.4345 \times 10^{-3}$
Returned Parameter Estimates after first Iteration Cycle of PCD Method			
$J_{\text{opt}}^{(1)}$ (kg.cm <sup>2</sup> )	3.02659	12.0721	21.292
$B_{\text{opt}}^{(1)}$ (Nm.rad <sup>-1</sup> .sec)	$2.10405 \times 10^{-3}$	$2.14214 \times 10^{-3}$	$2.06596 \times 10^{-3}$
Optimal Parameter Estimates returned from Powell’s Conjugate Direction Method			
$\hat{J}_{\text{opt}}$ (kg.cm <sup>2</sup> )	3.168237	12.29243	21.51233
$\hat{B}_{\text{opt}}$ (Nm.rad <sup>-1</sup> .sec)	$1.913588 \times 10^{-3}$	$1.989772 \times 10^{-3}$	$1.95168 \times 10^{-3}$
Min. Cost $\hat{E}_{\text{opt}}$	$3.278108 \times 10^{-2}$	$5.910952 \times 10^{-3}$	$2.456056 \times 10^{-2}$
Total No. of PCD Iterations	4	4	2
No. of Cost Func.-Eval <sup>‡</sup>	99	121	56
Initial Bracketing	24	31	15
Line Searching	74	89	40
Av. No. of Func Eval <sup>‡</sup> /Iter <sup>n</sup>	~25	~30	28
Total Time (sec)	3584	5234	2918
Average Iter <sup>n</sup> Time (sec)	896	1308.5	1459
Time/MSE Cost Eval. (sec)	36.202	43.256	52.106

**Table 1.**  
Returned BLMD dynamical parameter estimates via PCD algorithm.

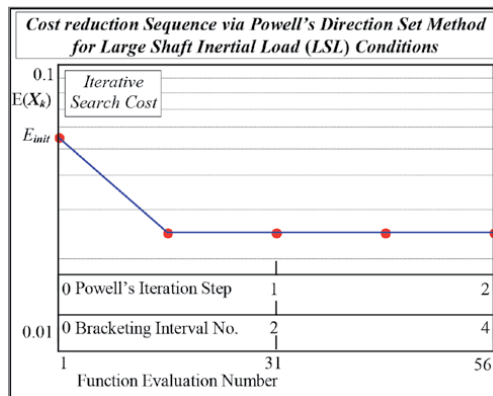




**Figure 6.**  
 NSL iterative cost reduction.

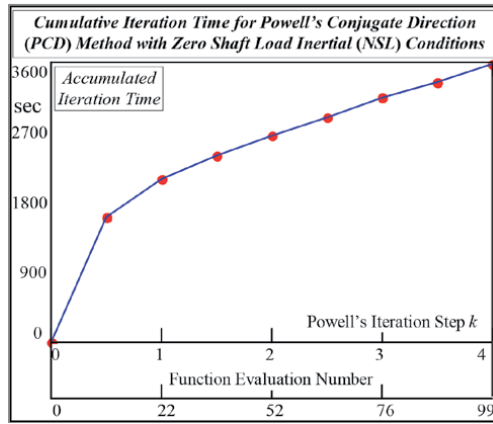


**Figure 7.**  
 MSL iterative cost reduction.

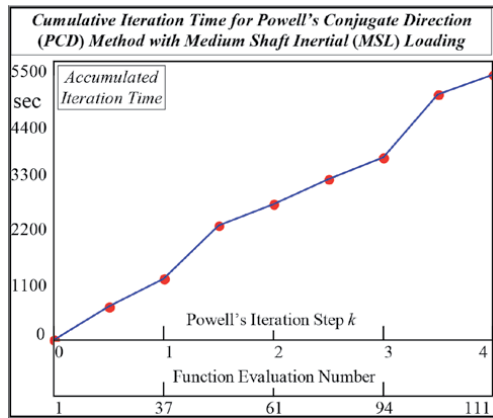


**Figure 8.**  
 LSL iterative cost reduction.

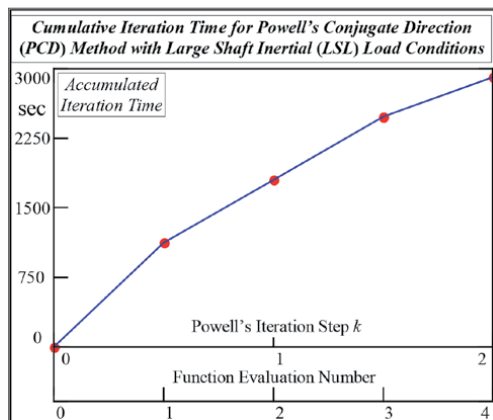
This capture restriction of the shaft velocity target data is due to the experimental constraints of the data acquisition system used, which limited the size of the data record acquired. These returned PCD statistics are also based on a fixed time step of



**Figure 9.**  
NSL cumulative iteration time.



**Figure 10.**  
MSL cumulative iteration time.



**Figure 11.**  
LSL cumulative iteration time.

1  $\mu$ s, in relation to BLMD model exercise in the MSE penalty cost formulation, for accurate realization of the PWM edge transitions and for computation benchmarking purposes in both the PCD and modified FSD methods of identification.

The shaft velocity data sets can also be deployed in the construction of MSE response surfaces  $E_{\omega}$  over the two dimensional  $[J, B]$  quantized parameter manifold [13], with mesh size details given in **Table 2**, as a secondary means of parameter extraction for qualification of the PCD method and verification of the returned PCD optimal estimates in terms of accuracy. These cost surfaces, pertaining to the various motor shaft inertial loads, are illustrated in **Figures 2–4** and appear to possess a parabolic ravine-like structure with an embedded elliptical stationary region stretched into a ‘line minimum’ shape [13]. These cost constructs can be used as the basis for an error analysis as given in **Table 4** by which the accuracy of PCD estimates are gauged.

The percentage relative error in the returned PCD optimal estimate  $\hat{J}_{opt}$  appears to increase with shaft inertial load but with no apparent trend noticeable in the damping parameter estimate. The percentage error in the identified  $\hat{B}_{opt}$  parameter, however, exceeds that for the inertia in all three test cases which highlights the problem in accurately extracting the damping coefficient due to presence of a pronounced ‘line minimum’ of admissible friction values predominantly in the  $B$  parameter direction.

The opposite error pattern will be shown to occur in the extraction of the inertia estimate, coupled with lower relative error in the identification of the friction coefficient, upon application of the modified form of the FSD method with details given in **Table 5**. This contrast in error pattern for the returned FSD parameters is due to better discriminating features of the corrugated FC cost surface in terms of its selectivity, with greater curvature in the  $B$  parameter direction for increased data capture length, and lower threshold parameter step sizes as shown in **Tables 7–9** in [13].

The observed phase-a current demand, feedback and controller o/p waveforms, displayed in **Figures 12–20** at critical internal nodes of the BLMD model [10], appear to be coherent with similar test data in the early phase of the transient response of an actual drive system to a unit torque demand step i/p. However there is an eventual loss of synchronism, with the evolution of the model step response

$\delta J^L = 15.7386 \times 10^{-3} \text{ kg.cm}^2 \equiv 0.51\% J_m;$ $J_m = 3.086 \times 10^{-4} \text{ kg.m}^2$ (NSL - Rotor Inertia)	$\delta B^L = 38.092 \times 10^{-6} \text{ Nm.rad}^{-1}.\text{sec} \equiv 1.78\% B_m;$ $B_m = 2.14 \times 10^{-3} \text{ Nm.rad}^{-1}.\text{sec}$		
Shaft Velocity Target Data $V_{\omega r}$	<i>No Shaft Load</i>	<i>Medium Inertia</i>	<i>Large Inertia</i>
Data Sample Rate $T_s$	20 $\mu$ s	40 $\mu$ s	49.6 $\mu$ s
Decimation Factor	20	40	50
No of Equivalent FC Cycles for Computation Benchmarking	$\sim 9.75$	$\sim 11.5$	$\sim 10.5$
Simulation Grid Size in $E_{\omega}$ [ $\Delta J, \Delta B$ ] as per [13]	<b>Figure 2</b> [ $2\delta J^L, \delta B^L$ ]	<b>Figure 3</b> [ $4\delta J^L, \delta B^L$ ]	<b>Figure 4</b> [ $7\delta J^L, \delta B^L$ ]
Surface Minimum Estimates			
$\bar{J}_{opt}$ (kg.cm <sup>2</sup> )	3.184	12.38	22.36
$\bar{B}_{opt}$ (Nm.rad <sup>-1</sup> .sec)	$1.875 \times 10^{-3}$	$1.875 \times 10^{-3}$	$1.875 \times 10^{-3}$
$\bar{E}_{opt}$	$3.26 \times 10^{-2}$	$5.819 \times 10^{-3}$	$2.453 \times 10^{-2}$

**Table 2.**  
 Details of velocity cost surface generation with parameter quantization.

FC Target Data	No Shaft Load	Medium Inertia	Large Inertia
No. of Cycles	~9.75	~11.5	~10.5
Data Sample Rate $T_s$	20 $\mu$ S	40 $\mu$ S	49.6 $\mu$ S
Decimation Factor	20	40	50
Simulation Mesh Size [ $\Delta$ , $\Delta B$ ]	[ $2\delta J^L$ , $\delta B^L$ ]	[ $6\delta J^L$ , $\delta B^L$ ]	[ $10\delta J^L$ , $\delta B^L$ ]
Surface Minimum $\bar{J}_{opt}$ (kg.cm <sup>2</sup> )	3.089	12.158	20.877
$\bar{B}_{opt}$ (Nm.rad <sup>-1</sup> .sec)	$1.921 \times 10^{-3}$	$1.921 \times 10^{-3}$	$1.921 \times 10^{-3}$
Min Cost $\bar{E}_{opt}$	$4.922 \times 10^{-2}$	$1.877 \times 10^{-2}$	$1.236 \times 10^{-2}$

**Table 3.**  
Quantized parameter FC response surface simulation details.

Motor Inertia $J_{opt}$	No Shaft Load	Medium Shaft Inertia	Large Shaft Inertia
$\left(\frac{J_{opt}-\bar{J}}{\bar{J}}\right) \times 100\%$	-0.495%	-0.707%	-3.79%
Damping Factor $B_{opt}$ $\left(\frac{B_{opt}-\bar{B}}{\bar{B}}\right) \times 100\%$	-2.058%	-6.121%	-4.09%
Global Cost $E_{opt}$ $\left(\frac{E_{opt}-\bar{E}}{\bar{E}}\right) \times 100\%$	0.555%	1.58%	0.125%

**Table 4.**  
Error analysis for returned PCD parameter estimates (% relative error).

Motor Inertia $J_{opt}$	No Shaft Load	Medium Shaft Inertia	Large Shaft Inertia
$\left(\frac{J_{opt}-\bar{J}}{\bar{J}}\right) \times 100\%$	-0.29%	-0.15%	-0.19%
Damping Factor $B_{opt}$ $\left(\frac{B_{opt}-\bar{B}}{\bar{B}}\right) \times 100\%$	1.98%	4.16%	1.98%
Global Cost $E_{opt}$ $\left(\frac{E_{opt}-\bar{E}}{\bar{E}}\right) \times 100\%$	5.99%	17.37%	5.42%

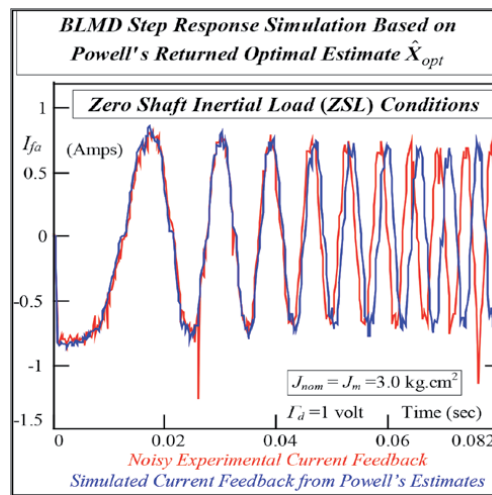
**Table 5.**  
Error analysis for returned FSD parameter estimates (% relative error).

towards steady state conditions, in the waveform comparison due to the impact of the estimated  $J/B$  dynamic time constant mismatch in relation to the intrinsic value  $\tau_m$  of the actual motor drive system.

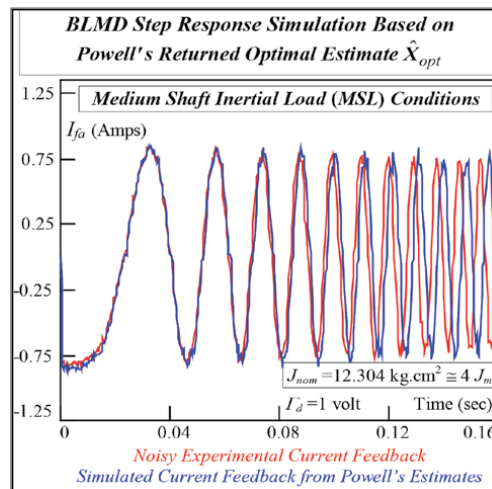
This is borne out by the low correlation measurement coefficients in **Table 6**, which gauges the degree of FM coherence, between the respective experimental and simulated trace responses. Further evidence of this mismatch, though small, can be visualized in the deviation of the shaft velocity characteristics depicted in **Figures 21–23**, despite the high correlation coefficient to the contrary, as the motor drive accelerates towards rated shaft speed. A possible explanation for the correlation discrepancy between waveform types may attributed to the application in this instance of shaft velocity target data in BLMD parameter extraction process which results in a good fit between shaft velocity waveforms based on returned estimates due to PCD cost minimization. The returned estimates, however, result in a less than satisfactory trace coherence measure of fit in the current related characteristics

Nominal Shaft Inertial Load $J_{nom}$ (kg.cm <sup>2</sup> )	No Shaft Load (NSL - Rotor) $J_m = 3.0$	Medium Shaft Load (MSL) 12.304	Large Shaft Load (LSL) 20.822
Current Feedback $I_{fa}$	Figure 12 47.4%	Figure 13 67.2%	Figure 14 61.0%
Current command $I_{da}$	Figure 15 44.2%	Figure 16 67.6%	Figure 17 53.0%
Current Controller o/p $V_{ca}$	Figure 18 18.1%	Figure 19 38.8%	Figure 20 20.4
Motor Shaft Velocity $V_{\omega r}$	Figure 21 93.7	Figure 22 99.1%	Figure 23 98.6%

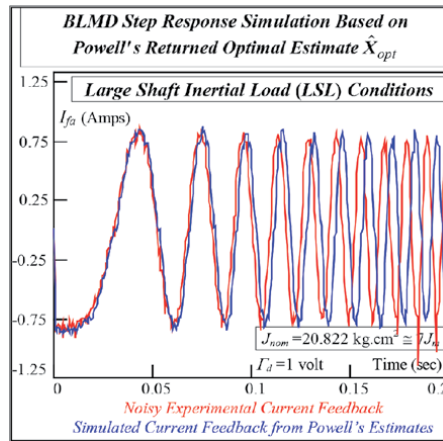
**Table 6.**  
 BLMD simulation trace coherence correlation coefficient  $\rho$ .



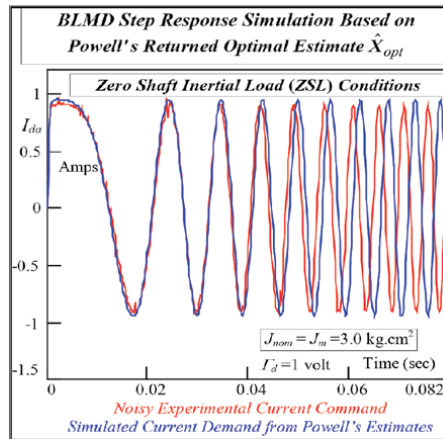
**Figure 12.**  
 NSL FC simulation.



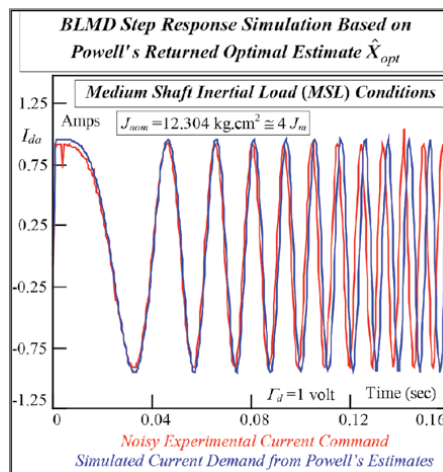
**Figure 13.**  
 MSL FC simulation.



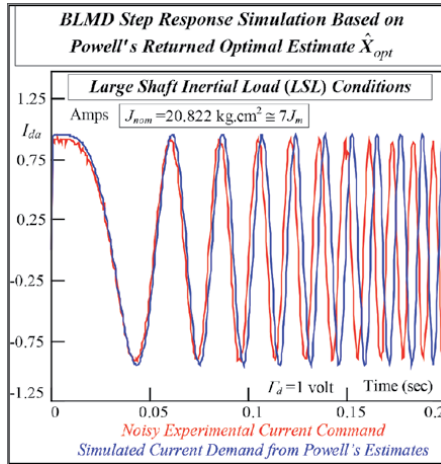
**Figure 14.**  
LSL FC simulation.



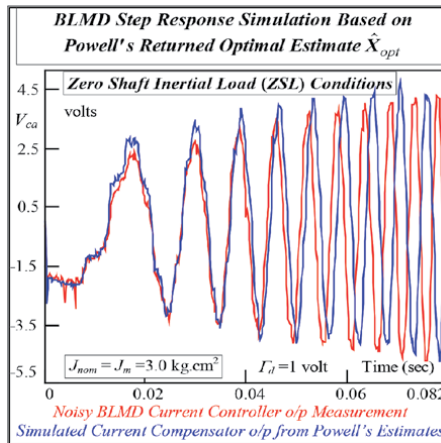
**Figure 15.**  
NSL current demand simulation.



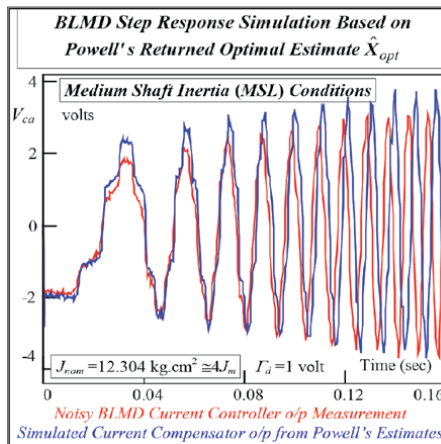
**Figure 16.**  
MSL current demand Simul<sup>n</sup>.



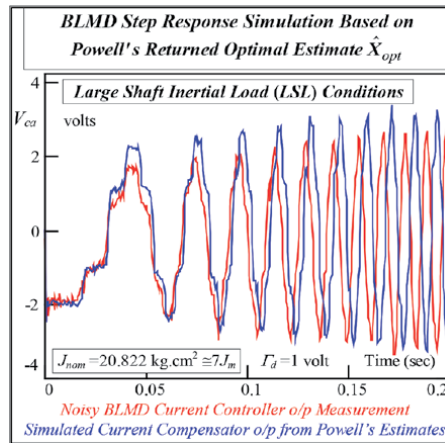
**Figure 17.**  
 LSL current demand Simul<sup>n</sup>.



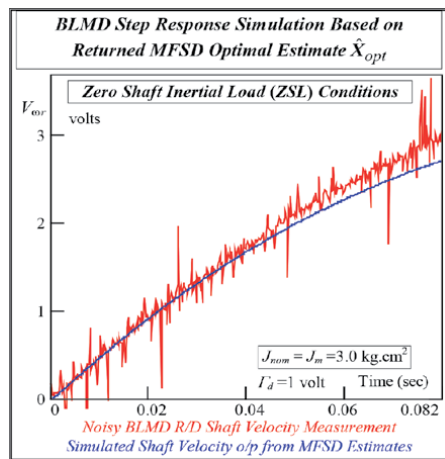
**Figure 18.**  
 NSL - current controller o/p.



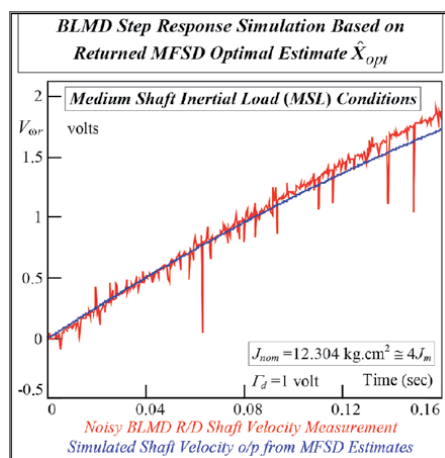
**Figure 19.**  
 MSL - current controller o/p.



**Figure 20.**  
LSL - current controller o/p.

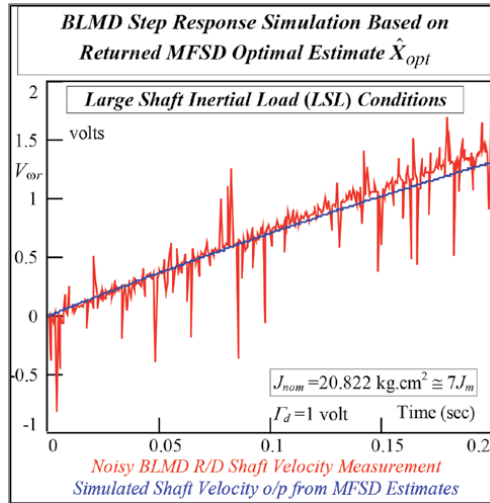


**Figure 21.**  
ZSL - shaft velocity simulation.



**Figure 22.**  
MSL - shaft velocity simulation.

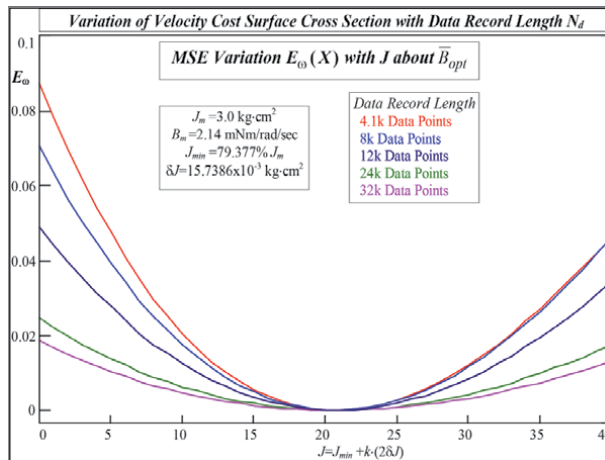




**Figure 23.**  
 LSL - shaft velocity simulation.

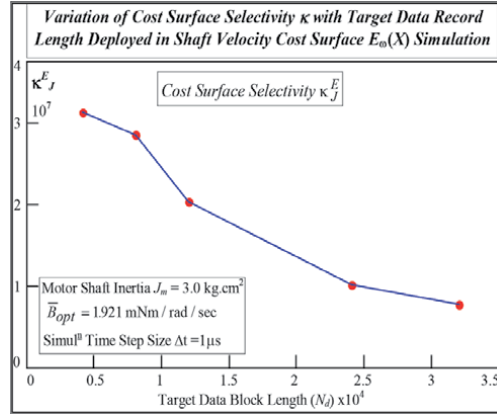
Data Record Length $N_d$	4095	8000	12000	24000	32000
Surface $E_{\omega}(\bar{J}_{opt})$ Curvature $\kappa_f^2$	$3.122 \times 10^7$	$2.843 \times 10^7$	$2.036 \times 10^7$	$1.026 \times 10^7$	$7.692 \times 10^3$
Convergence Metric $N_f^5$	10.3	10.8	12.8	18	20.8
Best Parameter Resolution Possible with PCD Algorithm for $\delta J^L = 15.739 \times 10^{-3} \text{ kg.cm}^2$					
Total Shaft Inertial Load $J$ ( $\text{kg.cm}^2$ )		Shaft Inertia $J_m = 3.0$	Medium Inertia 12.304	Large Inertia 20.822	
% parameter Resolution $\delta J^L/J$		0.525%	0.128%	0.0756%	

**Table 7.**  
 Cost surface selectivity measure – Figures 24, 25 and 26.

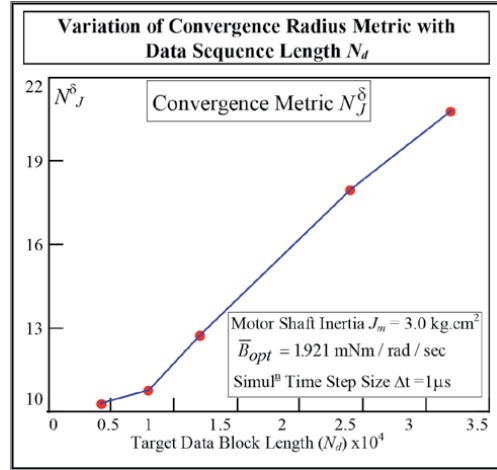


**Figure 24.**  
 Cost surface variation with data record length.

pertaining to the BLMD current loop operation in torque control mode. Furthermore the use of larger quantized parameter step sizes, in conjunction with shaft velocity target data in the PCD parameter extraction process, results in a greater spread of



**Figure 25.**  
Cost surface selectivity.



**Figure 26.**  
Global convergence variation.

returned optimal estimates about the global extremum compared with the threshold values deployed with current feedback data in the FSD algorithm.

The initial impact of the shaft velocity transient step response decays and is ultimately swamped with the onset of steady state conditions as maximum shaft speed is reached with lengthening data records. The quadratic shape of the response surface  $E_\omega$  is de-emphasized with reduced curvature, as more velocity target data is accumulated for MSE cost formulation, resulting in a loss of selectivity at the global minimizer.

This flattening of the cost profile about  $X_{opt}$  is clearly evident in **Figure 24** for response surface cross sections in the inertia parameter  $J$ , based on simulated BLMD model target data, with increased data record length. The resultant ill conditioned surface admits a multitude of possible global minimum estimates about  $X_{opt}$ . The variation in cost function selectivity can be obtained by fitting a quadratic polynomial approximation [13] to each of the response surface cross sections  $E_\omega(J)$  in **Figure 24** and estimating the curvature at the fitted parabolic vertex  $\bar{J}_{opt}$  as

$$\kappa_J^E = d^2Q(j)/dJ^2|_{J=\bar{J}_{opt}} = 2b_2 \quad (13)$$

<i>I<sub>fa</sub></i> Cost Surface Target Data [13]		Sample Size $N_i = 200 \approx 10\%N_C$
Parameters $x$	Inertia $J_m$ kg.cm <sup>2</sup>	Damping $B_m$ Nm/rad/sec
Nominal Values $x_m$	3.0	$2.14 \times 10^{-3}$
Parameter Bounds $\pm \Delta x_m$	$\pm 20\%$	$\pm 10\%$
Quantized Step Size $\delta x^L$	$9.1125 \times 10^{-3}$	$2.6712 \times 10^{-5}$
Population Size $N_x$	141	16
Results of Initial Exploratory Search of Parameter Space		
Histograms	<b>Figure 30</b>	<b>Figure 31</b>
Class Interval Size	$10\delta j^L$	$\delta B^L$
Minimum Value $x_{min}$	2.408	1.953
Maximum Value $x_{max}$	3.592	2.327
Actual Mean Value $x_{mean}$	3.041	2.153
Theoretical Mean Value $\bar{x}$	3.0163	1.9934
% Error in Mean Value	0.82%	8%
Actual Standard Deviation $\hat{\sigma}_x$	0.343	0.109
Theoretical Standard Deviation $\sigma_x$	0.518	0.750
Optimal Parameter Estimate $\hat{x}_{iopt}$	3.06379	1.953
FSD Exploratory Search Results of FC - MSE Cost Surface		
Initial Global Minimum Estimate $\hat{E}_{iopt}$		$5.473 \times 10^{-2}$
Maximum Value $E_{max}$		$6.936 \times 10^{-1}$
Mean Value $\bar{E}$		$3.6671 \times 10^{-1}$
Standard Deviation $\sigma$		$1.6538 \times 10^{-1}$
Initial Temperature $T_i$ ( $= k\sigma$ for $k = 10$ )		1.6538

**Table 8.**  
 Initial details of FSD exploratory search phase.

in **Table 7** with

$$Q(\bar{J}_{opt}) = 0 \quad \text{and} \quad dQ(J)/dJ|_{J=\bar{J}_{opt}} = 0. \quad (14)$$

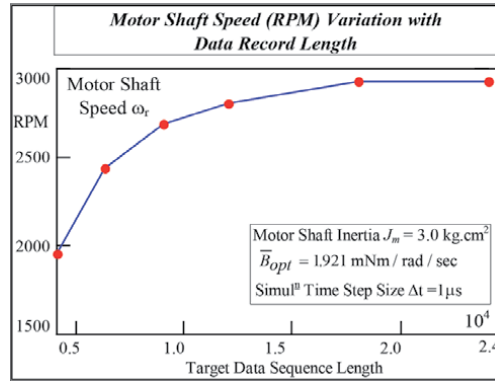
This curvature  $\kappa_J^E$  variation, shown in **Figure 25**, exhibits a decreasing quasi-linear dependency with target data length and indicates as a consequence poor response surface selectivity and large convergence radius in the stationary region containing the global minimum. An estimate of the global convergence radius  $r_J$  of the  $J$  parameter stationary zone can be obtained in terms of the fixed cost noise estimate  $\hat{\sigma}$ , due to inexact PWM simulation [13], and the curvature variation with different data training record sizes  $N_d$  via Eq. (13) as

$$r_J = \sqrt{\frac{8\hat{\sigma}}{\kappa_J^E}}. \quad (15)$$

This radial width can be referenced to the threshold step size  $\delta j^L$  in **Table 2**, in order to establish the scale of the bounded region of convergence, with metric variation for different data sequence lengths  $N_d$  given in **Table 7** as

$$N_j^\delta = r_j / \delta J^L \quad (16)$$

and displayed in **Figure 26**. This characteristic illustrates clearly the emergence of a pattern of reduced cost selectivity, which is mirrored as a radial extension of the global minimum region, with increased target data capture. The opposite trend prevails in **Figure 27** with the application of current feedback in the MSE penalty cost and is the main motivation for its use as a target function with a tighter bound, besides sensorless motor control issues, in the identification of the BLMD dynamics despite the FSD computational intensity. The use of a fixed threshold step size establishes the degree of accuracy possible in parameter resolution during BLMD system identification as shown in **Tables 7** and **9**. The relative percentage accuracy in the returned estimates improves with motor shaft inertial loading as tabulated with a greater resolution possible with the deployment of current feedback target data in MSE cost reduction. If a coarser step size  $\delta \mathbf{X}^*$  is adopted, which may be tolerable at large inertial loads without adversely affecting the percentage resolution in the returned estimates, both the FSD and PCD methods of parameter extraction can proceed much faster with smaller computational burdens. This results from the reduced number of the feasible lattice points to be searched in the quantized parameter domain with small changes in the percentage accuracy. The required parameter accuracy can be user defined in such circumstances at the start of the identification search routine and encoded in the relevant step size  $\delta \mathbf{X}^*$  as a measure of the desired coarseness of resolution.



**Figure 27.**  
Motor shaft speed variation.

Data Record Length $N_d$	6280	9000	12000	24000	32000
Fitted Coefficient $b_0$	$3.708 \times 10^9$	$7.053 \times 10^9$	$9.495 \times 10^9$	$1.93 \times 10^{10}$	$2.398 \times 10^{10}$
Selectivity Measure $S_j^E$	$3.379 \times 10^3$	$6.427 \times 10^3$	$8.652 \times 10^3$	$1.759 \times 10^4$	$2.186 \times 10^4$
Parameter Resolution Accuracy in FSD Estimation with $\delta J^L = 9.1125 \times 10^{-3} \text{ kg.cm}^2$					
Total Shaft Inertial Load $J$ (kg.cm <sup>2</sup> )	Shaft Inertia $J_m = 3.0$	Medium Inertia 12.304	Large Inertia 20.822		
% parameter Resolution $\delta J^L / J$	0.304%	0.074%	0.0438%		

**Table 9.**  
Cost surface selectivity measure.

## 7. Description of FSD method of parameter identification

The simulated annealing technique used here as a global optimization algorithm for motor parameter identification [22, 23] is based on the Fast Simulated Diffusion (FSD) method proposed by Sakurai et al. [24]. This method, motivated by quantum mechanics (QM), is modeled on the diffusion  $\nabla E(x_i)$  of a particle, with position trajectory  $x_i$ , across a barrier potential  $E(x_i)$  under the influence of a temperature controlled random driving force  $\sqrt{2T}dw_i$  resulting in Brownian motion. The particle ensemble dynamics [53] can be described by the Ito stochastic differential equation with space time coordinates  $(x, t)$  as

$$dx = -\nabla E(x)dt + \sqrt{2T}dw. \quad (17)$$

The down-hill gradient term  $-\nabla E(x)$  in Eq. (17), which is effective at low temperature, has the tendency to minimize the particle potential  $E(x)$  along the continuous “parameter” path  $x_i$  as

$$\frac{dx_i}{dt} = -\nabla E(x_i) \quad (18)$$

and thus improve the cost of the objective function  $E(x)$ . The stochastic perturbation component of motion  $dw_i$  along a particular trajectory  $x_i(t)$ , which is an essential coherent energy interaction at high temperature  $T$ , imparts enough momentum to enable the particle to “hill climb” its way out of a potential well and thus avoid local minimum capture with

$$dx_i = -\nabla E(x_i)dt + \sqrt{2T}dw_i. \quad (19)$$

If a proper cooling procedure [57] is implemented the global minimum potential  $E_{min}$  is reached after infinite time with a Gibb’s distribution having a limiting Boltzman probability density [58].

$$p(x) \propto \exp\left(-\frac{E(x)}{T}\right) \text{ as } t \rightarrow \infty \quad (20)$$

which peaks with a value of one at the global minimizer as  $T$  approaches zero. The global minimizer of  $E(x)$  can be obtained according to Aluffi-Pentini et al. [53] from inspection of the asymptotic value of a numerically computed sample trajectory  $x(t)$  as time  $t \rightarrow \infty$  via numerical integration of Eq. (17) from  $x(0)$  at  $t = 0$  as the stochastic temperature  $T$  is reduced very slowly to zero with time  $t$ .

The differential random process in Eq. (17) has been mapped into an algorithmic procedure for approximating the global minimizer [24] of multim minima objective functions with a large  $N$ -Dimensional parameter manifold. Notable applications of discrete SA type methodology include the time scheduling problem of a traveling salesman to  $N$  cities which is NP-complete [15] and to cell placement in integrated circuit design. The SA method, however, requires very large computing resources [49] for global minimum convergence in such problem solutions whereas the introduction of the gradient search component in Eq. (17) with an aggressive cooling schedule can result in significant gains in the reduction of CPU time. In general the true location in parameter space of the global minimum of a multim minima cost function cannot be guaranteed with certainty in practical applications of the SA algorithm but a good sub-optimal approximation can be obtained with reasonable computational effort which is an acceptable estimate  $\hat{X}_{opt}$  of the optimal parameter vector  $X_{opt}$ . Similar comments

pertain to the FSD global minimum approximation during BLMD parameter extraction over the FC cost surface which is described herein.

In BLMD model parameter optimization the particle energy  $E(x)$  is replaced by the FC cost function  $E_{fa}(\mathbf{X})$  and its positional  $x_i$  configuration space by the parameter set  $\mathbf{X}$  bounded by the permissible tolerance band  $\pm \Delta\mathbf{X}_m$ . The FSD algorithm relies on two modifications, instead of integrating Eq. (19) directly, to accelerate convergence and prune wayward or non profitable moves in parameter space. The first is based on the use of an accept/non-accept function rule after Metropolis et al. [59] where the probability of acceptance of the next move

$$\mathbf{X}_{k+1} = \mathbf{X}_k + d\mathbf{X} \quad (21)$$

in the parameter identification search is governed by the Boltzman's distribution. If the next move yields a lower cost  $E(\mathbf{X}_{k+1})$  than the current value  $E(\mathbf{X}_k)$  at  $\mathbf{X}_k$  then the new parameter configuration  $\mathbf{X}_{k+1}$  is accepted. Alternatively in an error increasing move a random number

$$R_n \in [0, 1] \quad (22)$$

is chosen for arbitration in the move selection process and the transition probability

$$P_{k+1} = \exp\left(\frac{-\{E(\mathbf{X}_{k+1}) - E(\mathbf{X}_k)\}}{T}\right) \quad (23)$$

is computed. If the resulting decision is

$$P_{k+1} > RN \quad (24)$$

then the displacement  $\mathbf{X}_{k+1}$  is accepted, as a controlled uphill step as part of the iterative improvement process for a better solution, otherwise it is rejected in which case  $\mathbf{X}_{k+1}$  has to be regenerated. The second modification involves the alternate application of the hill-descending gradient search term  $-\nabla E(\mathbf{X})$  with the temperature dependent stochastic term in Eq. (19) in the generation of the next move. This alternate application ensures limited hill-descent at high temperature than otherwise would be the case if the two terms were added together as in the conventional method where ineffective moves due to the stochastic term would negate gains made in the gradient search. Instead of calculating the direction of steepest descent  $-\nabla E(\mathbf{X})$ , which is expensive for large dimensional parameter space, the alternative directional entity

$$\langle \nabla E(\mathbf{X}) \cdot \hat{r} \rangle \hat{r} \quad (25)$$

is used instead because its expected value approaches  $-\nabla E(\mathbf{X})$  in the long term where  $\hat{r}$  is a unit vector along a randomly chosen axis in parameter  $\mathbf{X}$  space. A quadratic fit is then employed if the cost function is concave along the randomly chosen axis and the minimum is estimated via the Newton–Raphson method. If the cost function happens to be convex in the chosen direction a small  $d\mathbf{X}$  is first used and then doubled up a prescribed number of times until  $E(\mathbf{X}_k + d\mathbf{X})$  fails to decrease. This approach gives a crude but inexpensive estimate of the minimum.

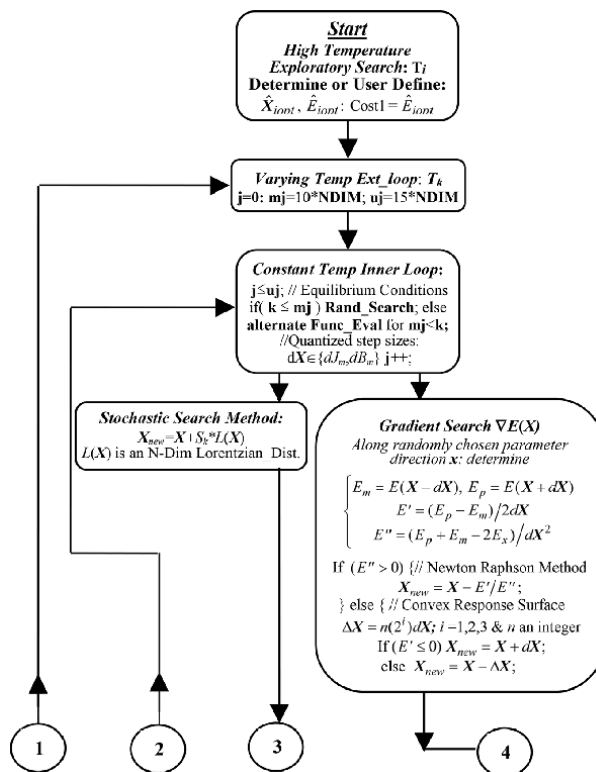
## 8. Application of FSD method to motor parameter estimation

The FSD iterative process proceeds slowly in accordance with a pseudo temperature control parameter  $T_k$ , with the same units as the cost surface, from an initial

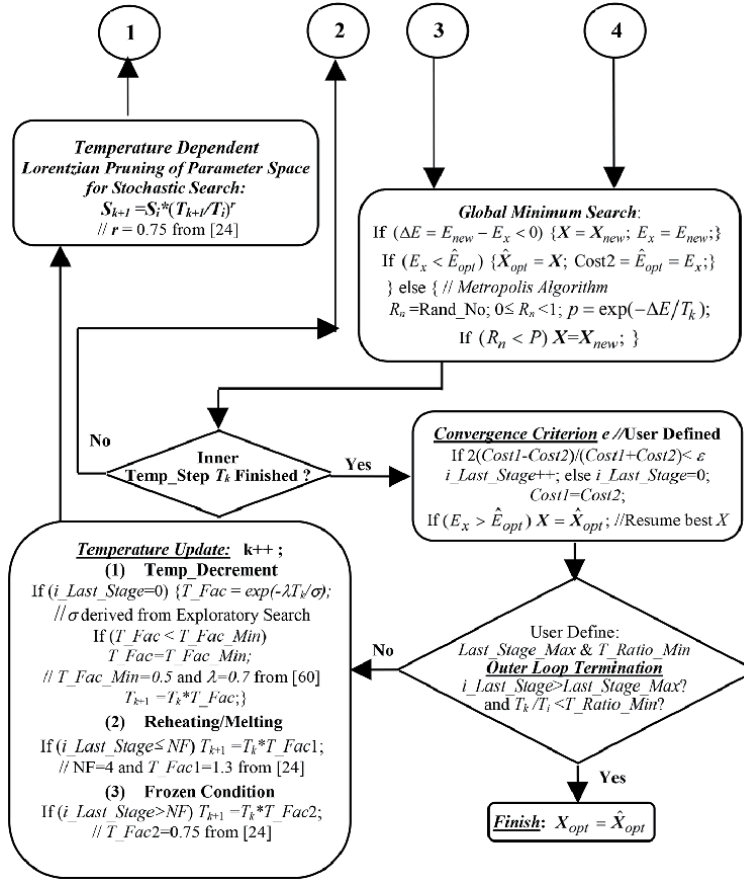
high temperature value  $T_i$ . This annealing technique permits the generation of a population of parameter vector  $\mathbf{X}$  adjustments at each anneal temperature step to simulate the effect of cooling tardiness for the parameter ensemble to reach a steady state configuration.

The cooling process also includes an equilibrium condition based on some minimum level of acceptance at each temperature step before progressing on to the next step. The temperature  $T_k$  is reduced very slowly, according to a prescribed anneal curve [60] or exponentially [15], to avoid premature quenching and resultant parameter configuration trapping in a metastable state. The outer temperature loop decrementation sequence and parameter adjustment acceptances at each inner loop equilibrium temperature constitute the simulated annealing schedule. The temperature reduction should be rapid with an early detection of equilibrium for effective cooling without quenching. Condensation followed by termination of the anneal method occurs when the cost function value tends to reduce slowly and remain unchanged for several consecutive temperature steps according to some stopping criterion. At this stage the parameter configuration has frozen to an optimal arrangement in the neighborhood of the global minimum cost.

The flowchart of the complete FSD algorithm is presented in **Figures 28** and **29** with details of the inner and outer temperature loops. The high temperature melting phase, used in initialization, and the cooling procedure for the FSD algorithm are adopted from Huang et al. [60]. The annealing start temperature  $T_i$  is statistically determined from the standard deviation  $\sigma$  of the cost function distribution over  $N_i$  sample points from an initial exploratory search of parameter space. During this hot phase the feasible parameter domain is uniformly sampled with the temperature  $T_i$



**Figure 28.**  
 Flowchart of fast simulated diffusion algorithm – Heating phase.



**Figure 29.**  
Flowchart of fast simulated diffusion algorithm – Cooling phase.

assumed high enough at ‘infinity’  $T_{\infty}$  such that all generated states are accepted. An adequate sample size is established from the tolerance bounds  $\pm \Delta \mathbf{X}_m$  imposed, which define a hypercube of feasible lattice points in parameter space, with interstitial distance based on the quantized step sizes [13]. A sample size of 10% of the total number determined [13, 28] as

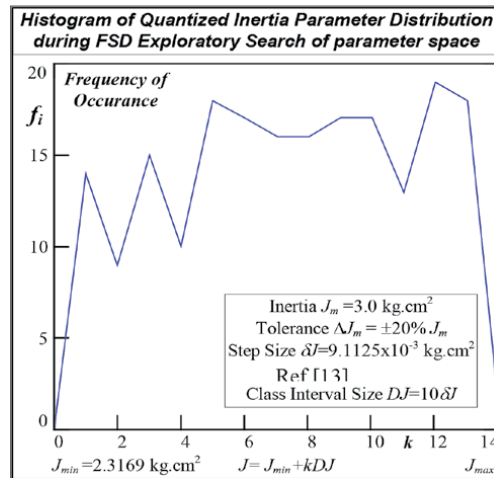
$$N_C = N_B \cdot N_J = \left( \frac{2\Delta B_m}{\delta B^L} \right) \left( \frac{2\Delta J_m}{\delta J^L} \right) = (16) \cdot (141) = 2256 \quad (26)$$

respectively, gives a good search coverage of parameter space for tolerance bounds given in **Table 8** with  $N_i = 200$ . This is evident from the histograms of the near uniform parameter search distributions shown in **Figures 30** and **31** with details in **Table 8** for interval sizes of  $\delta B^L$  and  $10\delta L^L$ . The uniformity of the exploratory search distribution can be checked from theoretical consideration of the population size  $N_x$  of quantized parameter values, for the given tolerance bounds, as

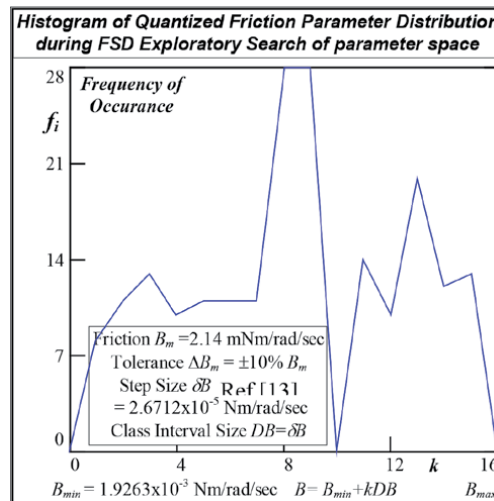
$$N_x = (2\Delta x_m) / \delta x^L \quad (27)$$

The mean parameter estimate obtained from a random search of discretized parameter space, with sample space size  $N_x$  and uniform probability of occurrence





**Figure 30.**  
 Histogram of J search value.



**Figure 31.**  
 Histogram of B search value.

$$p_x = 1/N_x, \tag{28}$$

is given by

$$\begin{aligned} \bar{x} &= \frac{1}{N_x} \sum_{j=1}^{N_x} X_j = \frac{1}{N_x} \sum_{j=1}^{N_x} [(x_m - \Delta x_m) + (j - 1)\delta x^L] \\ &= \left(\frac{N_x - 1}{2N_x}\right) \{2(x_m - \Delta x_m) + (N_x - 1)\delta x^L\} \end{aligned} \tag{29}$$

This theoretical estimate compares favorably with those in **Table 8** obtained from FSD simulation with low relative error percentages, which verifies the randomness quality of the initial search. The standard deviation  $\sigma_x$  of the parameter search estimates can be likewise determined theoretically via Eqs. (28) and (29) from the variance  $\sigma_x^2$  with

$$\begin{aligned}\sigma_x^2 &= \overline{x^2} - \bar{x}^2 = \frac{1}{N_x} \sum_{j=1}^{N_x} x_j^2 - \bar{x}^2 \\ &= (x_m - \Delta x_m)^2 + (N_x - 1) \cdot (x_m - \Delta x_m) \delta x^L + \frac{(\delta x^L)^2}{6} (N_x - 1)(2N_x - 1) - \bar{x}^2\end{aligned}\quad (30)$$

The standard error  $\hat{\sigma}_x$  in the simulated parameter search estimates in **Table 8** is of the same order of magnitude as that obtained from Eq. (30) with a sizeable discrepancy in the viscous friction coefficient which is possibly due to the limited  $B$  parameter sample space used. The initial temperature  $T_i$  is determined from the resultant sample cost distribution illustrated in **Figure 32**, with details in **Table 8**, as

$$T_i = k\sigma = 1.654. \quad (31)$$

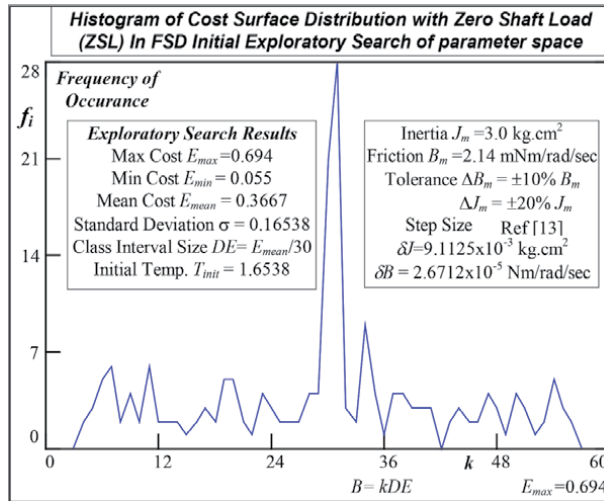
The anneal temperature scaling factor  $k$  is chosen [60] with a typical value of 10 on the basis of acceptance of parameter configuration costs, assumed normally distributed, worse than the present value by  $3\sigma$  with a successful Boltzmann jump probability of

$$P = \exp(-\Delta E/T_i) = \exp(-3\sigma/T_i) \approx 0.75 \quad (32)$$

such that

$$k = -3/\ln P = 10 \quad (33)$$

In this statistical gathering phase the best estimate  $\hat{\mathbf{X}}_{iopt}$  of the optimal parameter vector with cost  $\hat{E}_{iopt}$  is retained for initialization of the ensuing FSD algorithm if required. The cooling schedule is the most critical feature of the FSD method to guarantee global convergence and avoid trapping at local minima due to premature quenching. The update temperature algorithm consists of a sequence of temperature decrements and a condition to secure thermal equilibrium at each stage so that the average cost decreases in a uniform manner overall. The sequential temperature decrease



**Figure 32.**  
 Histogram of exploratory search costs.

$$T_{k+1} = \alpha_k * T_k \quad (34)$$

is based on the step reduction factor

$$\alpha_k = \exp\left(-\lambda \frac{T_k}{\sigma}\right) \quad (35)$$

with a typical value for  $\lambda$  of 0.8 and has been reported to work well in practice [24, 60] for a diversity of applications. The following numerical constants used in the FSD parameter extraction process have been obtained from heuristics and are known to give good performance of the algorithm with accurate estimates of global minimum convergence [60]. In practical applications the reduction process is typically lower bounded by

$$\min \alpha_k = 0.5 \quad (36)$$

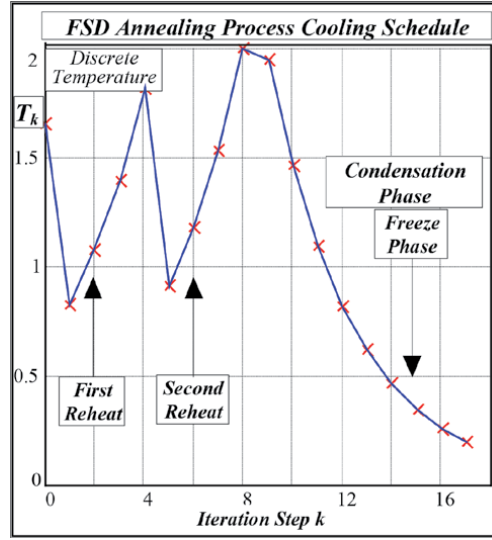
to prevent too rapid a reduction. Acceptance of large cost function jumps is a feature of the high temperature phase as in simulated annealing obviating the need for the ineffective gradient term in the iteration process and thus only a random search of parameter space is made initially for the first 10 external loops. At each temperature step the number of iterations performed is predefined at 15 times the dimensionality of parameter space (NDIM) to ensure equilibrium of the method. As the control temperature is reduced the volume of random search space is pruned, which is an attribute of the  $\sqrt{2T}$  term of the FSD equation in Eq. (17), in accordance with a temperature dependent multiplier proportional to

$$\left(\frac{T_{k+1}}{T_i}\right)^r, 0.5 \leq r \leq 1.0. \quad (37)$$

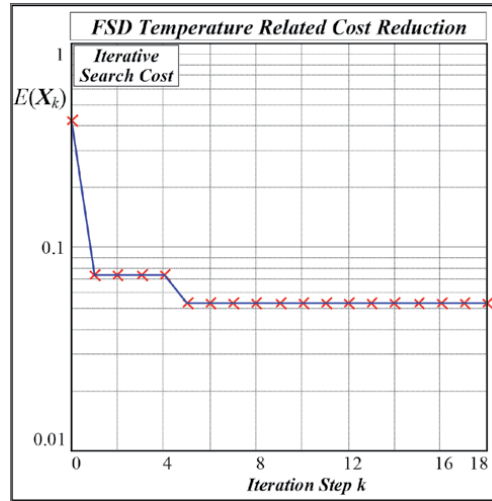
This enhances convergence towards the global parameter optimizer  $\mathbf{X}_{opt}$  and curtails non profitable random moves. Another useful feature of the FSD random search [24] is the application of a long tailed Lorentzian parameter distribution, which permits the occurrence of large cost jumps from local minima at low temperatures and consequent metastable state  $\hat{\mathbf{X}}_{opt}^i$  trap avoidance. At the termination stage of the FSD algorithm, when there is little observed change in the cost function over successive temperature steps, a controlled reheat phase is introduced with gradual temperature increase to reduce the risk of premature quenching and local minima trapping. The final phase of temperature decrease to the freeze condition results in convergence to the best approximation of the global minimiser.

## 9. Results obtained from FSD parameter identification

The extraction of the dynamical parameters for a typical brushless motor drive system with zero shaft load (NSL) inertia using the FSD method [26] is qualitatively illustrated in **Figures 33** and **34** with returned estimates summarized in **Table 10**. The application of a uniform random exploratory search of parameter space initially is rewarded by the provision of a good estimate of the optimal parameter set for initialization of the FSD algorithm. The returned estimate  $\hat{\mathbf{X}}_{iopt}$  in **Table 10**, which has a cost  $\hat{E}_{iopt}$  equal to 15% of the exploratory sample mean  $\bar{E}$ , in this case is very close to the best possible estimate  $\bar{\mathbf{X}}_{opt}$  of the optimal vector with the minimal



**Figure 33.**  
FSD cooling sequence record.



**Figure 34.**  
Iterative cost reduction sequence.

quadratic penalty  $\bar{E}_{opt}$  for the parameter tolerance range chosen. The best estimate  $\bar{\mathbf{X}}_{opt}$  for the optimal parameter set can be obtained by visual inspection from the experimental  $I_{fa}$  cost surface in **Figure 35** employing worst case quantized parameter step sizes in **Table 8**. This estimate  $\bar{\mathbf{X}}_{opt}$ , which will be referred to from here on as the ‘optimal’ vector for convenience and brevity of expression, is used as a reference against which the accuracy of the returned FSD parameter estimates can be judged. The proximity of  $\hat{\mathbf{X}}_{iopt}$  to  $\bar{\mathbf{X}}_{opt}$ , as determined from the error analysis in **Table 11**, is indicated by the degree of quantized resolution of parameter space with error differential

$$d\hat{\mathbf{X}}_{iopt} = [2\delta J^L, \delta B^L]^T. \quad (38)$$

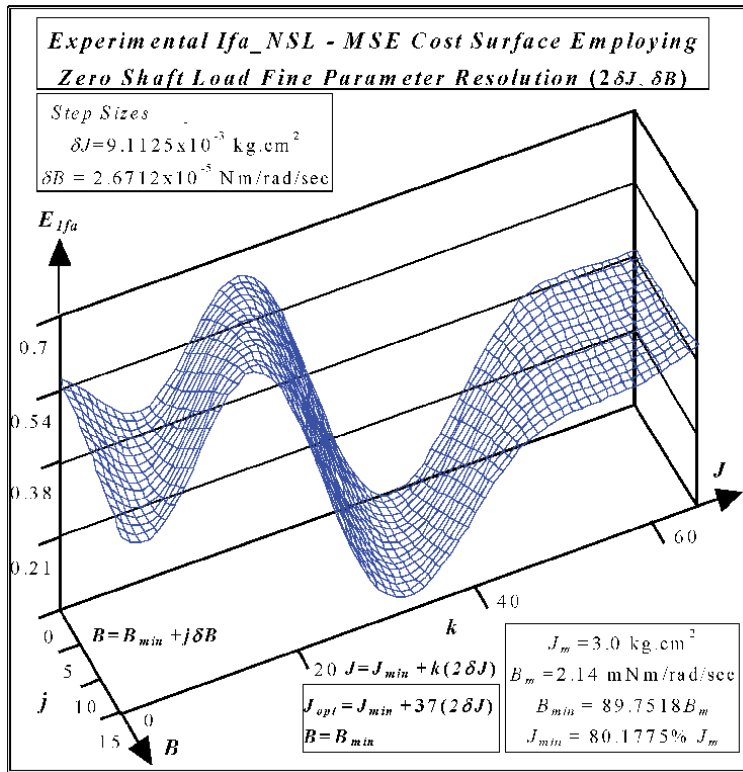
Optimal Estimates obtained via Experimental Cost Surface in <b>Figure 35</b>			
Quantized Parameter FC Response Surface Simulation Mesh Size $[2\delta J^L, \delta B^L]$			
$\bar{J}_{opt} = 3.0797 \times 10^{-4} \text{ kg}\cdot\text{m}^2$ $\bar{B}_{opt} = 1.9207 \times 10^{-3} \text{ Nm/rad/sec}$ $\bar{E}_{opt} = 5.0051 \times 10^{-2}$			
Initial Exploratory Search and Returned FSD Estimates - <b>Table 8</b>			
<i>Sample Statistics</i>	$N_i = 200$	$\sigma = 1.654 \times 10^{-1}$	mean $\bar{E} = 0.37$
$\hat{J}_{iopt} = 3.0638 \text{ kg}\cdot\text{cm}^2$ = 99.5% $\bar{J}_{opt}$	$\hat{B}_{iopt} = 1.953 \times 10^{-3} \text{ Nm/rad/sec}$ = 102% $\bar{B}_{opt}$	$\hat{E}_{iopt} = 5.473 \times 10^{-2}$ = 109% $\bar{E}_{opt}$	
Returned FSD Optimal Parameter Estimates			
FSD Initialization: $X_i$	$J_i = 82\% J_m$	$B_i = 109\% B_m$	
Observations			
After First Temperature Step $T_1$ :	→	Main Lobe Capture as in <b>Figure 36</b> $J_{opt}^{(1)} = 3.0341 \text{ kg}\cdot\text{cm}^2 = 98.5\% \bar{J}_{opt}$ $B_{opt}^{(1)} = 2.199 \text{ mNm/rad/sec} = 114.5\% \bar{B}_{opt}$ $E_{opt}^{(1)} = 7.346 \times 10^{-2} < \sigma$	
• After Fifth Temperature Step $T_5$ :	→	Cost Reduction Ceases	
• First Reheat Cycle Completed:	→	4 Temperature Steps {NF = 4} $J_{opt}^{(2)} = 3.07058 \text{ kg}\cdot\text{cm}^2 = 99.7\% \bar{J}_{opt}$ $B_{opt}^{(2)} = 1.9586 \text{ mNm/rad/sec} = 102\% \bar{B}_{opt}$ $E_{opt}^{(2)} = 5.295 \times 10^{-2} < E_{opt}^{(1)} < \sigma$	
• Total Number of Temperature Steps:	→	18 {First 10 used for Random Search}	
FSD Global Convergence Estimate: $\hat{X}_{opt} = \{J_{opt}^{(2)}, B_{opt}^{(2)}\}$			
Computational Details of FSD Algorithm in <b>Figures 28 and 29</b>			
No. of External Temperature Loops with Random Search only $m_j = 10$			
No. of Parameter Adjustments per Temp. Step to reach Equilibrium $u_j = 15\text{-NDIM} = 30$			
No. of MSE Evals. →	Random: 405	Gradient: 387	
Av. No. of BLMD Simul <sup>ns</sup> per Temp. Step: →	$(405+387)/18 = 44$		
Total Simulation Time for FSD Algorithm: →	20546 secs for a 486-DX-66 MHz CPU		
Execution Time per BLMD Simul. Trial: →	$20546/(405+387) \approx 26 \text{ sec}$		

**Table 10.**  
 Summary of returned FSD parameter estimates.

The effectiveness of the FSD method in achieving parameter ‘optimality’ over the FC undulating cost surface is demonstrated by deliberately initializing the search far from the global minimum in the neighborhood of a local minimum at  $X_i$  as given in **Table 10**.

A scatter diagram of iterative search costs levied by the FSD process is portrayed in **Figure 36** and contrasted with a cross section of the FC cost surface shown in **Figure 37** for variable  $J$  and fixed  $B = \bar{B}_{opt}$ . Application of a Boltzmann probability step transition in surmounting significant cost barriers, with cost elevations of  $2\sigma$  above  $\bar{E}$  which conceal the global minimum, is evident in the iterative search improvement process for a global optimum. Thus a means of escape from local minimum capture, which causes problems for traditional identification methods with non-optimal convergence, is provided for parameter extraction in an effort to secure the least mean squares (LMS) estimates.

In the early stages of the cooling schedule the temperature reduction is matched by a rapid reduction in iterative cost which eventually saturates as the parameter estimates approach the optimal LMS values. After the first temperature step the



**Figure 35.**  
MSE cost surface without shaft load.

BLMD Parameter	Motor Inertia $J_{opt}$ $\left(\frac{J_{opt} - J_{opt}}{J_{opt}}\right) \times 100\%$	Damping Factor $B_{opt}$ $\left(\frac{B_{opt} - B_{opt}}{B_{opt}}\right) \times 100\%$	Global Cost $E_{opt}$ $\left(\frac{E_{opt} - E_{opt}}{E_{opt}}\right) \times 100\%$
Exploratory Search	-0.515%	1.682%	9.348%
<i>FSD Method</i>			
Post Temp Step $T_1$	-1.479%	14.49%	46.77%
Post Temp Step $T_5$	-0.296%	1.976%	5.808%
Coherence of BLMD Simulation Waveforms with Experimental Test Data			
NSL - Correlation Coefficient $\rho$	Current Feedback $I_{fa}$ 94.62%	Current command $I_{da}$ 95.815%	Cur. Controller o/p $V_{ca}$ 94.9%
Shaft Velocity Correlation Coefficient $V_{or} = 97.9\%$ .			

**Table 11.**  
Error Analysis for Returned FSD Parameter Estimates.

estimation cost, which is less than the exploratory  $\sigma$ , is captured within the main lobe of the response surface containing the global minimum. Subsequent progress in cost improvement is minimal which is mainly due to the fact that the parameter estimate at  $\mathbf{X}^{(1)}_{opt}$  is relatively close to the global minimizer and that random searching is employed for the first ten temperature steps without the assistance of a gradient search. In addition to this the selectivity of the cost surface to  $B$  parameter variation is poor along the valley floor, which is flat near the global minimum as illustrated in **Figure 35** as per [13, 28], making it difficult for global minimum convergence.

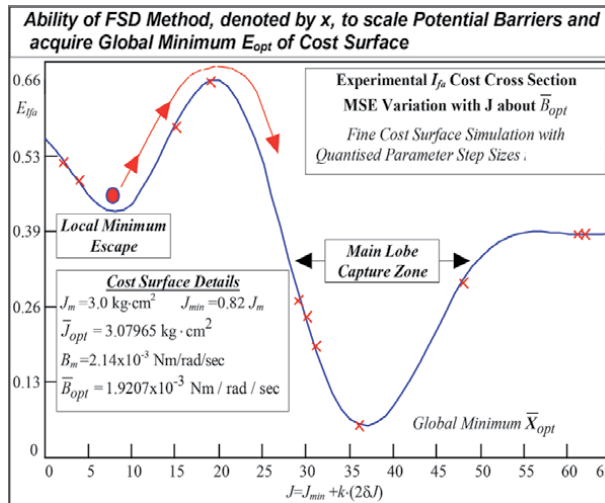


Figure 36. FSD parameter search with local minimum escape.

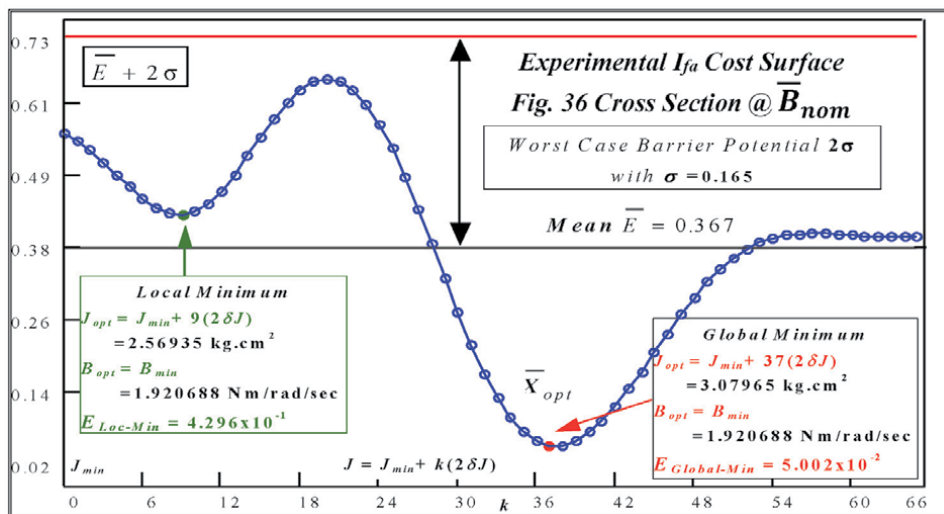


Figure 37. Acquisition of Global Minimum from experimental  $I_{fa}$  cost surface.

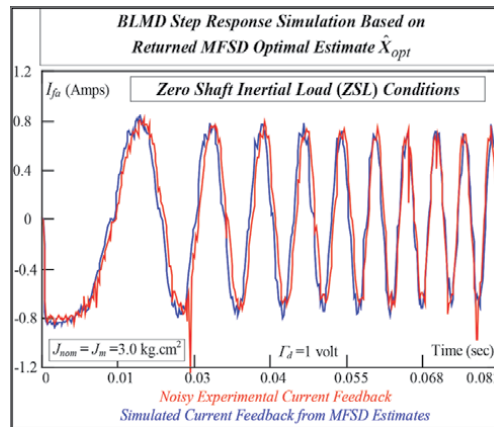
During this phase of cost immobility a reheat cycle is introduced [24], as part of the FSD cooling schedule shown in Figure 29 to overcome local minimum trapping, which is irrelevant in this case as the current estimate is circumjacent the optimal value. After the fifth iteration further cost reduction ceases with the global convergence estimate given by

$$\hat{\mathbf{X}}_{opt} = \mathbf{X}_{opt}^{(2)} \quad (39)$$

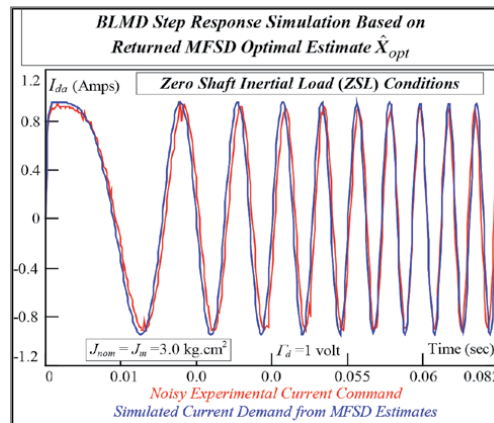
A second reheat phase followed by freeze conditions does not improve the parameter estimate. The total execution time on a 486-DX-66 MHz processor for the FSD method in this parameter identification procedure amounted to 20546 seconds with a computation burden of 792 cost function evaluations over 18 temperature steps. This translates on average into 44 BLMD model simulation runs per

temperature step with a processor runtime of 26 seconds per simulation trial. This puts into perspective the computational intensity and the cost in terms of CPU time of this optimization technique. A breakdown of the overall search effort deployed in the FSD process projects into 405 random search calculations with a punitive computational overhead of 387 evaluations necessary for anticipated downhill movement factored in only for the last 9 temperature steps. This overhead is manifested, after the random search epoch during which the parameter estimates have reached a minimum potential, in the lengthy reheat and condensation phases shown in **Figure 33**. The inadequacy of the gradient search procedure in this instance, with approximate global convergence already achieved, is compounded by the tendency of the FSD method to dither about in a region of parameter space where the response surface has poor selectivity in the  $B$  parameter with little iterative progress in cost reduction expected. The comparison of BLMD step response simulation with experimental drive test data, at internal observation nodes based on returned FSD optimal estimates, provides an excellent fit in terms of frequency and phase coherence as per **Figures 38–40** pertaining to BLMD current control.

These simulated step responses are almost identical to those based on returned estimates for the modified form of the FSD method (MFSD), with zero inertial load conditions (NSL), discussed in Section 11 below. These simulation traces are

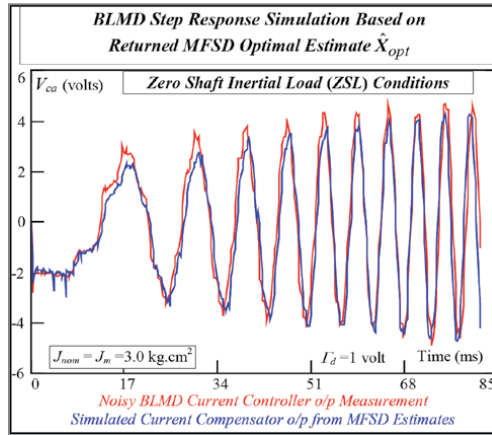


**Figure 38.**  
ZSL FC Simul<sup>n</sup> via MFSD estimates.

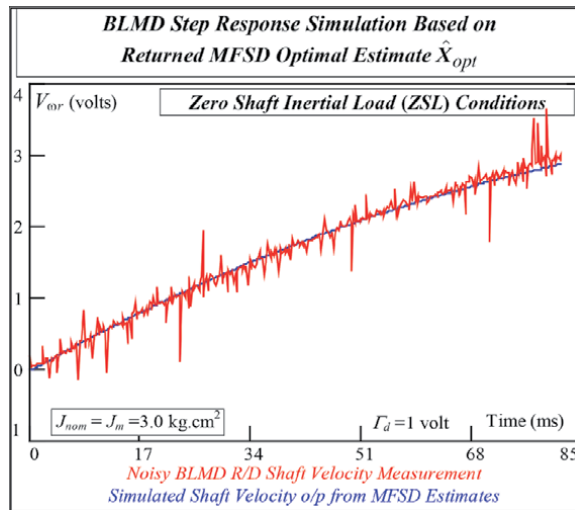


**Figure 39.**  
ZSL- current demand Simul<sup>n</sup>.

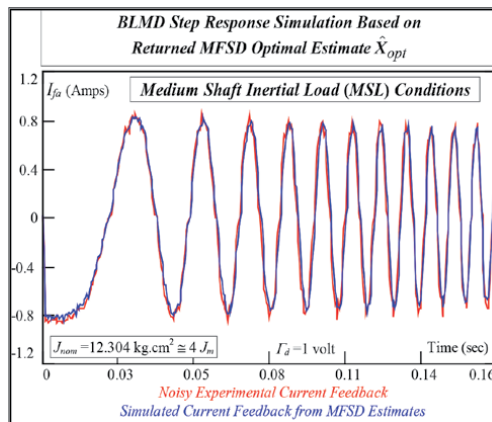




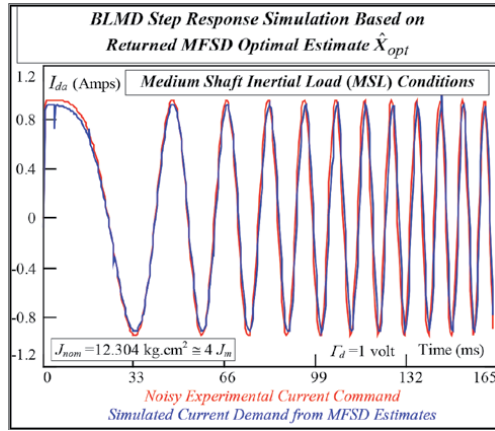
**Figure 40.**  
 ZSL-current controller Simul<sup>n</sup>.



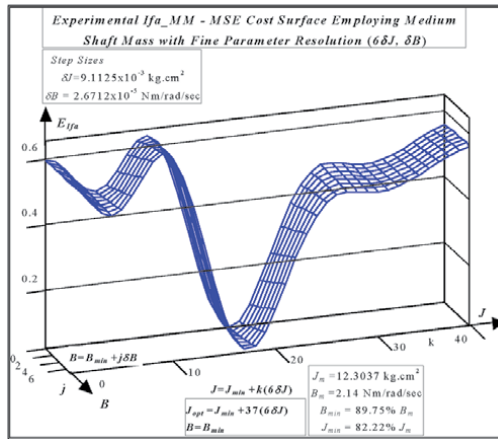
**Figure 41.**  
 ZSL – MFSD Shaft Velocity Simulation.



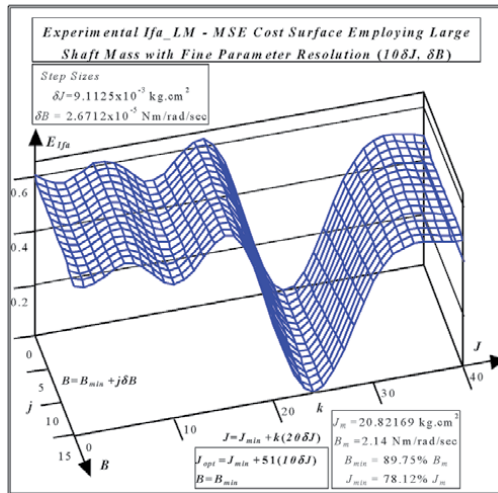
**Figure 42.**  
 FC simulation via MFSD estimates.



**Figure 43.**  
 MSL - current demand simulation.

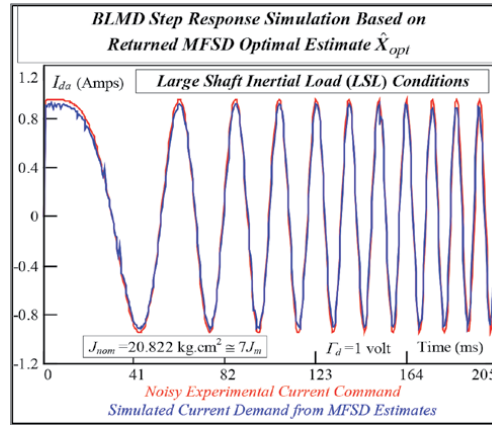


**Figure 44.**  
 Current controller o/p simulation.

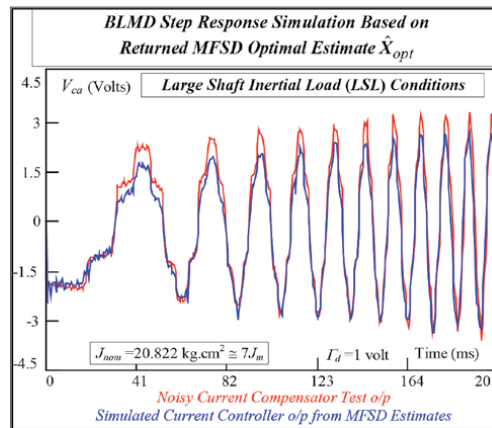


**Figure 45.**  
 FC-LSL Simulation via MFSD Estimates.

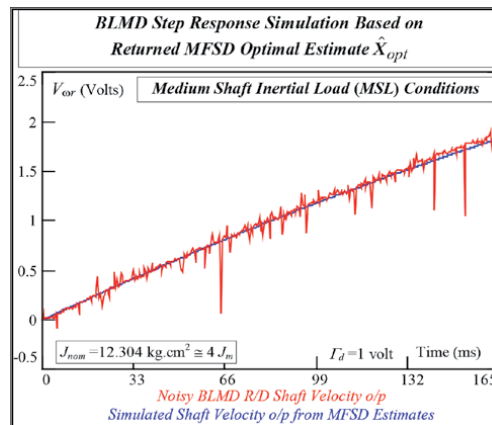
displayed together along with other waveforms for MFSD comparison purposes in **Figures 38–41** and **Figures 42–49** at different shaft inertial load bearing conditions. The goodness-of-fit measure is determined by the degree of correlation, expressed



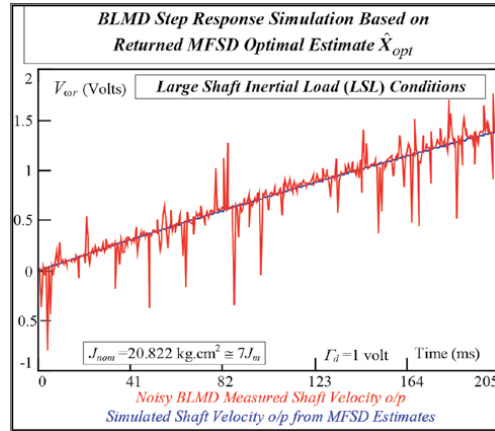
**Figure 46.**  
LSL - current demand simulation.



**Figure 47.**  
LSL - current controller simulation.



**Figure 48.**  
MFSD – MSL Shaft Velocity Simulation.



**Figure 49.**  
MFSD – LSL - Shaft Velocity Simulation.

as the correlation coefficient in **Table 11**, which exists between the simulated and experimental test data. This value is almost 95% for all three internal node observations which confirms the success of the FSD method as an accurate method of parameter extraction besides providing BLMD model validation. Further confidence in the FSD method is assured by the degree of correlation in **Table 11** of BLMD shaft velocity simulation with test data as shown in **Figure 41**.

## 10. Improved FSD method for motor parameter identification

A drawback with the general form of the FSD method described in Section 8 is the usage of a fixed number of external loops at high temperature for random searching only of parameter space without application of the gradient technique. The lack of adaptation of the method in its present form to the iteration statistics garnered during the initial high temperature phase gives rise to lengthy run time of the FSD algorithm. This results in an excessive number of cost calculations, which can be very expensive in central processing unit (CPU) time resources, accompanied by at least one reheat phase before termination of the FSD method without cognisance of the gains made at each temperature step and the shape of the cost surface based on motor feedback current.

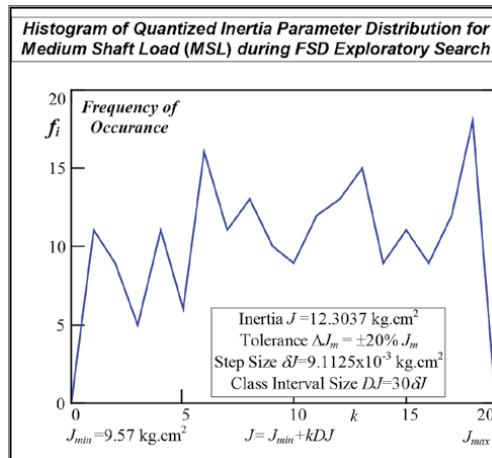
It is evident from **Figure 34** that significant reductions in the objection function are obtained during the first couple of high temperature steps only, with very little improvement thereafter. A more lucrative strategy [27] can be based on adaptation of the FSD method during the high temperature random search phase in the absence of good initialization of the parameter vector and on the initial exploratory search statistics where the maximum value  $E_{max}$ , mean  $\bar{E}$  and standard deviation  $\sigma$  of the resultant cost distribution are known.

The initial exploratory search statistics, for three known cases of motor shaft inertia to be ‘identified’, are listed in **Table 12**. The histograms associated with the initial exploration of inertial parameter space, along with those for the related costs, are shown in **Figures 50–53**. These histograms possess the same uniform randomness attributes as that for the zero shaft load inertia case discussed in Section 8 with cost distribution depicted in **Figure 32**.

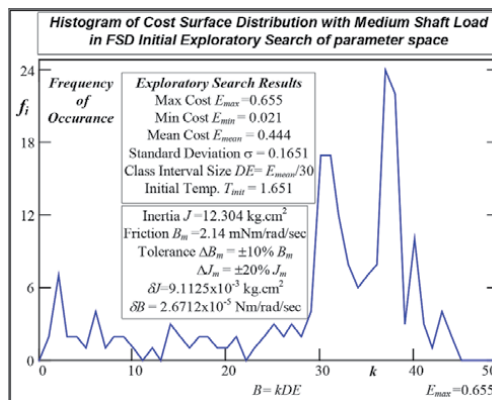
It is evident that all the tabulated exploration costs fall within  $\pm 3\sigma$  of the sample mean  $\bar{E}$  with worst case estimates  $E_{max}$  less than  $\bar{E} + 2\sigma$ , for different shaft load inertia as illustrated in **Figure 37**, which according to Chebyshev’s theorem of

Total Nominal Inertial Shaft Load $J_{nom}$ (kg.cm <sup>2</sup> )	No Shaft Load NSL - (0.0)	Medium Shaft Load MSL - (9.304)	Large Shaft Load LSL - (17.822)
	3.0	12.304	20.822
Tolerance: $\pm\Delta X_{nom} = [\pm 20\% J_{nom}, \pm 10\% B_{nom}]$		Sample Size: $N_i = 200$	
<i>Quantized Parameter Space:</i>			
Sample Percentage	9.542%	2.315%	1.19%
Max Cost $E_{max}$	0.694	0.655	0.653
Mean Cost $\bar{E}$	0.367	0.444	0.457
Min Cost $\hat{E}_{iopt}$	$5.473 \times 10^{-2}$	$2.054 \times 10^{-2}$	$1.497 \times 10^{-2}$
Standard Deviation $\sigma$	$1.654 \times 10^{-1}$	$1.651 \times 10^{-1}$	$1.694 \times 10^{-1}$
<i>Initial Optimal Estimates</i>			
$\hat{J}_{iopt}$ (kg.cm <sup>2</sup> )	3.064	12.121	20.767
$\hat{B}_{iopt}$ (Nm.rad <sup>-1</sup> .sec)	$1.953 \times 10^{-3}$	$1.953 \times 10^{-3}$	$2.033 \times 10^{-3}$

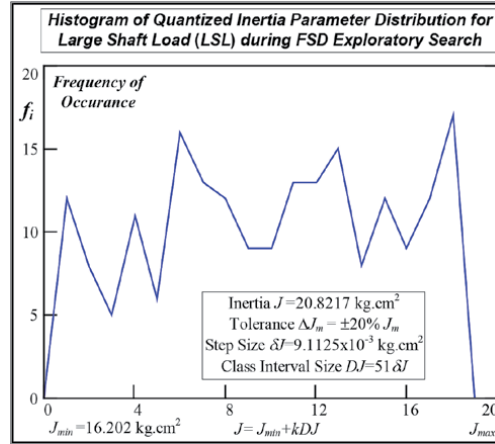
**Table 12.**  
 Statistical estimates for initial exploratory search of parameter space.



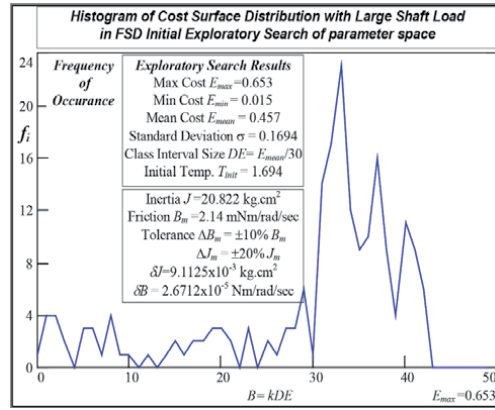
**Figure 50.**  
 Histogram of MSL-J search values.



**Figure 51.**  
 Histogram of MSL exploratory costs.



**Figure 52.**  
Histogram of LSL-J Search Values



**Figure 53.**  
Histogram of LSL exploratory costs.

measurements [13, 28, 61] should account for at least 89% of all costs. Furthermore exploratory parameter estimates  $\bar{\mathbf{X}}_{iopt}$  with costs less than the rms deviation  $\sigma$  about the mean are very close to being optimal and are in the ambit of global minimum convergence given the shape of the FC response surface. Such estimation costs, which are less than  $\bar{E} - \sigma$ , are confined to the main lobe of the objective function which has for an apex the global minimizer  $\bar{\mathbf{X}}_{opt}$ . For the purpose of validation and demonstration of the effectiveness of the MFSD method as a parameter extraction tool in identifying various motor shaft inertial loads the exploratory least squares parameter estimates in **Table 12** are ignored in the initialization process.

Instead the method is deliberately initiated far from the global estimates in all three cases at

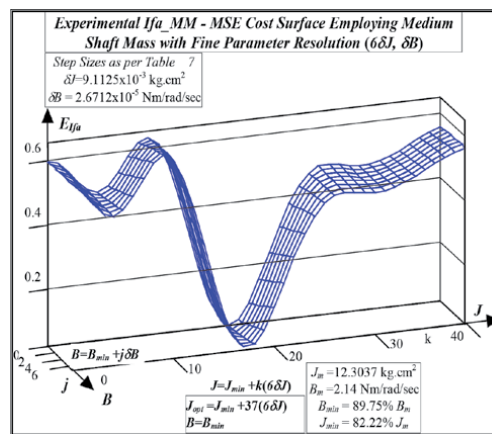
$$\mathbf{X}_{init} = [82\%J_{nom}, 109\%B_m]^T, \quad (40)$$

based on parameter tolerance bounds given in **Table 12** with initial temperature given by

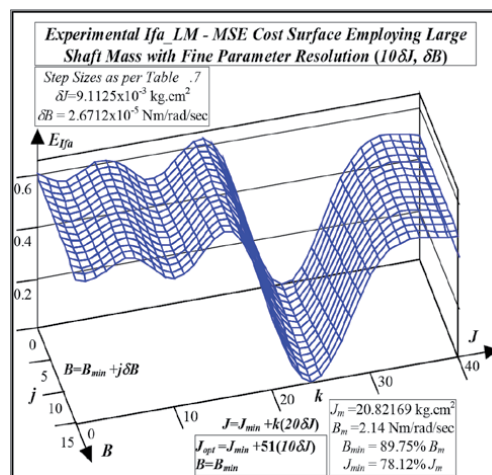
$$T_i = T_{init} = 10\sigma, \quad (41)$$

where the objective function value is large and in the vicinity of a local minimum. The best optimal parameter estimates are obtained separately by inspection from quantized parameter FC cost surface simulations, utilizing the same experimental FC target data, for comparison purposes with the MFSD method as shown in **Figures 35, 54** and **55** for different inertial loads. The modified version of the FSD algorithm uses a discretized parameter space with step sizes as listed in [13] and incorporates the following improvements:

- a. proceed with normal random probing of parameter space for the first high temperature step as in **Figures 28** and **29** with the best parameter estimate retained. At the end of this iterative sequence a ‘greedy’ search is performed, from the current best estimate obtained thus far, along each of the parameter directions in turn using the gradient method with a quantized parameter step size to take advantage of any possible further cost improvement that may accrue. The resultant optimal parameter estimates are trimmed to the nearest step size.



**Figure 54.**  
 MSE surface with medium shaft load.



**Figure 55.**  
 MSE surface with large shaft load.

Total Nominal Shaft Load Inertia $J_{nom}$ (kg.cm <sup>2</sup> )	Figure 35 - NSL No Shaft Load $J_m = 3.0$	Figure 54 - MSL Medium Inertia 12.304	Figure 55 - LSL Large Inertia 20.822
Parameter Initialization: [ $J_i = 0.82J_{nom}, B_i = 1.09B_m$ ] with $B_m = 2.14 \times 10^{-3}$			
Initial Cost $E_i$	0.414	0.590	0.616
Returned FSD Optimal Parameter Estimates			
$\hat{J}_{opt}$ (kg.cm <sup>2</sup> )	3.07965	12.13966	20.8375
$\hat{B}_{opt}$ (Nm.rad <sup>-1</sup> .sec)	$1.95863 \times 10^{-3}$	$2.00082 \times 10^{-3}$	$1.95863 \times 10^{-3}$
Min Cost $\hat{E}_{opt}$	$5.21677 \times 10^{-2}$	$2.20282 \times 10^{-2}$	$1.30295 \times 10^{-2}$
No. of Temp Steps to reach $\hat{E}_{opt}$	2	2	2
No. of Func_Evals	110	109	103
Random Search	45	45	45
Gradient Search	65	64	58
Total Time (sec)	2834	5562	6499
Average Iteration Time $t_{ITER}$	1417	2781	3250
Simul <sup>n</sup> Time/Func. Eval. $t_{SIM}$	25.76 s	51.03 s	63.10 s

**Table 13.**

Modified FSD method of motor shaft inertial load parameter extraction.

- b. If the current Least Mean Squares (LMS) error is less than  $\sigma$  then the alternate application of hill descent with random search is pursued during subsequent temperature steps in the iterative improvement process towards optimality. This approach is adopted because the best estimate available with cost  $\sigma$  most likely resides within the capture zone of the main cost lobe as depicted in **Figure 37** for example. If the converse is true then random searching is continued as in step (a) until sufficient cost reduction has been achieved.
- c. A gradient search at the termination of each temperature step is maintained for further gain in cost reduction. This has the effect of forcing and hastening the convergence of the parameter estimates towards optimality. If however there is no cost improvement for several temperature steps, typically 3 and pending step (b), the gradient search is ceased and the FSD process enters the termination stage after the fourth step which corresponds with the onset of the freeze condition in the normal FSD method. The termination process can be speeded up by halting the algorithm at this point without the necessity of the freeze condition as subsequent parameter convergence information bears out in **Table 13**. This three step stalled cost reduction phase allows for improved convergence in the  $B$  parameter estimate along the valley of the response surface, where it is very flat in this parameter direction, without undue expenditure of computational effort. Secondly the anneal temperature rises during this interregnum, in line with the reheat phase of the FSD method, nullifying the possibility of false minima trapping of parameter estimates.

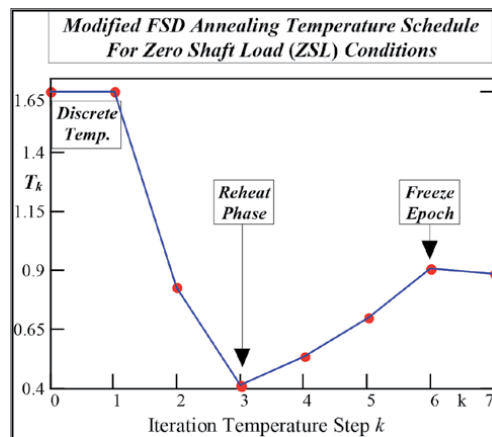
Once the motor friction coefficient  $B$  has been resolved further pruning of CPU run time can be availed of if only the inertial parameter is to be identified, for which the FC response surface is more sensitive near the global minimum, by abrupt termination of the FSD method after two non profitable temperature steps.



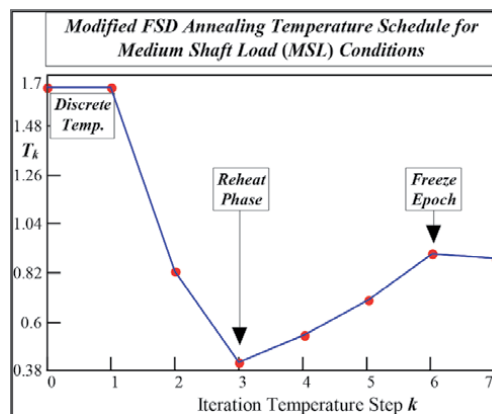
## 11. Parameter convergence results for the modified FSD method

The optimized parameter estimates of the motor dynamics, returned by the modified FSD method [25], with convergence details for three cases of shaft load inertia to be identified are summarized in **Table 13**. The number of iterative temperature steps and functional evaluations required, for the motor parameter identification process to reach minimum potential, in each of the three cases of shaft load inertia are almost identical for the modified FSD method with considerable savings effected in computational effort over the conventional FSD approach.

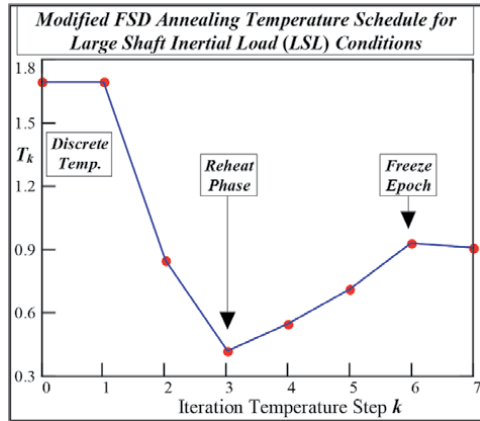
The anneal temperature profiles realized during the MFSD extraction process of the three shaft inertial loads are depicted in **Figures 56–58**. Global optimality, to within the limits of the quantized parameter step sizes listed in [13], is achieved in all three cases in at most two temperature steps as shown in **Figures 59–61**. The CPU runtime however increases with the mechanical time constant  $\tau_m$  given in [9], for a similar number of motor simulation trials as indicated in **Figure 62**, in sympathy with the time duration of the observed FC target data used as an argument in the objective function formulation.



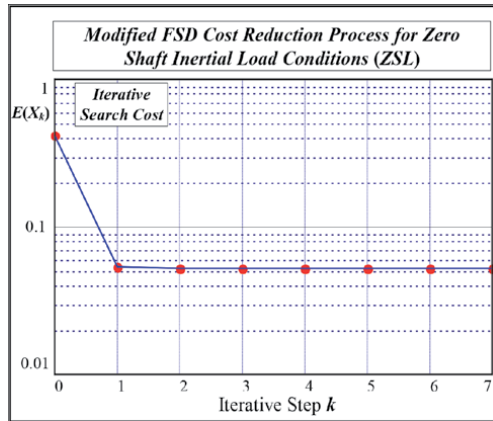
**Figure 56.**  
*NSL cooling temperature history.*



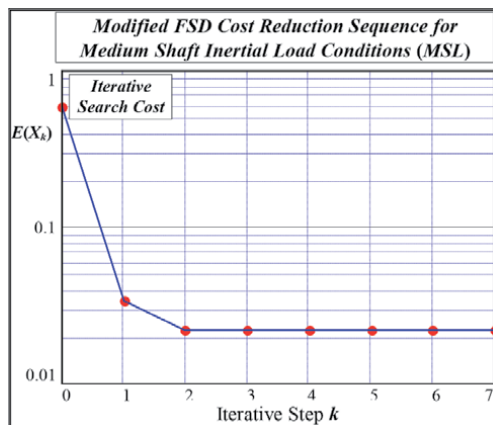
**Figure 57.**  
*MSL cooling temperature history.*



**Figure 58.**  
MSL cooling temp. Sequence.

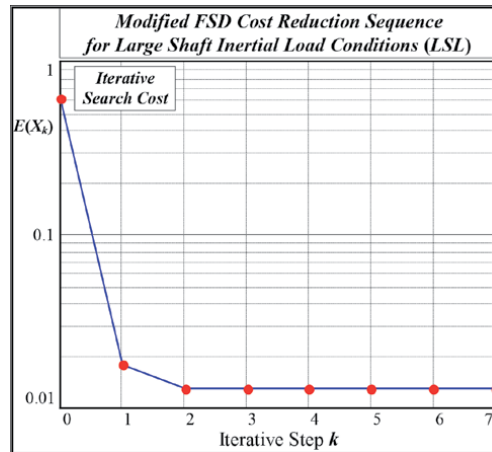


**Figure 59.**  
Iterative NSL cost reduction.

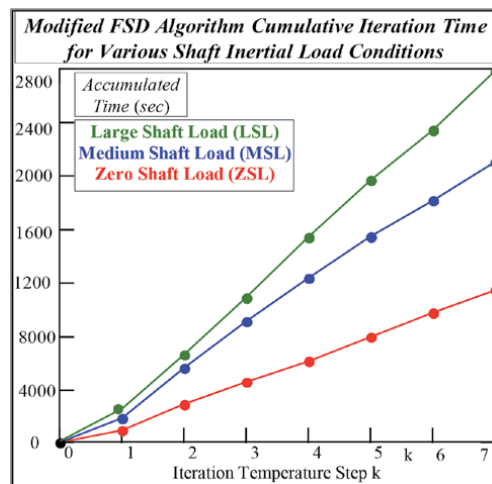


**Figure 60.**  
Iterative MSL cost reduction.

While in general increasing the number of FC cycles improves the accuracy of the extracted global parameter estimates, measurement constraints limited the size of the FC transient data record to 10 cycles used in the MSE cost formulation. The



**Figure 61.**  
 Iterative LSL cost reduction.



**Figure 62.**  
 MFSD computation time.

length of each data training record is fixed at 4095 sample points ( $N_d$ ) with a normalized time duration of approximately 10 machine cycles for reference purposes and response surface comparison with details given in **Table 3**.

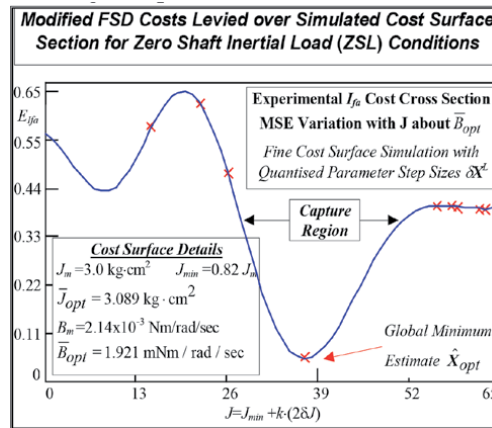
Response surface simulation, although computationally expensive, provides an alternative route of accurately obtaining the optimal parameter vector by means of inspection of the surface minimum cost. It can thus be used as a yardstick by which the overall convergence performance of the MFSD method can be contrasted over a range of motor shaft load inertia. The cost simulations are based on an initial crude parameter mesh size given in **Table 3** with a  $1 \mu\text{s}$  time step  $\Delta t$  in all cases as shown in **Figures 35, 54** and **55** with further refinement down to quantized step sizes necessary in the vicinity of the global minimum for resolution of the optimal parameter set.

The optimal shaft inertia values extracted by this approach are used in the error analysis given in **Table 5** as a reference by which the accuracy of the FSD optimal estimates are gauged. The parameter estimates extracted by the FSD optimization technique are very accurate in the shaft inertia only with fractional

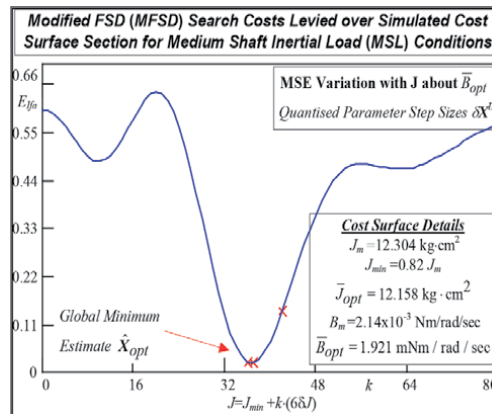
percentage relative errors achieved in global convergence performance as indicated in **Table 5**.

The errors appear to decrease with increasing inertial load coupled with a general trend in the global cost reduction. The error performance in **B** parameter estimation is ‘poor’ by comparison and is responsible for the increase in relative global cost in column 2 of the error summary. The wider error margins can be inferred from the deployment of the larger quantized relative step size listed in [13] and attributed to poor selectivity of the cost surface, in the friction coefficient, in the region of the global minimum as discussed in [13]. The quality of the MFSD probe of parameter space in obtaining the global estimates and evading local minimum capture is demonstrated in **Figures 63–65** by the asterisked search costs over cross sections of the simulated FC response surfaces for contrast.

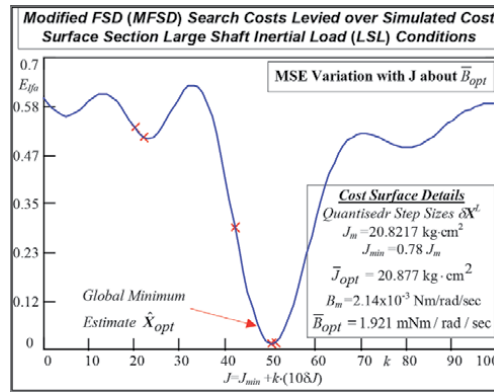
The accuracy of the returned MFSD estimates can also be checked by employing these optimal parameter vectors in BLMD dynamical simulation and resultant comparison with experimental motor drive step response data. The model simulation traces are compared with observed test data, on the basis of accuracy of fit at critical internal nodes of the drive system, for the dual purpose of validation of the MFSD identification method and enhancement of BLMD model confidence which are essential intrinsic component features of system identification [35].



**Figure 63.**  
FSD identification of rotor dynamics.



**Figure 64.**  
FSD identification of inertial MSL.



**Figure 65.**  
 FSD identification of inertial LSL.

Nominal Shaft Inertial Load $J_{nom}$ (kg·cm <sup>2</sup> )	No Shaft Load (NSL) $J_m = 3.0$	Medium Inertia (MSL) 12.304	Large Inertial Load (LSL) 20.822
Current Feedback $I_{fa}$	<b>Figure 38</b> 94.6%	<b>Figure 42</b> 98.6%	<b>Figure 45</b> 99.4%
Current command $I_{da}$	<b>Figure 39</b> 95.8%	<b>Figure 43</b> 97.8%	<b>Figure 46</b> 99.7%
Current Controller o/p $V_{ca}$	<b>Figure 40</b> 94.9%	<b>Figure 44</b> 99%	<b>Figure 47</b> 98.7
Motor Shaft Velocity $V_{\omega r}$	<b>Figure 41</b> 97.9%	<b>Figure 48</b> 99%	<b>Figure 49</b> 93.8%

**Table 14.**  
 BLMD simulation trace coherence correlation coefficient  $\rho$  in [13].

The goodness of fit of the simulation traces, at various BLMD model observation nodes in [13] in terms of frequency and phase coherence, with sampled test data is obvious from **Figures 38–40** and from **Figures 42–47** for BLMD feedback current, current demand and compensator outputs at different inertial loads. This measure of trace coherence, indicative of MFSD parameter extraction accuracy, can be gauged by the excellent correlation coefficient for the fixed amplitude swept frequency waveforms listed in **Table 14 (Figures 38–41, 42–49)**. The accuracy of the MFSD method is further substantiated in **Figures 41, 48** and **49** by the correlation of the model shaft velocity step response with BLMD output test data.

## 12. FC response surface selectivity with MFSD application

The above application of the modified version of the FSD method in inertial mass  $J$  identification evolved from the nature of the FC objective function deployed. The fixed sample number  $N_i$  (200) of random searches during the exploratory phase, or the first temperature step employing 15\*NDIM random searches with user supplied initialization only and an absent exploratory phase, tended to give less coverage of quantized parameter space for the cases with increased shaft inertia presented in **Table 12**. It therefore seems more difficult to find a good starting vector, as it takes longer to probe a sufficient volume of parameter space, from which to anchor all subsequent searches. The modified method would, it appears,

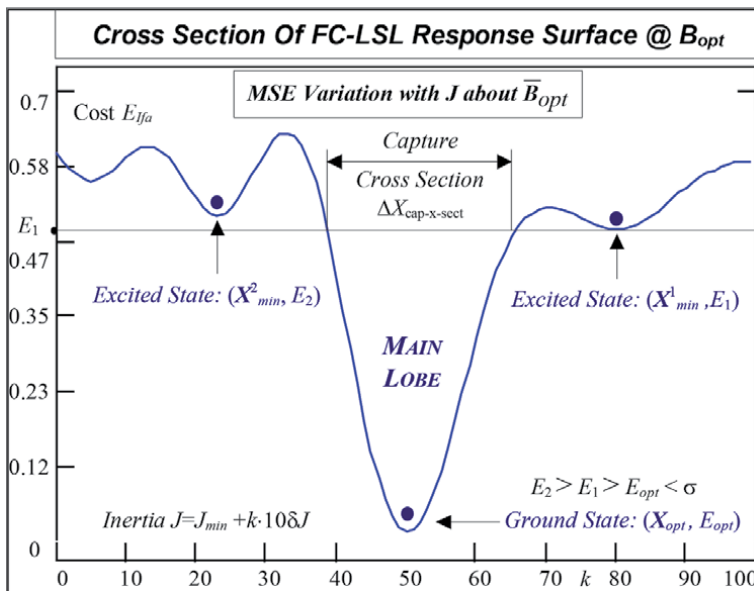
then terminate prematurely if trapped in a false minimum potential well after four anneal steps without adequate sampling of the parameter domain thus generating uncertainty as to the quality of returned estimates. However this is not the case when the topography of the cost surface in the observed FC variable is examined for different shaft inertial loads.

The number of feedback current cycles is fixed, at about ten in **Table 3**, as this influences the FC response surface “nuclear capture cross section” [62] shown in **Figure 66**. This can be defined as the waist of the main lobe containing the global minimum ‘energy’  $E_{opt}$ , with value less than the exploratory search  $\sigma$ , and the penalty costs of all excursions into the parameter space kernel forming a nucleus about  $X_{opt}$  with values less than the next excited level  $E_1$  or local minimum above the ground state  $E_{opt}$ . The probability of capture  $P_c$  of a random search can be formulated as the ratio of the capture cross section to the tolerance band of parameter variation in [13] about the nominal value  $x_m$  as

$$P_C = \frac{\Delta x_{cap-x-sect}}{x_{max} - x_{min}} = \frac{tdx^L}{2mdx^L} = \frac{t}{2m} \quad (42)$$

where it is assumed that the least important  $B$  parameter value has been resolved. Thus for a uniform random search the impingement of target parameter space, given by  $NP_c$ , increases directly with the number of trials  $N$ . Once the barrier potential  $E_1(X)$  to the kernel of capture is breached at any stage during the MFSD iterative improvement process, the related parameter vector is then installed as the best estimate by which all subsequent searches are adjudicated on for any further cost decrease. At the end of an anneal step a gradient search is performed which forces the best estimate to date, if trapped in the main lobe region, to converge towards the global minimum.

The capture cross section percentages, expressed in terms of allowed parameter tolerance and quantization step size, for the three cases of motor shaft inertia are presented in **Table 15** along with details of search cost capture during



**Figure 66.**  
FC cost surface ‘nuclear’ capture zone.

Capture Cross Section at $B_{opt}$	No Shaft Load Figure 35 38.97%	Medium Inertia Figure 54 30.15%	Large Inertia Figure 55 29.69%
FSD Target Acquisition			
Exploration Phase Trapped States %	127 out of 200 63.5%	85 out of 200 42.5%	71 out of 200 35.5%
First Anneal Step Random Search Only: %Trapping	15 ex 30 50%	16 ex 30 53.33%	12 ex 30 40%
First 2 Temp Steps Random + Mixture: %Capture	81 out of 110 73.64%	75 out of 109 68.81%	78 out of 103 75.73%

**Table 15.**  
 Estimation of % capture cross section from simulated FC cost surfaces.

implementation of the modified FSD method. The tabulated cross section percentages stabilize at about 30% for medium to large mass loading which shows that the probability of reaching the target zone is independent of the number of searches conducted owing to the shape of the optimal cost surface sections employed. This removes any uncertainty as to the effectiveness of the FSD method of entering the global region and possible lockup in an excited state in the event of a short anneal sequence where a large quantized parameter space is concerned.

Also the various FSD entrapment estimates at different phases of the search algorithm provide a rough independent confirmatory measure of the probability of capture where a uniform random search of parameter space is conducted. The percentage of all search excursions into the optimal parameter nucleus, in both the optional exploration phase and the mandatory random initial anneal step, exceed the cross section estimates in all cases and thus enhances user confidence in the modified method. The rationale behind this approach is based on the FC cost surface topography, which has a ravine like capture region shown in [13], with both  $J$  and  $B$  parameters considered. This target zone, which has essentially parallel contour lines in the direction of the ‘line minimum’ in [13], can be condensed into a single dimension in the percentage ratio carve up of parameter space for target acquisition calculations in the  $J$  parameter as shown **Figure 66**.

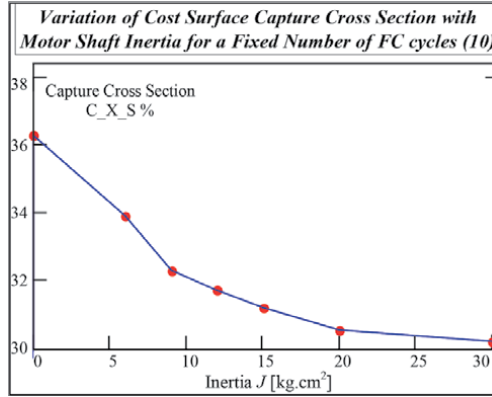
Further confidence in the capture cross section estimate for different shaft inertial loads can be gained by employing fixed length data records in the observed current feedback  $I_{fa}$ . This is based on ten FC machine cycles as the target reference for FC cost surface simulation in the  $J$  parameter at  $B_{opt}$  for estimation purposes as summarized in **Table 16** with data sampling rates  $f_s$  and displayed in **Figure 67**.

The probability of capture can be more accurately defined by the weighted contribution of all tabulated cross section estimates at different inertial loads as

$$P_C = \frac{\sum_i J_i * C\_X\_S}{\sum_i J_i} = 0.31. \quad (43)$$

$I_{fa}$ [ $f_s$ ]	50 kHz	37 kHz	31 kHz	27 kHz	24 kHz	21 kHz	17 kHz	15 kHz
Inertia kg.cm <sup>2</sup>	3.0	6.0	9.0	12.0	15.0	20.0	30.0	40.0
C_X_S	36.3%	33.9%	32.3%	31.7%	31.2%	30.5%	30.2%	30.3%

**Table 16.**  
 % capture cross section ( $C\_X\_S$ ) estimates from  $I_{fa}$  target data.



**Figure 67.**  
*C\_X\_S variation with inertia.*

The adoption of a fixed threshold step size  $\delta J^L$  during MFSD parameter extraction results in improved parameter resolution, given by the reduction in the variability estimate as

$$V_x^L = \frac{\delta x^L}{x}, \tag{44}$$

with increased shaft inertial loads as shown in **Table 9**.

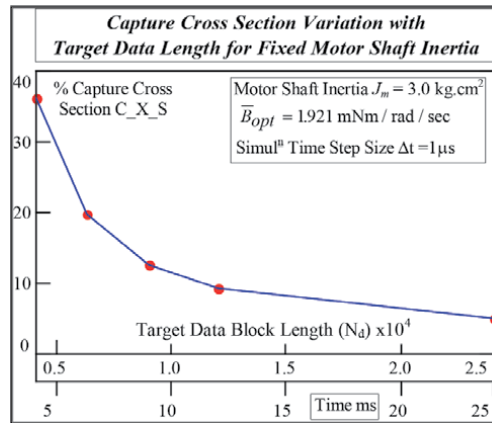
This percentage reduction results in better accuracy of the returned  $J$  parameter values for increased shaft inertial loads with smaller relative errors as indicated by error analysis in **Table 5**. The selectivity of the FC objective function in parameter identification can be improved somewhat by increasing the length of the  $I_{fa}$  sampled data record used as the target reference in the least squares cost formulation. The convergence domain is reduced with the data record length in an inverse time relationship thus improving the accuracy of the returned parameter estimates. This is articulated by the accompanying reduction in capture cross section thus restricting the search for optimality to a smaller kernel within the parameter nucleus surrounding  $X_{opt}$ . The raison d'être is the eclipsing of the finite length motor speed exponential ramp up transient by the ever lengthening presence of steady state speed saturation. The capture cross section variation for different simulated data file lengths, with a fixed sampling rate of 50 kHz and  $\pm 20\%$  parameter tolerance, is illustrated in **Table 17** for nominal rotor inertia  $J_m$  anchored at  $\bar{B}_{opt}$ .

The accompanying graph in **Figure 68** shows that the capture cross section of the optimal parameter kernel decreases with increasing data block size and ultimately stabilizes to a constant value for modest lengths which is indicative of the fact that the motor has reached maximum shaft speed. The increased selectivity is accompanied by a plurality of excited states, shown in **Figure 69**, which eventually level off in number as the shaft speed reaches its maximum value. This

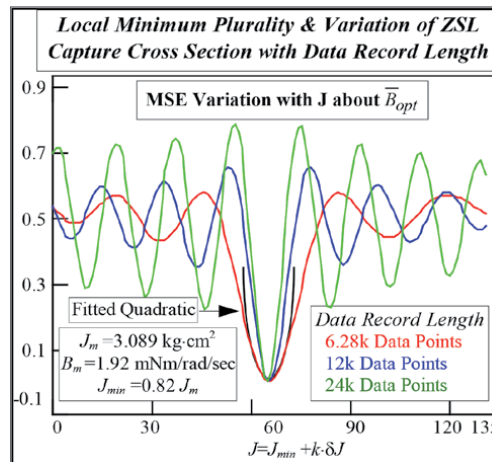
Parameter Vector Fulcrum: $X = [J_m, \bar{B}_{opt}]^T = [3.0 \times 10^{-4}, 1.921 \times 10^{-3}]^T$					
No. Samples	4120	6280	9000	12000	24000
Shaft Speed RPM	1953	2441	2686	2808	2930
C_X_S	36.3%	19.69%	12.56%	9.24%	5.1%

**Table 17.**  
*Response surface selectivity improvement with target data length.*





**Figure 68.**  
*C<sub>X\_S</sub> versus data length.*



**Figure 69.**  
*Local minimum proliferation.*

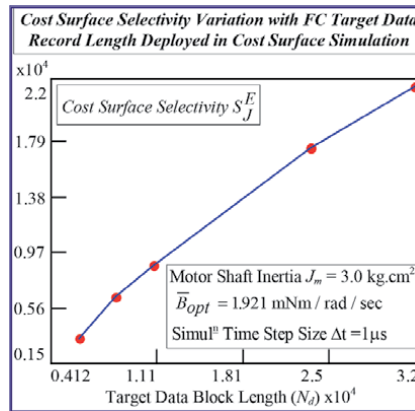
manifestation is due to the contributory interference effect of the two frequency modulated sinusoids, during the machine speed step response, in the penalty function construct as explained in [13]. The selectivity is obtained by fitting a quadratic polynomial, expressed in terms of the parameter  $J$ , to the main lobe of the cost surface cross section centred on the vertex  $\bar{X}_{opt}$  as shown in **Figure 69** with

$$E(J) - \bar{E}_{opt} = b_0 (J - \bar{J}_{opt})^2 \quad (45)$$

and least squares coefficient  $b_0$  determined [9, 10]. The selectivity measure is based on the first order variation  $\delta E(J)$  of the penalty cost function about the global minimum estimate  $\bar{X}_{opt}$  in terms of the worst case parameter step size  $\delta J^L$  as

$$S_J^E \Big|_{\bar{X}_{opt}} = \frac{\delta E}{\delta J} = b_0 \delta J^L. \quad (46)$$

This measure indicates a quasi linear dependency with increased data record length  $N_d$  as detailed in **Table 9** and shown in **Figure 70**. Further increase in data



**Figure 70.**  
FC cost surface selectivity.

record length causes the capture cross section to tend asymptotically to zero, with the nearest local minimum to the global entity approaching the same minimum energy as the ground state  $E_{opt}$ , due to the gradual onset of steady state conditions. The increased data length results in a proliferation and clustering of genuine local minima, other than noise induced cost surface granularity, about the global extremum as the initial transient step response is swamped by the onset of steady state motor speed.

It is noteworthy that a judicious application of the Fast Fourier Transform (FFT) to the FC target data, after the initial transient has partially expired, provides an estimate of the motor speed variation with residual data record length as shown in **Figure 27**. The BLMD shaft speed can be approximated by the FFT of the FC step response over a short time span at the end of each FC data sequence where machine speed is quasi constant and shaft velocity fluctuations have almost disappeared. The exponential buildup of the motor shaft speed with data record length/time in **Figure 27** is indicative of the characteristic response to the torque demand step input [9, 10].

### 13. Conclusions

The FSD optimization technique has been shown, although computationally intensive, to be a very accurate and effective parameter identification method over a noisy cost surface with embedded local minima. This novel method avoids the convergence difficulties associated with the application of classical optimization techniques during parameter extraction by providing an anneal temperature related Boltzmann probability of escape from local minimum capture. The accuracy of the FSD returned parameter estimates is independently confirmed with values found to be in substantial agreement with those extracted through cost surface simulation, with worst case relative error in the damping parameter  $B$  of 2%, for known BLMD dynamics. Further confidence enhancement in the FSD identification strategy is provided by the degree of correlation ( $\sim 94\%$ ) of BLMD model simulations, based on returned FSD parameter estimates, with observed test data. The application of the Powell Conjugate Direction method, as an alternative competitive identification strategy to the FSD method, with a simple unimodal velocity cost function has been discussed. The identified BLMD dynamics, though reasonable by comparison with known nominal values in that relative errors are less than 6% in **Table 4**, are not as

accurate as those returned by the FSD method. This discrepancy is apparent for BLMD waveform simulation over long spans, using returned PCD parameter estimates for three known cases of inertial load to be identified, with poor correlation estimates spanning 20–70% as per **Table 6** for FM related current traces. The reasons, relating to cost surface selectivity and returned parameter accuracy, underpinning the choice of target data in the MSE objective function have been discussed. The obscuration of shaft velocity transient response detail by the onset of steady state conditions, as a consequence of lengthy data records, results in impairment of the PCD method in terms of reduced cost surface selectivity and loss of returned parameter accuracy with increased convergence metric in **Table 7**. The reverse trend emerges during FSD optimization, with better selectivity and improved parameter accuracy with smaller capture cross section, for current feedback usage with this being the preferred target data choice in MSE cost formulation.

The effectiveness of the modified version of the FSD algorithm has been demonstrated in the extraction of the BLMD dynamics with enhanced speed in global convergence which is less than 15% of the runtime for the original FSD algorithm based on zero inertial shaft load conditions. The accuracy of the method is verified with initialization far from the global minimum at the edge of the parameter tolerance band and supported by the returned results with relative errors less than 0.3% as per **Table 5** for three known cases of motor shaft inertia. The correlation accuracy of BLMD waveform simulation with measured data is greater than 94% in **Table 14** for all cases of returned inertial parameter estimates by the FSD method. The computation intensity of the method can be reduced by adopting coarser quantization step sizes, instead of the specified noise threshold value, within an acceptable bounded parameter relative error. This improved FSD method can be potentially used, as an effective optimizer in the dynamical parameter extraction phase, to facilitate autotuning during embedded system commissioning.

## Acknowledgements

The author wishes to acknowledge

1. Eolas (Science Fundation Ireland) – The Irish Science and Technology Agency – for research funding.
2. Moog Ireland Ltd. for brushless motor drive equipment for research purposes.

## Author details

Richard A. Guinee  
Munster Technological University (MTU), Cork, Ireland

\*Address all correspondence to: [richard.guinee@cit.ie](mailto:richard.guinee@cit.ie)

## IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] EL\_Sharkawi M A. Advanced control techniques for High Performance Electric Drives. In: Control and Dynamic Systems. Academic Press; 1991. Vol 44; p. 59-130.
- [2] Friedrich G, Kant M. Choice of drives for electric vehicles: a comparison between two permanent magnet AC machines. IEE Proc.-Elec. Power Appl.. May 1998; Vol. 145; No. 3.
- [3] Demerdash N A, Nehl T W, Maslowski E. Dynamic modelling of brushless dc motors in electric propulsion and electromechanical actuation by digital techniques. In: IEEE/IAS Conf. Rec. CH1575-0/80/0000-0570; 1980. pp. 570-579.
- [4] Dote Y. Servo Motor and Motion Control using Digital Signal Processors. Prentice Hall International;1990.
- [5] Kuo B C, Tal J. Incremental Motion Control Systems: DC Motors and Control Systems. Illinois: SRL Publ. Co.; 1978.
- [6] DC Motors, Speed Control, Servo Systems. 5th ed. Electrocraft Corp.; 1980.
- [7] Vas P. Sensorless Vector and Direct Torque Control. Oxford Science Publications: Oxford Univ. Press; 1998.
- [8] Dawson D M, Hu J, Burg T C. Nonlinear Control of Electric Machinery. Marcel Dekker Inc.; 1998.
- [9] Guinee R A. Mathematical Modelling and Simulation of a PWM Inverter Controlled Brushless Motor Drive System from Physical Principles for Electric Vehicle Propulsion Applications [Internet]. In: Electric Vehicles – Modelling & Simulation. Intech Open Open Access Publisher; Aug 2011. p. 233-286. Ch11. Available from: <https://www.intechopen.com/books/electric-vehicles-modelling-and-simulations/mathematical-modelling-and-simulation-of-a-pwm-inverter-controlled-brushless-motor-drive-system-from>
- [10] Guinee R A. Extended Simulation of an Embedded Brushless Motor Drive (BLMD) System for Adjustable Speed Control Inclusive of a Novel Impedance Angle Compensation Technique for Improved Torque Control in Electric Vehicle Propulsion Systems [Internet]. In: Electric Vehicles–Modelling & Simulation. IntechOpen Open Access Publisher; Aug 2011. p. 417-466. Ch19. DOI: 10.5772/29821. Available from: <https://www.intechopen.com/books/electric-vehicles-modelling-and-simulations/extended-simulation-of-an-embedded-brushless-motor-drive-blmd-system-for-adjustable-speed-control-in>
- [11] Guinee R A, Lyden C. Accurate Modelling and Simulation of a High Performance Permanent Magnet Adjustable Speed Drive System for Embedded Systems. In: 8th European Conference on Power Electronics and Applications (EPE '99); Sep. 1999; Lausanne. Switzerland.
- [12] Guinee R A, Lyden C. Accurate Modelling And Simulation Of A DC Brushless Motor Drive System For High Performance Industrial Applications. In: IEEE International Symposium on Circuits and Systems (ISCAS '99); May/ June 1999; Orlando. Florida.
- [13] Guinee R A. Mathematical Analysis for Response Surface Parameter Identification of Motor Dynamics in Electric Vehicle Propulsion Control [Internet]. In: New Generation of Electric Vehicles. IntechOpen Open Access Publisher; Dec 2012; p. Ch11. DOI: 10.5772/54483. Available from: <https://www.intechopen.com/books/new-generation-of-electric-vehicles/mathematical-analysis-for-response->

surface-parameter-identification-of-motor-dynamics-in-electric-ve

- [14] Guinee R A, Lyden C. Motor Parameter Identification using Response Surface Simulation and Analysis. In: Proc. of American Control Conference (ACC 2001); June 25-27; 2001; Arlington. VA. USA.
- [15] Kirkpatrick S, Gelatt Jr. C D, Vecchi M P. Optimization by Simulated Annealing. Science; May 1983; Vol.220; No.4598.
- [16] Johnson D S, Aragon C R, McGeogh L A, Schevon C. Optimization by simulated annealing: an experimental evaluation; Part 1, graph partitioning. Operations Res.; Nov-Dec 1989; Vol. 37; p. 865-892.
- [17] Ellams P, Mansell A D. Simulated annealing in the analysis of resonant DC link inverters. In: IEE Proc.-Electr. Power Appl.; May 1994; Vol. 141; No. 3,
- [18] Sorkin G B. Simulated Annealing on Fractals: Theoretical Analysis and Relevance for Combinatorial Optimization. In: Proc. 6<sup>th</sup> ann. MIT Conf. Advanced Res. VLSI.; Apr 1990; p. 331-351.
- [19] Hu Y H, Pan S W. SaPOSM: An Optimization Method Applied to Parameter Extraction of MOSFET Models. IEEE Trans on Computer-Aided Des. of Integ. Cir. And Sys.. Oct. 1993; Vol. 12; No. 10.
- [20] Vai M -K, Ng M F D. A technology-independent device modelling program using simulated annealing optimization. In: Proc. C.I.C.C.'89; May 1989; p 381-384.
- [21] Chandy J A, Kim S, Ramkumar B, Parkes S, Banerjee P, An Evaluation of Parallel Simulated Annealing Strategies with Application to Standard Cell Placement. IEEE Trans on Computer-Aided Des. of Integ. Cir. And Sys.. Apr. 1997; Vol. 16; No. 4.
- [22] Guinee R A, Lyden C. A Novel Application of the Fast Simulated Diffusion Algorithm in Brushless Motor Parameter Identification. In: The 3<sup>rd</sup> IEEE European Workshop on Computer-Intensive Methods in Control and Data Processing; Sep. 1998; UTIA. Prague Czech Republic.
- [23] Guinee R A, Lyden C. A Novel Application of the Fast Simulated Diffusion Algorithm for Dynamical Parameter Identification of Brushless Motor Drive Systems. In: The 2000 IEEE International Symposium on Circuits and Systems (ISCAS 2000); May 2000; Geneva. Switzerland.
- [24] Sakurai T, Lin B, and Newton A R. Fast Simulated Diffusion: An Optimization Algorithm for Multiminimum Problems and Its Application to MOSFET Model Parameter Extraction. IEEE Trans. on Computer-Aided Design. Feb. 1992; Vol. 11.
- [25] Guinee R A, Lyden C. Parameter Identification of a Motor Drive using a Modified Fast Simulated Diffusion Algorithm. In: Proc. of the IASTED Intern. Conf. on Modelling and Simulation; May 1998; Pittsburgh. Pa. USA. IEEE; p. 224-228.
- [26] Guinee R A, Lyden C. A Novel Application of the Fast Simulated Diffusion Optimization Technique for Brushless Motor Parameter Extraction. In: International Conference on Control (UKACC); Sep 2000; Cambridge Univ..
- [27] Guinee R A, Lyden C. Parameter Identification of a Brushless Motor Drive System using a Modified Version of the Fast Simulated Diffusion Algorithm. In: Proc. of American Control Conference (IEEE ACC-1999); June 1999; San Diego. USA.
- [28] Guinee R A. Mathematical Modelling, Simulation and Parameter Identification of a Permanent Magnet

- Brushless Motor Drive System. Ph.D Thesis, National Microelectronics Research Centre (NMRC) – Tyndall Institute: University College Cork - National University of Ireland; 2003.
- [29] Leu M C, Liu S, Zhang H. Modelling, Analysis and Simulation of Brushless DC Drive System. In: Winter Annual Meeting of ASME; 89-WA/DSC-1; 1989; San Fran. USA.
- [30] Ljung L. Issues in System Identification. In: IEEE Control Systems IEEE; Jan 1991. P. 25-29. 0272-1708/91/0100-0025,
- [31] Ljung L. A Discussion Of Model Accuracy In System Identification. International Journal Of Adaptive Control and Signal Processing. J. Wiley & Sons; 1992. Vol.6.
- [32] Ljung L. System identification: Theory For The User. PHI; 1987.
- [33] Moog Brushless Technology: Brushless Servodrives User Manual. Moog GmbH; 1988. D-7030 Böblingen. Germany. D310.01.03 En/De/It 01.88.
- [34] Holtz J. Sensorless Position Control of Induction Motors - an Emerging Technology - invited paper. In: Proc. of the 24<sup>th</sup> Annual Conf. of the IEEE Indus. Electronics Society (IECON '98); Sep. 1998; Aachen, Germany.
- [35] Fletcher R. Practical Methods of Optimization. 2nd ed. J.Wiley & Sons; 1993.
- [36] Fletcher R. Function minimization without evaluating derivatives - a review. Computer J. 1965:8:33-41.
- [37] Powell M J D. A method for minimizing a sum of squares of non-linear functions without calculating derivatives. Computer J. 1965:11: 302-304.
- [38] Swann W H. Direct Search Methods. Numerical Methods for Unconstrained Optimization. London: Academic Press; 1972.
- [39] Goldfeld S M, Quandt R E, Trotter H F. Maximization by Quadratic Hill-Climbing. Econometrica. July 1966: Vol. 34; No.3.
- [40] Box M J. A comparison of several current optimization methods, and the use of transformations in constrained problems. Computer J. 1966;9:67-77.
- [41] Spendley W, Hext G R, Himsworth F R. Sequential Applications of Simplex Designs and Evolutionary Operation. Technometrics. 1962; Vol. 4: No. 4.
- [42] Walsh G R. Methods of Optimization. John Wiley & Sons; 1979.
- [43] Bunday B D. Basic Optimization Methods. London: E. Arnold Publishers; 1984.
- [44] Fiacco A V, McCormick G P. Computational Algorithm for the sequential unconstrained minimization technique for nonlinear programming. Management Science. July 1964; Vol. 10: No. 4.
- [45] Spang H A. A review of minimization techniques for nonlinear functions. SIAM Review. Oct. 1962: Vol. 4, No. 4.
- [46] Gill P E, Murray W. Algorithms for the solution of the nonlinear least squares problem. SIAM J. Numer. Anal.. Oct 1978; Vol. 15: No. 5
- [47] Adby P R, Dempster M A H. Introduction to Optimization Methods. London: Chapman and Hall; 1974
- [48] Shanno D F. Conjugate gradient methods with inexact searches. Mathematics of Operations Research. Aug 1978; Vol. 3: No. 3.
- [49] Press W H, Flannery B F, Teukolsky S A, Vetterling W T. NUMERICAL RECIPES in C. CUP; 1990.

- [50] Levenberg K. A method for the solution of certain nonlinear problems in least squares. *Quart. Appl. Math.*.1944; No. 2: p. 164-168.
- [51] Doganis K, Scharfetter D L. General Optimization and Extraction of IC Device Model Parameters. *IEEE Trans. on Elect. Dev.*. 1983; Vol. 30: No. 9.
- [52] Stewart G W. A modification of Davidson's minimization method to accept difference approximations of derivatives. *J. Assoc. Comput. Mach.*. 1967; 14: p.72-83.
- [53] Aluffi-Pentini F, Parisi V, Zirilli F. Global optimization and stochastic differential equations. *J. Optimization Theory and Applications*. Sep. 1985; vol. 47: no. 1
- [54] Fletcher R. *Optimization*. London: Academic Press; 1969.
- [55] Hestenes M R. *Conjugate Direction Methods in Optimization*. New York:Springer-Verlag; 1980.
- [56] Brent R P. *Algorithms for Minimization Without Derivatives*. PHI; 1973.
- [57] S. Geman and C-R. Hwang, "Diffusions for global optimization", *SIAM J. Contr. Optimization*, Vol. 24, No. 5, Sep. 1986, pp. 1031-1043.
- [58] Eisberg R, Resnick R. *Quantum Physics*. NY: J. Wiley & Sons; 1974.
- [59] Metropolis M, Rosenbluth A, Rosenbluth M, Teller A, Teller E. Equation of State Calculations by Fast Computing Machines. *J. Chem. Phys.*. June 1953; vol. 21.
- [60] Huang M D, Romeo F, Sangiovanni-Vincentelli A. An efficient general cooling schedule for simulated annealing. In: *Proc. ICCAD'86*; Nov. 1986.
- [61] Mosteller F, Rourke R E K, Thomas G B. *Probability with Statistical Applications*. Addison-Wesley Publ. Co.; 1961
- [62] Glasstone S, Edlund M C. *The Elements of Nuclear Reactor Theory*. Macmillan & Co. Ltd; 1956.





---

Section 3

# Enabling Technologies

---



# Visibility-Based Technologies and Methodologies for Autonomous Driving

*Said Easa, Yang Ma, Ashraf Elshorbagy, Ahmed Shaker, Songnian Li and Shriniwas Arkatkar*

## Abstract

The three main elements of autonomous vehicles (AV) are orientation, visibility, and decision. This chapter presents an overview of the implementation of visibility-based technologies and methodologies. The chapter first presents two fundamental aspects that are necessary for understanding the main contents. The first aspect is highway geometric design as it relates to sight distance and highway alignment. The second aspect is mathematical basics, including coordinate transformation and visual space segmentation. Details on the Light Detection and Ranging (Lidar) system, which represents the ‘eye’ of the AV are presented. In particular, a new Lidar 3D mapping system, that can be operated on different platforms and modes for a new mapping scheme is described. The visibility methodologies include two types. Infrastructure visibility mainly addresses high-precision maps and sight obstacle detection. Traffic visibility (vehicles, pedestrians, and cyclists) addresses identification of critical positions and visibility estimation. Then, an overview of the decision element (path planning and intelligent car-following) for the movement of AV is presented. The chapter provides important information for researchers and therefore should help to advance road safety for autonomous vehicles.

**Keywords:** Lidar, traffic visibility, infrastructure visibility, high-precision maps, sight distance, highway design, technologies

## 1. Introduction

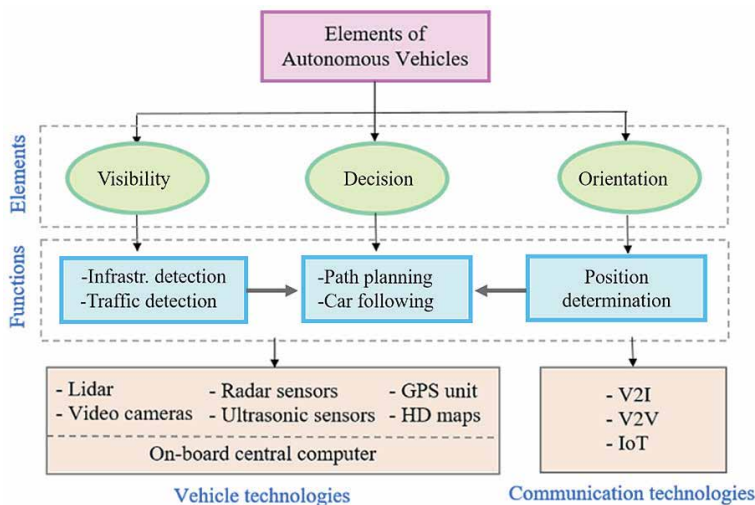
Autonomous vehicle (AV) will reduce human errors and is expected to lead to significant benefits in safety, mobility, and sustainability [1–5]. The technical feasibility of automated highways has been demonstrated in San Diego, California in 1997 [6, 7]. The technology is emerging around the world on both passenger and freight sides. Autonomous vehicles have already started to appear on the roads across the globe. Clearly, as the AV market expands, transportation professionals and researchers must address an array of challenges before AV soon becomes a reality. Several government and industry entities have deployed demonstrations and field tests of the technology. Centres for testing and education, products, and standards for AV have also been established. Currently, researchers, scientists, and engineers are investing significant resources to develop supporting technologies.

The self-driving system involves three main elements: orientation, visibility, and decision, as shown in **Figure 1**. For each element, certain functions are performed with the aid of one or a combination of technologies. For AV orientation, the position of the vehicle is determined mainly using Global Navigation Satellite System (GNSS). To increase reliability and accuracy, this technology is supplemented with other data gathered from specific perception technologies, such as cameras and internal measurement devices (tachometers, altimeters, and gyroscopes).

The visibility of AV, sometimes referred to as perception, involves infrastructure detection (e.g. sight obstacles, road markings, and traffic control devices) and traffic detection (other vehicles, pedestrians, and cyclists). High-precision (HP) maps help the AV not only perceive the environment in its vicinity, but also plan for the turns and intersection far beyond the sensors' horizons. The other main technology that represents the 'eye' of the AV is the Light Detection and Ranging (Lidar). Other technologies used for infrastructure visibility include video cameras which can detect traffic signals, read road signs, keep track of other vehicles, and record the presence of pedestrians and other obstacles. Radar sensors can determine the positions of other nearby vehicles, while ultrasonic sensors (normally mounted on wheels) can measure the position of close objects (e.g. curbs). Lidar sensors can also be used to identify road features, like lane markings. Three primary sensors (camera, radar, and Lidar) work together to provide AV with visuals of its 3D environment and help detect the speed and distance of nearby objects. The visibility information is essential for the safe operation of the AV. For example, the information can be used to ensure that adequate sight distance (SD) is available and if not to take appropriate actions.

The decision of AV is based on the data generated by the AV sensors and HP maps. Examples of the decisions that are made by AV include routing, changing lanes, overtaking vehicles, stopping at traffic lights, and turning at an intersection. The decision is made by an on-board centralized artificial intelligence (AI) computer that is linked with cloud data and include the needed algorithms and software.

For the AV to function effectively, it must communicate with the environment as part of a continuously updating, dynamic urban map [3]. This involves vehicle-to-vehicle (V2V) and vehicle-to-infrastructure (V2I) communications. Internet of Things (IoT) technology provides telematics data about the condition and



**Figure 1.** Elements of autonomous vehicles and their functions and technologies.

performance of AV, such as real-time location and speed, idling duration, fuel consumption, and the condition of its drivetrain.

This chapter mainly focuses on the visibility of AV and briefly addresses the decision element. Section 2 presents important visibility fundamentals related to highway geometric design and mathematical tools. Section 3 presents details on the Lidar system, which represents the ‘eye’ of the AV. This section also introduces a new Lidar system that has great potential for high precision mapping for autonomous vehicles. Section 4 describes the methodologies related to infrastructure visibility and traffic visibility. Section 5 presents some implementation aspects and Section 6 presents the conclusions.

## 2. Visibility fundamentals

### 2.1 Highway sight distance

In the geometric design guides by American Association of State Highways and Transportation Officials and Transportation Association of Canada [8–10], the available sight distance for human-driven vehicles is measured from the driver’s height above the pavement surface. For autonomous vehicles, the driver’s eye is the Lidar. The ability of the AV to ‘see’ ahead is critical for the safe and efficient operation. Sufficient sight distance must be provided to allow the AV to stop, avoid obstacles on the roadway surface, overtake slow vehicles on two-lane highways, and make safe turns/crossings at intersections. For human-driven vehicles, the required sight distance is based on driver’s perception-reaction time (PRT), vehicle speed, and other factors. A value of 2.5 s is used for human-driven vehicles and a value of 0.5 s has been assumed for AV in the literature [11]. This smaller reaction time will result in shorter required sight distances for AV.

For autonomous vehicles, four basic types of SD are defined as follows:

- *Stopping Sight Distance (SSD)*: This is the distance traveled by the AV during system reaction time and the braking distance from the operating speed to stop.
- *Passing Sight Distance (PSD)*: This is the distance, including system reaction time, required by the AV on rural two-lane roads to allow the vehicle to pass a slower vehicle by using the opposing lane.
- *Decision Sight Distance (DSD)*: This is the distance that allows the AV to maneuver or change its operating speed or stop to avoid an obstacle on the roadway surface.
- *Intersection Sight Distance (ISD)*: This is the distance along a cross road with a right of way that must be clearly visible to the crossing AV so that it can decide and complete the maneuver without conflicting with the cross-road vehicles.

Since the AV response time is less than the driver’s PRT, the required SD for autonomous vehicles would be shorter than that for human-driven vehicles. The impact of AV on highway geometric design has been explored in a preliminary manner by making simple assumptions regarding system reaction time and Lidar field of view [11]. The study focused on SSD, DSD, and vertical curves.

The required SD for autonomous vehicles, Lidar height, and object height influence the design and evaluation of highway vertical, horizontal, and 3D alignments, as shown in **Figure 2**. The Lidar height above the pavement surface is critical in

determining the available sight distance. For crest vertical curves (**Figure 2a**), since the required SD for autonomous vehicles are shorter, the required Lidar height for safe operation,  $h_L$ , would be somewhat less than the design driver's eye height,  $h_1$ . Thus, by placing the Lidar at or above  $h_L$ , the AV can safely operate on existing highways without the need for modifying their design. For a sag vertical curve with an overpass (**Figure 2b**), where the truck driver's eye height controls the traditional curve design, the required Lidar height for safe operation would be somewhat larger than the design driver's eye height. Thus, by placing the Lidar at or below  $h_L$ , the AV can safely operate on existing highways. For horizontal curves (**Figure 2c**), the Lidar height is not important for detecting horizontal obstacles, except when cut slopes are present.

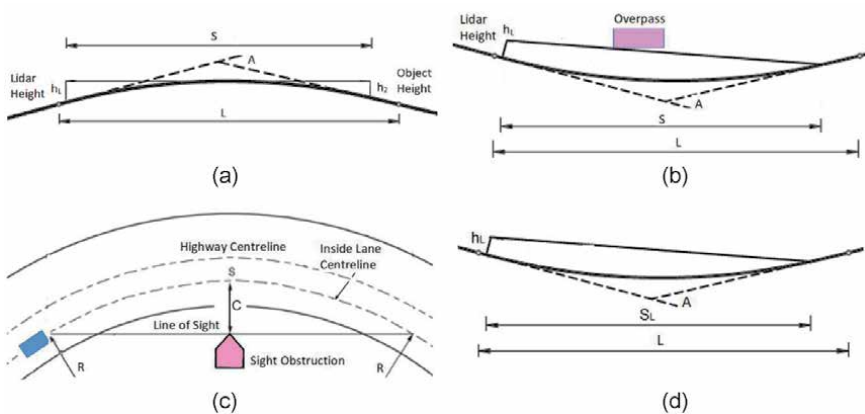
For human-driven vehicles, sag vertical curves (**Figure 2d**) are designed based on the distance that the vehicle headlamp can illuminate ahead of the vehicle, where the angle of light beam upward from the vehicle plane is considered equal to  $1^\circ$ . What is interesting is that for AV the Lidar height is irrelevant to the operation on sag vertical curves. The reason is that Lidar detection does not require light and therefore an obstacle ahead of the autonomous vehicle can be detected under all light conditions. It is necessary, however, to ensure that the effective Lidar range is greater than the required SD.

Note that the geometry of SD shown in **Figure 2** represents individual horizontal or vertical alignments. For 3D alignment (combined horizontal and vertical curves), the AV can directly determine the available SD and compare it with the required SD. The sight line in this case may be obstructed by the pavement surface or by obstacles on the roadside, such as a building or a cut slope.

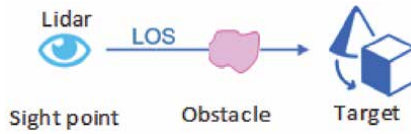
## 2.2 Mathematical tools

### 2.2.1 Overview of visibility modeling

Quantitative visibility estimation (VE) involves four basic components as shown in **Figure 3**: (1) sight point, (2) target points (e.g. lane marking, traffic sign, and stalled vehicle), (3) line of sight (LOS) which connects the sight point with target points, and (4) obstacle data that obstruct LOS (e.g. vegetation, barrier, and building). The purpose of VE is to determine whether the obstacles affect the visibility of the targets to the sight point in a mathematical manner.



**Figure 2.** Effect of Lidar height and sight distance on the operation of AV on different highway alignments. Note:  $A$  = algebraic difference in grade,  $C$  = lateral clearance,  $h_L$  = Lidar height,  $h_2$  = object height,  $S_L$  = effective Lidar range,  $L$  = vertical curve length,  $R$  = radius of horizontal curve,  $S$  = sight distance. (a) Crest vertical curve (b) sag vertical curve with overpass, (c) horizontal curve and (d) sag vertical curve.



**Figure 3.**  
 Basic components of quantitative visibility estimation.

The accuracy of VE is closely associated with the precision of obstacle data. In the past, due to a lack of effective and affordable technique to collect dense and precise geospatial data, the obstacle data were conventionally represented by digital surface and terrain models generated from sparse geospatial points. However, it was proved that neither model could handle objects with complex shapes and may yield inaccurate VE results in some cases [12–14].

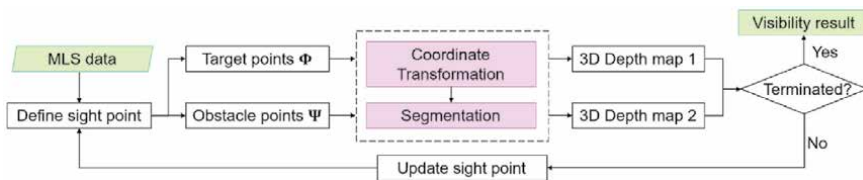
Over the past five years, mobile laser scanning (MLS) data have been recognized as a reliable data source for conducting visibility-related analyses [15–18]. MLS point clouds enable a very accurate and precise representation of real-world environment. In this case, a number of indoor computerized estimations on MLS point clouds are viable to replace risky and time-consuming outdoor field measurements [16]. In addition, MLS data are also the main data source for manufacturing HP maps which are indispensable to AV. Therefore, the mathematical model for estimating visibility in autonomous driving is built based on MLS data.

The general workflow of VE using MLS data is graphically presented in **Figure 4**. For a given eye position, both the target points and obstacle points undergo the same procedure (i.e. coordinate transformation and segmentation) and two respective depth maps of the same size are generated. The visibility estimation is achieved through comparing two 3D depth maps. The process of VE is described in detail next.

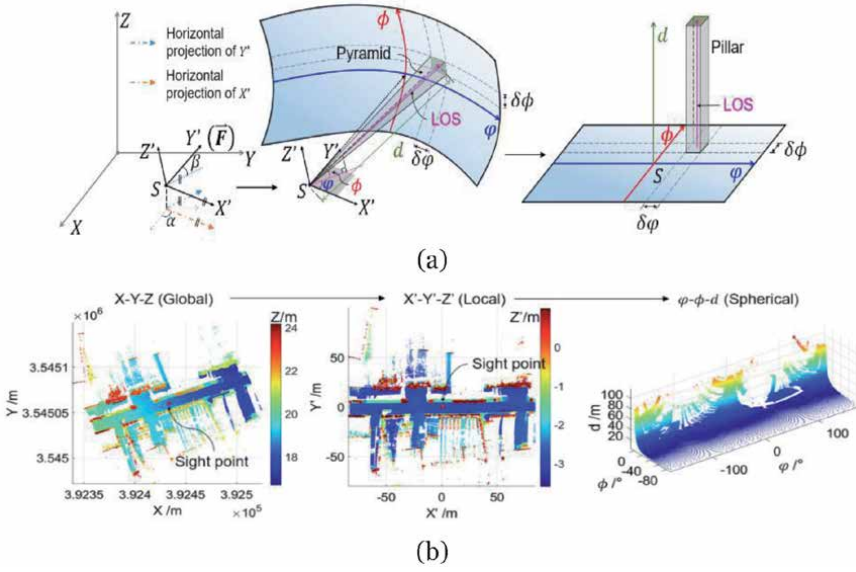
### 2.2.2 Coordinate transformation

To illustrate, consider VE at a single position. Let  $S$  be the sight point. Let  $\vec{F}$  be the forward direction. Let  $\Phi$  and  $\Psi$  be the set of target points and obstacle points, respectively. To reduce the computational complexity, a local coordinate frame is established first as depicted in **Figure 5a**. Set  $S$  as the origin. Then, the  $Y'$  and  $X'$  axes are set along  $\vec{F}$  and the horizontal vector normal to  $\vec{F}$ , respectively. Finally, the  $Z'$ -axis is set along the upward direction perpendicular to the  $X'SY'$  plane. The local coordinates of points ( $S, \Phi$  and  $\Psi$ ) are obtained via:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \mathfrak{R} \cdot \begin{bmatrix} x - x_s \\ y - y_s \\ z - z_s \end{bmatrix} = \begin{bmatrix} \cos \alpha \cdot \cos \beta & -\cos \beta \cdot \sin \alpha & \sin \beta \\ \sin \alpha & \cos \alpha & 0 \\ -\cos \alpha \cdot \sin \beta & \sin \alpha \cdot \sin \beta & \cos \beta \end{bmatrix} \cdot \begin{bmatrix} x - x_s \\ y - y_s \\ z - z_s \end{bmatrix} \quad (1)$$



**Figure 4.**  
 Workflow of VE using MLS data.



**Figure 5.** Coordinate transformation: (a) different coordinate systems and (b) visualization of points in different frames (generated with Matlab 2020b; the same for **Figures 6, 7, and 8-10**).

where  $\mathfrak{R}$  = the rotation matrix,  $(x_s, y_s, z_s)^T$  = geodetic coordinates of the sight point,  $(x, y, z)^T$  and  $(x', y', z')^T$  = geodetic and local coordinates of points ( $S$ ,  $\Phi$  and  $\Psi$ ), respectively, and  $\alpha, \beta$  = rotation angles around the  $Z$  and  $X$ -axes, respectively.

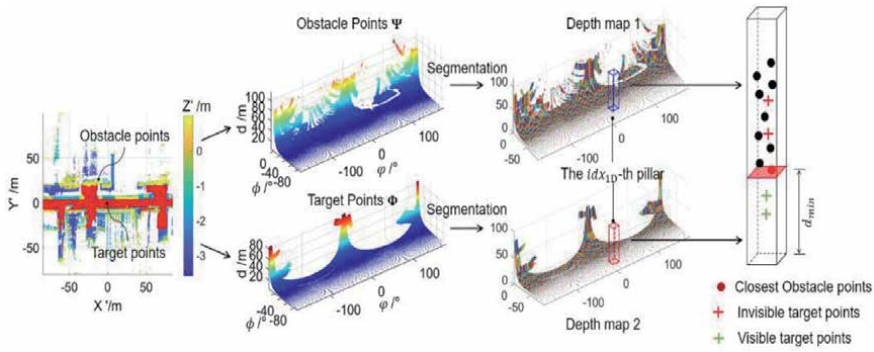
To enable a more intuitive and efficient estimation of target visibility, the local Cartesian coordinates are further converted into spherical coordinates as follows,

$$\begin{cases} \varphi = \arctan \frac{y'}{x'} \\ \phi = \arctan \frac{z'}{\sqrt{x'^2 + y'^2}} \\ d = \sqrt{x'^2 + y'^2 + z'^2} \end{cases} \quad (2)$$

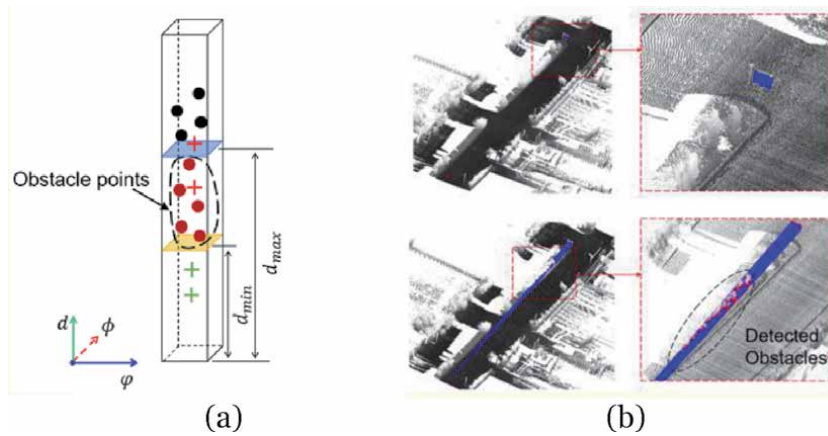
where  $(\varphi, \phi, d)^T$  = azimuth, elevation, and radius of the spherical coordinates shown in **Figure 5a**.

Points in different coordinate frames are visualized in **Figure 5b**. The sight point marked in figure corresponds to  $S$  in **Figure 5a**. In the  $\varphi$ - $\phi$ - $d$  space,  $\varphi$  and  $\phi$  refer to the horizontal and vertical angles between a LOS and the forward direction, respectively, while  $d$  measures the depth from the sight to the target point. Conventionally, as shown in **Figure 5a**, a sight cone or sight pyramid is constructed around a given LOS in a local frame to detect whether the obstacle points therein touch that LOS [12, 17]. However, when the number of target points is very large, the practice of building sight pyramids and then searching for points therein in dense MLS point clouds is quite time-consuming. With the coordinate transformation, a sight pyramid in the  $x'$ - $y'$ - $z'$  space is equivalent to a pillar in the  $\varphi$ - $\phi$ - $d$  space. In this case, it is reasonable to assume that when  $\delta\varphi$  and  $\delta\phi$  are small, only the point with the minimum  $d$  can be seen by an observer in a pillar. This process is more straightforward than detecting the closest point in a sight pyramid in the  $x'$ - $y'$ - $z'$  coordinate system.





**Figure 6.**  
 Comparison of two depth maps.



**Figure 7.**  
 Sight obstacle detection: (a) principle and (b) results.

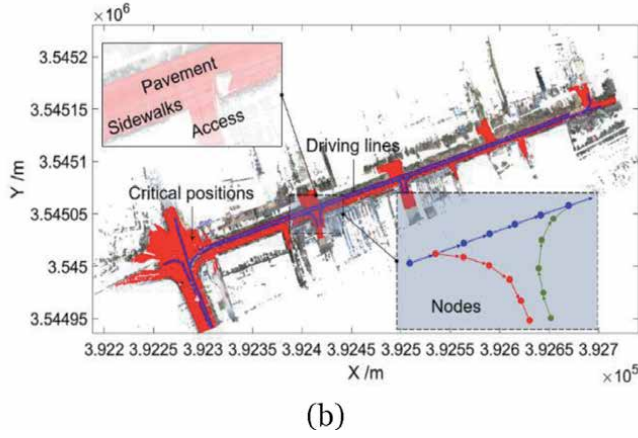
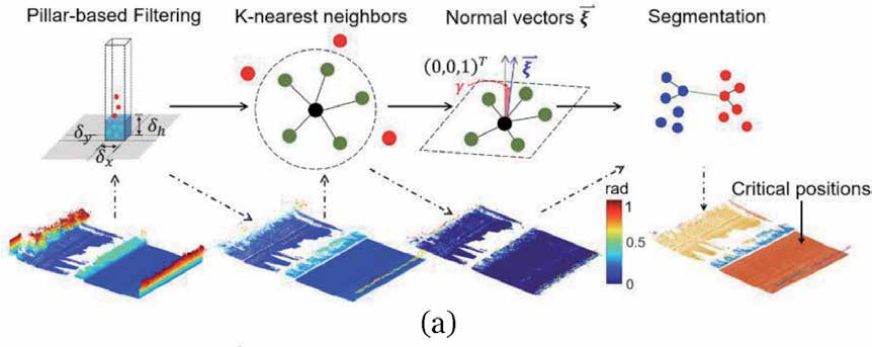
### 2.2.3 Visual space segmentation

Based on the preceding assumption, the main task is to segment points in the  $\varphi$ - $\phi$ - $d$  space into different pillars. To reduce the computational complexity and thus improve the efficiency, a linear index based segmentation procedure was developed and applied to partition points. The procedure assigns a new one-dimensional (1D) index to each point and can divide point cloud data into pillars or voxels. Taking the voxel generation for illustration, the segmentation procedure comprises three main steps as follows:

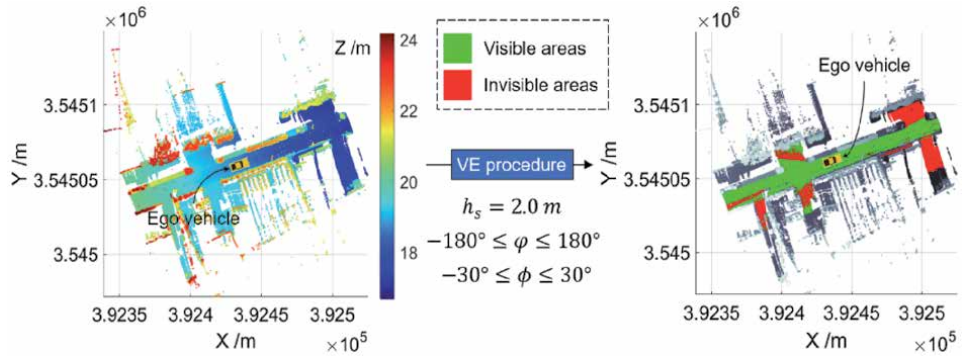
*Step 1:* Grid the original 3D data  $(a, b, c)^T$  using the following formula:

$$\begin{cases} a_g = \lfloor \frac{a}{\delta_a} \rfloor \cdot \delta_a \\ b_g = \lfloor \frac{b}{\delta_b} \rfloor \cdot \delta_b \\ c_g = \lfloor \frac{c}{\delta_c} \rfloor \cdot \delta_c \end{cases} \quad (3)$$

where  $\lfloor \cdot \rfloor$  = a function that rounds the number to an integer,  $\delta_a, \delta_b, \delta_c$  = user-defined voxel dimensions along the  $a, b$ , and  $c$  axes, respectively. **Figure 11a** shows the dense points that are converted to the gridded data.

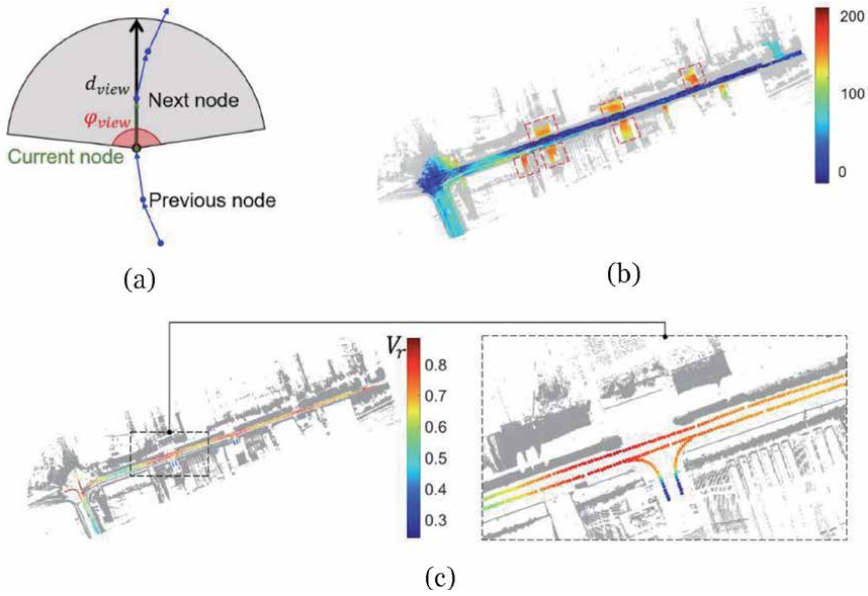


**Figure 8.** Extraction of critical positions: (a) general process of identifying planar ground points and (b) critical positions and driving lines.

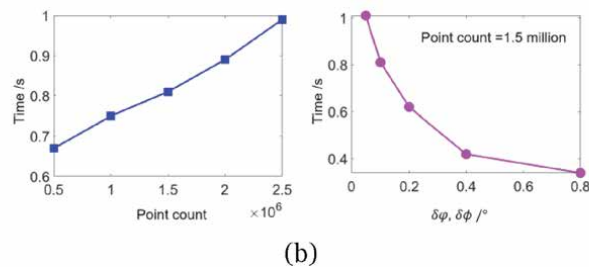
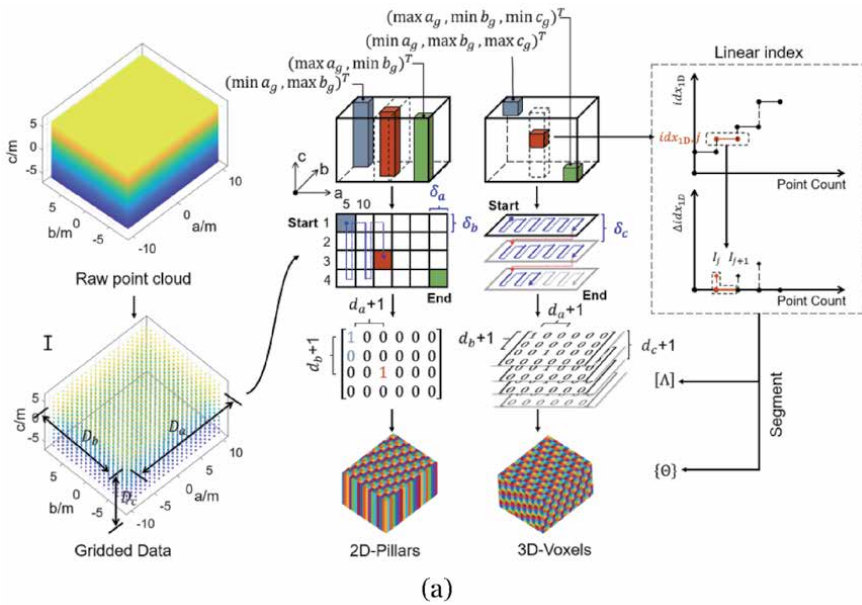


**Figure 9.** VE results at a single node.

Step 2: Let  $D_a, D_b,$  and  $D_c$  be the distances along the  $a, b,$  and  $c$  axes from  $(\min a_g, \min b_g, \max c_g)^T$  to  $(\max a_g, \max b_g, \min c_g)^T$ , respectively. Let  $d_a, d_b,$  and  $d_c$  be the distances along the  $a, b,$  and  $c$  axes from  $(\min a_g, \min b_g, \max c_g)^T$  to a certain grid point  $P$ , respectively. A null  $(D_b + 1) \times (D_a + 1) \times (D_c + 1)$  cell matrix  $\{\Theta\}$  and the same-size numerical matrix  $[\Lambda]$  are then constructed. The location of the grid point  $P$  in  $\{\Theta\}$  or  $[\Lambda]$  is  $(d_b + 1, d_a + 1, d_c + 1)^T$ . The location can be converted to a linear index  $idx_{1D}$  as follows



**Figure 10.** Visualization of VE results: (a) VE process, (b) cumulative times of invisibility, and (c) visibility ratio.



**Figure 11.** Linear index-based segmentation: (a) process and (b) time performance.

$$idx_{1D} = d_c \cdot (D_b + 1) \cdot (D_a + 1) + d_a \cdot (D_b + 1) + d_b + 1 \quad (4)$$

where  $d_a \leq D_a, d_b \leq D_b, d_c \leq D_c$  and  $idx_{1D} \leq (D_b + 1) \cdot (D_a + 1) \cdot (D_c + 1)$ .

In a matrix, each element can be accessed through either its 3D location or the linear index  $idx_{1D}$  [19–21]. Using Eq. (4), a 1D index is obtained for each point.

*Step 3:* Sort the points by increasing  $idx_{1D}$ , where the points falling inside a voxel will share the same  $idx_{1D}$ , as shown in **Figure 11a**. By computing the difference  $\Delta idx_{1D}$  between two consecutive  $idx_{1D}$ , the points in a certain voxel is determined by detecting the positions where  $\Delta idx_{1D}$  jumps. Suppose

$I\{I_1, I_2 \dots I_j \dots I_m | j > 1, m \leq (D_b + 1) \cdot (D_a + 1) \cdot (D_c + 1) - 1\}$  is the set of indices of  $\Delta idx_{1D} > 0$ . The  $I_j$ -th to  $I_{j+1}$ -th points are placed in the  $idx_{1D_j}$ -th element of  $\{\Theta\}$ . In the meantime, set the  $idx_{1D_j}$ -th element of  $[\Lambda]$  equal to 1. Regarding the pillars generation, the 1D index is written as

$$idx_{1D} = d_a \cdot (D_b + 1) + d_b + 1 \quad (5)$$

Due to the high computational simplicity, detecting the jumps of  $\Delta idx_{1D}$  to segment 3D data is faster than using the exhaustive searching or the kd-tree neighbors searching algorithm. Besides, a corresponding binary matrix is generated in the process of segmentation, which may aid in detecting either the indices of non-empty pillars or the connected components (CC) when necessary.

**Figure 11b** shows the time performance of the linear index based segmentation procedure on a computer with the processor of Intel®371 Core™ i5–6600, RAM of 16 GB and GPU of Nvidia®GTX-960. As  $\delta\varphi$  and  $\delta\phi$  are both set as  $0.1^\circ$ , the time the segmentation procedure takes is linearly correlated with number of points. Notably, it can process 2 million points within 1 seconds (s). Given the same dataset, the processing time of the indexing operation is negatively and exponentially correlated with the angular resolutions. However, the efficiency is still satisfactory as it takes less than 1 s to handle 1.5 million points when both  $\delta\varphi$  and  $\delta\phi$  are set as  $0.05^\circ$ .

Using the procedure shown in **Figure 11a**, the target points  $\Phi$  and obstacle points  $\Psi$  in the  $\varphi - \phi - d$  space can be separately partitioned into numerous pillars and thus two 3D depth maps are generated as shown in **Figure 6**. It is noteworthy that the starting and the terminating points (see **Figure 11**) for calculating the linear indices of  $\Phi$  and  $\Psi$  are identical (i.e. the size of generated  $[\Lambda]$  or  $\{\Theta\}$  is the same). As such, there is a one-to-one correspondence between pillars of depth maps 1 and 2. The binary matrix  $[\Lambda]$  of depth map 2 can help determine the positions or indices of pillars that have target points (i.e.  $[\Lambda] == 1$ ). Then the obstacle points in the corresponding pillar of depth map 1 can be retrieved via the linear index and compared with the target points as noted in **Figure 6**. Let  $d_{min}$  bet the minimum distance of the closest obstacle point to the sight point. As previously mentioned in Section 2.2.2, when  $\delta\varphi$  and  $\delta\phi$  are small, the target points are invisible (marked in red in Figure) if their  $d$ -values exceed  $d_{min}$ , otherwise they are visible (marked in green in Figure).

### 3. Visibility equipment

#### 3.1 Overview

A growing number of diverse end users requires 3D data as many new outdoor and indoor application domains benefits from three dimensional (3D) maps. The new application domains include AV, smart cities, asset management, augmented virtual reality, as-built drawings, and even gaming industry. Active and passive

sensors are the main sensor types that can be used for scanning the surrounding environment and generating 3D data. Lidar scanners are the main active sensors, while optical cameras are the main passive sensors. The Lidar sensors are considered the standard sensors used for 3D scanning and mapping. Unlike passive sensors, Lidar scanners do not require an external source of light to illuminate the target. Thus, it can scan day or night under all lighting conditions and with a higher resiliency to scan in adverse weather conditions. The technological advances in Lidar scanners and its miniaturization are rapidly growing and can be deployed for several transportation application domains, such as AV, road furniture mapping, road condition assessment, and 3D visualization.

### 3.2 Lidar operation

A Lidar scanner is an active Remote Sensing Sensor (RSS) that uses light as the source of target illumination [22]. The Lidar unit emits a pulsed light beam or a continuous light wave that hits the object and reflects back to the sensor. The precise measurement of the range from the sensor to the target might follow one of two methods. The first method involves the accurate measurement of the Time of Flight, which is the time interval that has elapsed between the emission of a short (but intense) light pulse by the sensor and its return after being reflected from an object to the sensor. Then, the range is given by

$$R = v t / 2 \quad (6)$$

where  $R$  = range (m),  $v$  = speed of light (m/s), and  $t$  = time interval measured from the emission to reception of the light pulse (s), where  $v = \sim 299,792,458$  m/s.

The precision of the time measurement determines the precision of the range measurement which is given by

$$\Delta R = t \Delta v / 2 + v \Delta t / 2 \quad (7)$$

where  $\Delta R$  = range precision,  $\Delta v$  = speed of light precision, and  $\Delta t$  = time interval precision. This method is commonly used in most Lidar systems.

The precision of the time measurement determines the precision of the range measurement, with the high-end Lidar scanners (surveying scanners) capable of measuring single pulse timing to a 3 picosecond (ps) accuracy and thus achieving a 1 mm range resolution. The range precision is given by

$$R = (M\lambda + \Delta\lambda) / 2 \quad (8)$$

where  $M$  = integer number of wavelengths ( $\lambda$ ), and  $\Delta\lambda$  = fractional part of the wavelength.

Generally, laser scanners survey the surrounding environment by steering the light beam through a mirror or a prism mechanism to cover one direction (e.g. vertical direction). To provide a sequence of profiles around the vertical axis of the laser unit and generate a 3D point cloud of the area around the laser unit, a controlled and measured motion in another direction (e.g. azimuth direction in the static terrestrial laser scanning case) is applied. Nevertheless, if the laser unit is mounted on a moving platform, the controlled and measured motion in the azimuth direction may be substituted by the platform movement.

Beam divergence is yet another factor that affects the 3D point cloud generation [22]. The light beam is collimated when emitted from the laser unit, but as the light beam propagates, the beam radius or diameter increases and the increase is related

to the distance the beam travels. Beam divergence is an angular measure that relates this increase to the distance it travels. The divergence affects the footprint that is measured by the beam. Thus, the measured distance represents a wider area on the target and, in turn, will decrease the specificity of the measured distance, as it will miss any position variation within the footprint. The effect is further highlighted in the case of long-range sensors, such as those used in Airborne Laser Scanning (ALS). This explains the narrow beam divergence needed in the sensors used for ALS, which is typically 0.5 mrad or less.

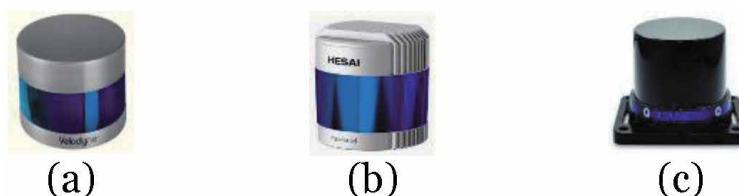
The Lidar sensor can be used in a mobile or static mode. In the mobile mode, the sensor is mounted on a moving platform such as AV or included as part of a 3D mapping system for mobile mapping and 3D map generation. In the static mode, also known as roadside Lidar for traffic monitoring, the sensor is mounted at a fixed location at an intersection to measure real-time traffic volume and speed. In the static mode, the sensor can use power over Ethernet as a versatile way of providing power to the sensor. The sensed data can then be classified using AI and big data algorithms as pedestrian/cyclist/vehicles objects, where a unique ID is then assigned to each classified object. The object can be continuously tracked using the position, direction, and speed. A complete solution that empowers a smart traffic flow management can be designed by integrating the Lidar within a perception software and IoT communications, thus improving mobility.

### 3.3 Types of Lidar scanners

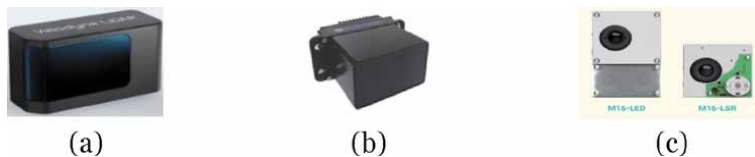
Spinning multi-beam Lidar scanners, which are relatively recent, have been introduced to meet AV industry requirements. Unlike 2D laser scanners that depend on the platform movement which covers the third dimension, the sensors have multiple beams with pairs of an emitter and a receiver for each beam. Each beam is oriented at a fixed vertical angle from the sensor origin [23]. The multi-beam mechanism spins mechanically around a spinning axis to cover 360° horizontal field of view. The frequency of the spinning axis frequency of the Velodyne Lidar sensors reaches 20 Hz, thus enabling a fast and rich 3D point cloud of the vehicle's environment and enhancing its 3D perception. Three examples of these sensors are shown in **Figure 12**.

The solid-state flash Lidar is yet another Lidar technology. This technology illuminates large areas simultaneously and measures the reflected energy on a photonic phased array in an analogy of the digital camera complementary metal-oxide-semiconductor sensor. Unlike the previously-discussed sensor data measurement mechanisms, solid-state Lidar sensors do not have any moving parts and the miniaturization allows on-chip lasers [23]. Three examples of these sensors are shown in **Figure 13**.

It is worth noting that the current rapid increase in laser sensor technology is being driven by the AV application domain. The solid-state Lidar sensors are very



**Figure 12.** Examples of spinning multi-beam laser sensors: (a) Velodyne VLP-32c, (b) Hesai Pandar64, and (c) Quanergy M8. Source: (a) [www.velodyne.com](http://www.velodyne.com) (b) [www.hesai.tech](http://www.hesai.tech), and (c) [www.quanergy.com](http://www.quanergy.com).



**Figure 13.**  
Examples of solid state Lidar: (a) Velodyne Velarray; (b) Quanergy S3, and (c) Leddartech M16-LSR. Source: (a) [www.velodyne.com](http://www.velodyne.com), (b) [www.quanergy.com](http://www.quanergy.com), and (c) [www.leddartech.com](http://www.leddartech.com).

promising as they are specifically designed for vehicle environment grid occupancy detection and collision avoidance. As the technology advances, however, it is anticipated that the solid-state Lidar technology will have a substantially positive effect on the 3D mapping field. Two Livox Lidar sensors are available: Mid-70 which has zero blind spot and Avia which has a detection range of up to 450 m along with the multi-scanning modes (repetitive and non-repetitive). These sensors better meet the needs for low-speed autonomous driving (Mid-70 sensor) and such applications as topographic surveying and mapping and power line inspection (Avia sensor). The sensors are shown in **Figure 14**.

### 3.4 New Lidar system

3D maps are a key infrastructure component needed for smart cities and smart transportation applications. Mobile mapping allows the generation of 3D maps for very large areas, which are considered cost-prohibitive to be mapped using conventional mapping equipment. Active sensors are the primary RSS in a number of mobile mapping systems. A mobile mapping system (MMS) is a moving platform (typically a vehicle) where the Lidar-based system is mounted on. This setup allows capturing data of large areas quickly, which is needed if conventional terrestrial mapping is to be used. Mapping-grade MMS normally achieves sub-meter accuracy, while survey-grade MMS achieves cm-level accuracy, where cost is typically US \$400 k and US\$1 M, respectively.

Several factors hinder the current mobile mapping systems usage for many user segments, including major investment, operational cost, unease of deployment, and required level of expertise. The design, development, and implementation of a new Lidar-based generic 3D mapping system has been carried out at the Department of Civil Engineering, Ryerson University [24, 25]. The developed system uses the relatively low-cost recently released RSS. The optimized selection of the sensors, the smart integration of the 3D mapping system components on both the hardware and software allow a higher level of versatility, ease of deployment, and substantial cost reduction, compared to commercially systems, while maintaining a comparable accuracy metrics. The system can be used on drones, cars, or as a stationary device. The characteristics of the new system are shown in **Table 1**.



**Figure 14.**  
Examples of the Livox Lidar sensors: (a) Livox Mid-70 and (b) Livox Avia. Source: [www.livoxtech.com](http://www.livoxtech.com).

Feature	Value
System accuracy	5 cm @ 50 m ( $1\sigma$ )
Weight (battery included)	1.5 kg
Diameter	10.3 cm
Height	15 cm
HFOV (VFOV)	360° (+15° to -15°)
Data	300,000 pts./sec (single return) 600,000 pts./sec (dual return)
Power consumption	19 W (Autonomy ~1.2 hr)
Operating Temperature:	-10°C to +60°C

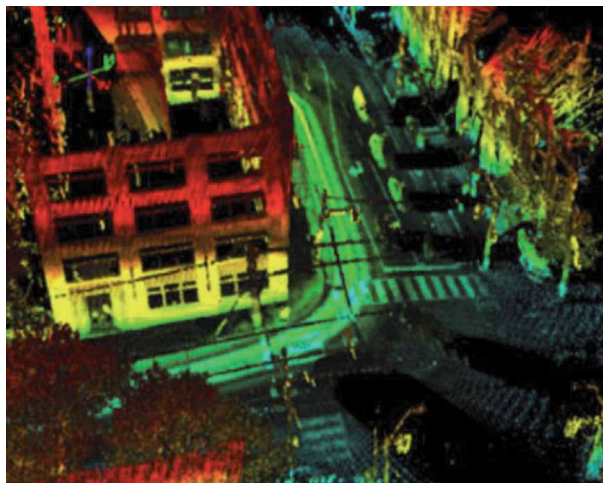
**Table 1.**  
*Features of new Lidar 3D mapping system.*

A sample of the 3D point cloud for a block at Ryerson University campus collected by the new system and colorized by height is shown in **Figure 15**. The figure clearly shows thin features such as overhead electricity wires, the traveled way, sidewalk, pedestrian crosswalk, traffic lights, and other vehicles.

The new Lidar-based system holds great potential for implementation in 3D map generation for AV. The drive behind the Lidar sensor technology advancement is to build a very low-cost, miniaturized sensor that can provide the AV with the ability of a robust 3D perception of the environment. A number of different competing factors need to be optimized, including sensor characteristics, such as range, range precision, horizontal and vertical fields of view, and size and power consumption. The main objective is to allow obstacle detection in a fast and reliable manner even at long ranges, thus ensuring safe vehicle operation.

### 3.5 Potential applications

The new Lidar sensors can benefit a multitude of application domains within the transportation sector. Those sensors are used to create the 3D maps that serve as the



**Figure 15.**  
*3D point cloud of part of Ryerson University campus, showing street furniture (color-coded by height).*



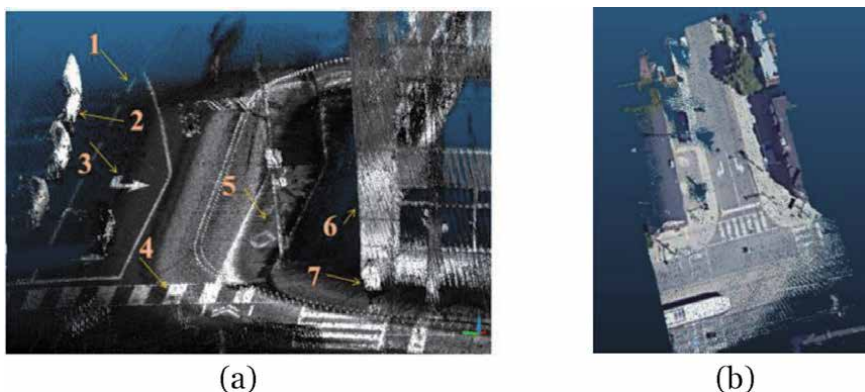
infrastructure needed for smart cities and smart transportation applications. Besides, Lidar sensors provide AV with its ability to perceive the 3D environment.

In addition, laser scanners can measure the amount of energy that has been reflected from the target after being illuminated with the scanner emitted pulse. The amount of reflected energy from each measured point constitutes the intensity as measured by the sensor. This measurement can prove valuable in a number of applications as asset management and lane marking extraction. The measured intensity depends on the target reflectivity, which affects the amount of energy reflected that can be detected by the sensor. As the target reflectivity decreases, the amount of reflected energy diminishes, thus weakening the signal returned to the sensor and deeming the target undetectable. The range and incidence angle to the target also affect the measured intensity.

Pavement markings have a different reflectance than the asphalt, thus allowing the automated extraction of the markings from the Lidar measured object intensities. This enables lane mapping and also serves as a tool that can be deployed for pavement marking assessment and evaluation. Street signs and furniture can be automatically extracted from the Lidar measured 3D point cloud. This is done using the fusion of the measured intensities, coupled with some machine learning algorithms that uses features geometry which is also measured by Lidar sensors.

A sample of a 3D point cloud intensity showing different features is depicted in **Figure 16**. The numbers in the figure refer to objects as follows: lane markings (1), vehicles (2), road surface arrows (3), zebra crossings (4), bike lane (5), a building (6), and pedestrian (7). Some highlighted features from the Lidar 3D point cloud intensity depicted in **Figure 16a**. As noted, the Lidar data intensity can be very useful in the automated feature extraction of street furniture, pavement marking assessment and evaluation which acts as an enabler for automated asset management. Note that a true color 3D Lidar point cloud can be produced by the fusion of the Lidar 3D point cloud and optical imagery, as shown in **Figure 16b**.

The fusion of Lidar 3D point cloud and optical imagery can prove very useful in such applications as digital twin and 3D visualization with and enhanced visual appeal. In addition, the different modality between Lidar data and optical imagery improves automated 3D point cloud classification as the classification process uses object geometry coupled with reflectance. The continued advances in Lidar sensor technologies, IoT, AI, and big-data algorithms and with faster wireless communications technologies that support cellular data networks 5G (even the 6G networks expected in 2030s), with data transfer rate of speeds of



**Figure 16.**  
*Lidar 3D point cloud: (a) intensity and (b) true color.*

~95 Gb/s, the dream of mass deployment of an advanced driver assistance system level 5 can become a reality.

## **4. Visibility methodologies**

### **4.1 Infrastructure visibility**

#### *4.1.1 High-precision 3D maps*

High-precision maps have been identified as one of the key technologies for autonomous driving [26]. These maps need to be purpose-built and highly accurate, with great level of details, and be updated in real-time. These maps, which are particularly made for autonomous driving, provide true-ground-absolute accuracy, normally at centimeter-level (5 cm or better), and contain details organized in multiple map layers, such as base map, geometric map, semantic map, map priors and real-time knowledge [27]. Technologies that can be leveraged for creating such maps include aerial imagery, aerial Lidar data, mobile Lidar mapping, and Unmanned Aerial Vehicles equipped with lightweight sensors. On the ground, HP maps are mainly created using vehicles equipped with high-tech instruments [26]. Efforts on HP mapping in 3D include Civil Maps, HERE, DeepMap, TomTom HP Maps, Mobileye, Uber Localization and Mapping, and Mapper, while most of them follow crowdsourcing models.

In addition to supporting localization and navigation, HP 3D maps can be leveraged for lane network construction, dynamic object detection, prediction of motion of vulnerable road users, and visibility analysis for the correlation between environmental visibility and navigational comfort under autonomous driving [28, 29]. The so-called self-healing mapping systems allows a detailed inventory of road features and objects, part of road infrastructure, on the side of the road for real-time determination of infrastructure visibility.

To provide dynamic routing for any situation, Geographic Information System (GIS) requires updated road network data, real-time traffic information, vehicle current location, and destination. Real-time vehicle location is obtained using Radar and Camera sensors using localization techniques such as simultaneous localization and mapping (SLAM) that can localize vehicle with high precision and map its exact location with respect to the surrounding environment. Localization can not only determine vehicle location, but also map landmarks to update HP maps. Depending on the use of sensors, the algorithm can be Lidar-based, Lidar and camera integrated, or camera-based.

Building and maintaining detailed HP maps in advance presents a very appealing solution for autonomous driving. However, exiting technologies may at the best provide near real-time mapping, and may have missed critical information, such as road markings and dynamic changes of road infrastructure. Real-time SLAM provides better solution to dynamic updating of HP maps for locating vehicles and mapping the surrounding environment.

#### *4.1.2 Sight obstacle detection*

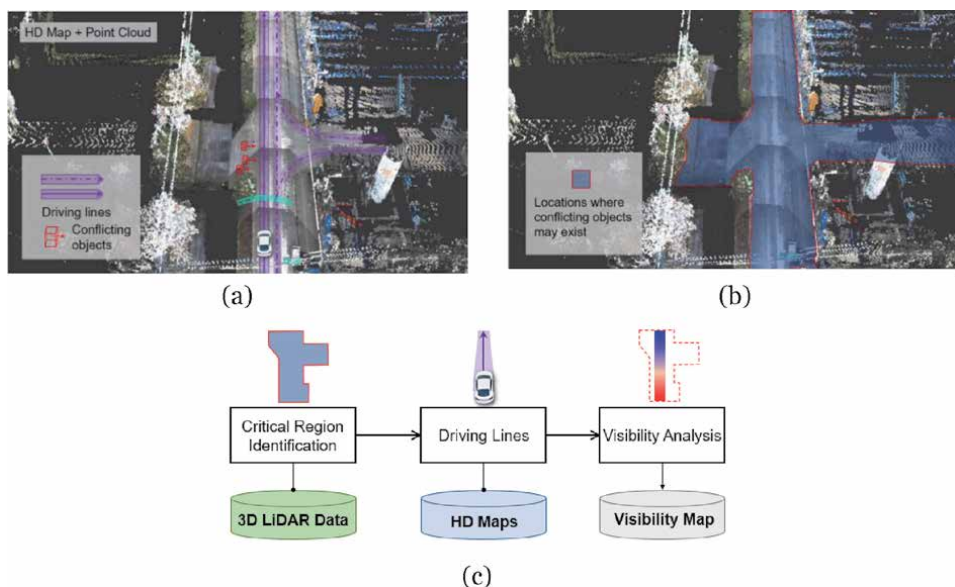
The technique for estimating visibility of target points described in Section 2.2 can be extended to detect sight obstacles that restrict infrastructure visibility (e.g. traffic signs). Specifically, as shown in **Figure 7a**, the visibility of the target points are determined by comparing their distances to the sight point with the minimum distance (i.e.  $d_{min}$ ) of the closest obstacle point to the sight point in each pillar.

The sight obstacles can be detected in a similar way. The respective depth maps of the obstacle points and target infrastructure points are first generated following the steps presented in Section 2.2. Then, at the stage of comparing the depth maps, another parameter is calculated, as shown in **Figure 7a**. Let  $d_{max}$  be the maximum distance of the target points from the sight point. The obstacle points farther than  $d_{max}$  will not affect the visibility of the target points. On the contrary, the obstacle points whose  $d$ -values lie between  $d_{min}$  and  $d_{max}$  will obstruct the target infrastructure points whose  $d$ -values exceed  $d_{min}$ . The sight obstacle detection is completed when all pillars are estimated. **Figure 7b** shows the detected sight obstacles that affect the visibility of a traffic sign. In this case, when combined with the techniques for automatically identifying traffic signs from MLS data, the visibility estimation procedure can help understand the visibility of traffic signs to the AV along a road corridor.

## 4.2 Traffic visibility

### 4.2.1 Overview

Guaranteeing adequate traffic visibility is crucial to the safe operation of AV. If the sight lines to the conflicting objects (e.g. vehicles, pedestrians, and cyclists) are obstructed by some obstacles (**Figure 17a**), the AV may not identify the objects at a safe distance, which may lead to a collision. It is quite challenging for AV vehicles to predict where pedestrians or cyclists may be present in a complex road scene in real-time [30]. In this case, it is meaningful to identify the positions where road users may exist in real-world using dense MLS points and investigate the visibility of these locations to the AV in advance (see **Figure 17a** and **b**). Then the georeferenced visibility information can be incorporated into HP maps, which may help AV make some proactive measures to reduce the collision risk at locations with unsatisfactory traffic visibility.



**Figure 17.** Traffic visibility estimation for AV: (a) reasons for estimating traffic visibility, (b) critical positions, and (c) general workflow of traffic visibility estimation.

As shown in **Figure 17c**, the general process of traffic visibility analysis comprises two main components: (1) critical positions and (2) driving lines. The critical positions correspond to the target points in the VE process (see **Figure 4**). The driving lines in HP maps which aid in navigating AV can help derive sight points in this case. Traffic visibility estimation in autonomous driving involves estimating the visibility of some critical positions along the pre-defined driving lines to the AV and generating the visibility map.

#### 4.2.2 Critical position identification

In this example, the planar ground points are identified as the locations where vehicles, pedestrians, or cyclists may exist. A common workflow shown in **Figure 8a** is used to extract the critical positions. More specifically, MLS data are first partitioned into a number of pillars with the linear-index based segmentation technique illustrated in **Figure 11**. Next, a pillar-wise filtering is performed that the points with  $\delta_h$  (user-defined, e.g., 0.2 m) higher than the lowest point are removed. The remaining points are considered as the rough ground points. Then, the kd-tree data structure is applied to find k-nearest neighbors for each data point. The neighboring points are used to derive the normal vector  $\vec{\xi}$  of each point [31]. Let  $\gamma$  be the angle between  $\vec{\xi}$  and the vertical direction (i.e.  $(0, 0, 1)^T$ ). The points with  $\gamma \leq 5^\circ$  are considered as the horizontal and planar points.

Because we mainly focus on the locations where the road users may exist and may conflict with the AV, a distance-based segmentation method is applied to remove some isolated planar point clusters. The point cluster with the largest size is determined as the ‘critical positions’. MLS data of a 500 m long urban road section is used to illustrate the process of VE. The extracted critical positions using the procedure shown in **Figure 8a** are marked in red in **Figure 8b**. The critical positions covers pavement surface, sidewalks, and some planar surfaces connected to the pavement. It is also recommended to use the semantic segmentation techniques powered by deep learning to identify the critical positions more accurately.

#### 4.2.3 Visibility estimation

The driving lines of AV are also plotted in **Figure 8b**. Each driving line is composed of many consecutive nodes. At each node, suppose that the sight point overlap with the node on the horizontal plane. However, the elevation of sight point  $h_s$  is adjustable to accommodate different height of Lidar sensors mounted on the autonomous driving vehicles. Then, the VE procedure described previously is applied to estimate the visibility of critical positions to the sight point. **Figure 9** presents the VE results at a single node. The visible and invisible critical positions are marked in green and red, respectively, while the potential obstacle points are marked in gray. In **Figure 9**, the range of  $\varphi$  is  $[-180^\circ, 180^\circ]$ , where both the front and rear positions are estimated. Different limits can be set on the range of  $\varphi$  to simulate varied horizontal viewing angles. The VE procedure at a single node can be extended to estimate the locations where the sight triangle is clear at intersections [16]. Specifically, if there are invisible (red) regions (see **Figure 9**) inside the sight triangle, it means the sight triangle is not clear.

The VE procedure is executed along the driving lines node by node (see **Figure 10a**) to gain a better understanding of traffic visibility for AV. In this phase, the variables involved are  $h_s = 1.6$  m,  $d_{view} = 100$  m,  $\varphi_{view} = -60$  to  $60^\circ$ ,  $\phi_{view} = -30$  to  $30^\circ$ ,  $\delta\varphi = 0.1^\circ$ , and  $\delta\phi = 0.1^\circ$ . The users may also adjust  $d_{view}$ ,  $\varphi_{view}$  etc. to investigate different situations.

The VE results are visualized in **Figure 10**. In this example, two types of quantitative information are derived based on VE results at each node. **Figure 10b** maps the cumulative times of invisibility  $N_{invis}$  at a point level. The initial  $N_{invis}$  of each target point is zero. During the process of VE,  $N_{invis} = N_{invis} + 1$  each time when the target point is not seen by the ego vehicle. A large  $N_{invis}$  means the target point is invisible to the AV for many times. The magnitude of  $N_{invis}$  is visualized with colors in **Figure 17b** which may help to identify the possible locations of the blind areas for the ego vehicles. As marked with rectangles, the visibility of these locations are poorer, which means that the AV may need to decelerate when approaching these locations to avoid potential right-angle collisions. Also, the blind-area results can be estimated in conjunction with collision data in future studies to better understand road safety.

**Figure 10c** shows the variation of the visibility ratio  $V_r$  along the driving lines. The  $V_r$  measures the ratio of visible targets to all targets at a single position. A low  $V_r$  may indicate that a majority of target points are invisible to the AV. Because  $V_r$  can be integrated with the driving lines, the AV can know where the visibility ratio is relatively low based on its location. In that case, the ego vehicle can decelerate in advance at a location with very low  $V_r$  to reduce the probability of colliding with a running pedestrian from a blind area.

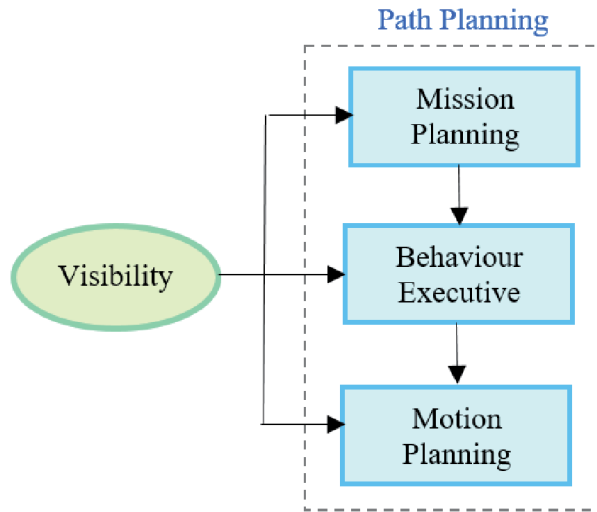
## 5. Decision

### 5.1 Path planning

This chapter has addressed so far the visibility element of autonomous vehicles, which is the main focus of the chapter. It is useful, however, to highlight the decision element (see **Figure 1**), which involves path planning. This element is considered to be the main challenge of autonomous vehicles. Path planning enables the autonomous vehicle to find the safest, most convenient, and most economic route between two points. One of the main modeling approaches to define such a route is Model Predictive Control (MPC). There is a variety of MPC, but basically the model solves a finite-time constrained optimal control problem in a receding horizon using nonlinear optimization. An application of path planning and control for overtaking on two-lane highways using nonlinear MPC can be found in [32].

There are two approaches for path planning (hierarchical and parallel) [33]. In the hierarchical approach the autonomous vehicle completes long-term mission and reduce the workload of motion planning. The input higher-level mission is decomposed into sub-missions and then passed to the next level down. The hierarchical model helps to resolve many complicated problems, yet it might slow down the work of a vehicle's feedback control and complicate the performance of sophisticated maneuvers. In the parallel approach, the tasks are more independent and can proceed simultaneously. In this approach, each controller has dedicated sensors and actuation mechanisms. The advantages of this approach: (1) the controllers are running at high frequency, which makes them safe and stable, (2) a high level of smoothness and performance is achieved by the controllers, and (3) the approach is relatively inexpensive and does not require using complicated motion planning devices. However, for some purposes, the hierarchical approach is more efficient. **Figure 18** illustrates the hierarchical approach, based on [33].

Extensive research has been conducted for path planning for automated parking, which is now a reality in many cities [34–38]. Automated parking aims to eliminate the influence of human factors, improve the quality and accuracy of control, and reduce the maneuver time by optimizing vehicle path in restricted parking zones. The vehicle to anything technology (V2X) allows communication between AV and

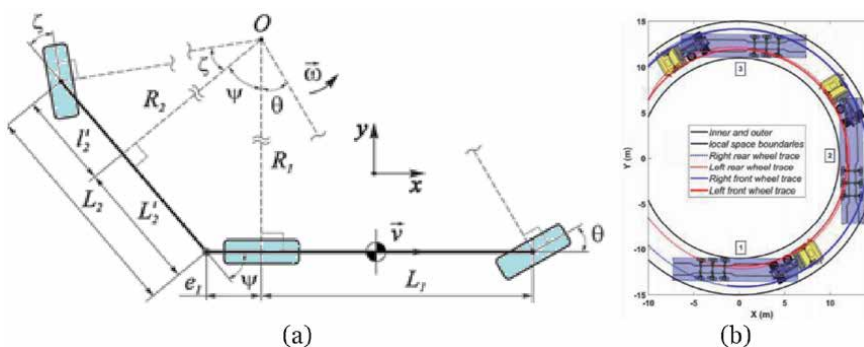


**Figure 18.**  
*Hierarchical approach for path planning.*

a building through data exchange to find an unoccupied space and generate the route to a destination. For large-sized trucks, the tasks involve predicting stable and safe passing on road curves and forecasting precise control for docking. Numerous approaches that consider different control strategies, sensory means, and prediction algorithms (e.g. geometric, fuzzy, and neural) for predicting vehicle parking path have been developed.

Recently, the bicycle kinematic models for vehicle motion have been used for path planning of automated parking [39]. Three basic types of vehicles were considered: passenger car, long wheelbase truck, and articulated vehicles with and without steered semitrailer axes. The authors presented a system of differential equations in matrix form and expressions for linearizing the nonlinear motion equations that increased the speed of finding the optimal solution. An original algorithm that considered numerous constraints was developed for determining vehicle permissible positions within the closed boundaries of the parking area using nonlinear MPC that finds the best trajectories.

**Figure 19** shows the kinematics of the curvilinear motion and the simulation results that validated the proposed model. Note that in this study, kinematic vehicle models were used instead of dynamic models. Kinematic models assume that no slip



**Figure 19.**  
*Modeling and simulation of autonomous docking of tractor-semitrailer vehicle with semitrailer's steered axles: (a) kinematics of curvilinear motion and (b) simulation results [39].*

occurs between the wheels and the road. This assumption is reasonable for vehicles moving at low speeds, which is the case for parking. However, dynamic models should be used for path planning of autonomous vehicles on highways since they are more accurate.

## 5.2 Intelligent car-following

Autonomous vehicles and connected automated vehicles (CAV) with advanced embedded technology can deliver safe and effective traffic movements [40]. However, there will be a transition period when AV and human-driven vehicles would share the common road space. Therefore, it would be extremely critical to organize the interactions which will likely be generated from that vehicle mix. Due to the anticipated asymmetric response of these two modes of vehicles, many possible combinations of these vehicles are possible. These combinations may give rise to interactions due to longitudinal and transverse movements.

To ensure safe interactions, the intelligent vehicles may include adaptive cruise control (ACC) and cooperative adaptive cruise control (CACC) systems. These systems mainly assist the acceleration control for the longitudinal movements. The systems control the acceleration based on the distance gap and the speed difference between the current vehicle and the vehicle ahead (leader), where the vehicle accelerates and decelerates based on the speed changes of the leader. For the ACC system, the distance and speed are obtained using radar, Lidar, or video cameras. For the CACC system, V2V communications are used to share the acceleration, deceleration, braking capability, and vehicle positions [41]. This communication significantly shortens the time headway (0.5 s) of the CACC vehicle compared to the ACC vehicle (1.4 s).

To obtain deeper insights into car-following behavior, micro simulation studies were conducted to estimate the impacts of AV and CAV using a variety of assumptions [41]. Car-following models involving intelligent vehicles were developed to evaluate their traffic impacts on such aspects as capacity and level of service, traffic stability, travel time, and vehicle speed. Several studies estimated the energy, environmental, and safety impacts using surrogate safety measures, such as travel speed, time-to-collision, and post-encroachment-time.

New car-following models for AV and CAV were developed by modifying available traditional car-following models to mimic the intelligent vehicle characteristics. The Intelligent Driver Model (IDM) [42], a time-continuous car-following model for the simulation of freeway and urban traffic, is the most commonly used model for simulations of intelligent vehicles. The basic function of the model is given by

$$a_{IDM}(s, v, \Delta v) = \frac{dv}{dt} = a \left[ 1 - \left( \frac{v}{v_o} \right)^\delta - \left( \frac{s^*(v, \Delta v)}{s} \right)^2 \right] \quad (9)$$

$$s^*(v, \Delta v) = s_o + vT + \frac{v\Delta v}{2\sqrt{ab}} \quad (10)$$

where  $s$  = current distance gap to the leader (m),  $s_o$  = minimum distance gap (m),  $v$  = current speed (m/s),  $v_o$  = desired safety speed (m/s),  $\Delta v$  = speed difference between the current vehicle and the leader (m/s),  $\delta$  = parameter that determines the magnitude of acceleration decrease (usually set equal to 4),  $T$  = desired safety gap (s),  $a$  = maximum acceleration (m/s<sup>2</sup>), and  $b$  = comfortable deceleration rate (m/s<sup>2</sup>). Note that  $s^*(v, \Delta v)$  is the desired distance gap (m).

Another car-following model that is commonly used models for conducting cooperative intelligent-vehicle simulations is the MICroscopic Model for Simulation

of Intelligent Cruise Control (MIXIC) [43]. Both IDM and MIXIC have been used as the benchmark models for combined AV and human-driven vehicles for various market penetration rates ranging from 0 to 25% [44, 45]. However, detailed calibration procedures of model parameters used in these models are not available due to the limited or unavailability of empirical data.

On these lines, substantiating vehicle's longitudinal motion in mixed traffic conditions is even more critical using available car-following model parameters. In the present context, the car-following models used in simulation packages need to be suitably modified. Further, most of these models are straightforward in their approach with a limited set of parameters. The simplification of available model parameters might not affect the simulation results to the extent the intelligent vehicles may impact the traffic streams in real-field experiments. For example, most car-following models tend to have fixed parameters and will average out driving characteristics. If such models are used for AV, they can reflect neither the driving style of the vehicles' actual drivers nor the contexts in which they drive.

In this direction, the intelligent car-following behavior models, which can concurrently consider all necessary parameters and adapt to specific field actions must successfully integrate possible mixed traffic scenarios. Hence, there is a strong need to modify the existing conventional models to handle the higher order of stochasticity due to mixed driving conditions. Alternatively, data-driven methodologies can be used to provide high-quality trajectory data for varying scenarios of automation. These data may be generated from experiments by creating test beds or even using driving simulators with human-machine interaction. For this purpose, the trajectory data available from simulation models may also be useful. Considering the recent highlights of AI, even intelligent car-following models can benefit from data learning methodologies. Clearly, the impact analysis of intelligent vehicles should consider the many uncertainties due to mixed traffic conditions. Although a few new models have been developed or modified to incorporate car-following and lane-changing attributes of intelligent vehicles, empirical data are warranted for model calibration.

## **6. Conclusions**

This chapter has presented an overview of the visibility-based technologies and methodologies for autonomous driving. Based on this overview, the following comments are offered:

1. The Lidar height is a key parameter that affects the visibility of the road ahead and in turn sight distance that influences the design of horizontal and vertical alignment as well as intersections. To ensure safety of AV when operating with human-driven vehicles on existing highways, the Lidar height should generally be larger than the current design heights for passenger cars and less than that for trucks.
2. The new Lidar system developed at Ryerson University holds great potential for generating 3D maps for AV. The system is cheap, allows AV a robust 3D perception of the environment, and allows obstacle detection in a fast and reliable manner even at long ranges.
3. The infrastructure and traffic visibilities can be estimated for autonomous vehicles based on a combination of MLS data and the driving lines of HP maps. The generated visibility map can be incorporated into the HP maps and help



the autonomous vehicle develop some proactive speed control strategies at locations where the visibility is unsatisfactory.

4. The uncertainty of the input variables can substantially affect the reliability of the design elements. Therefore, reliability analysis should be incorporated in all the tasks related to SD determination to ensure safe operation of the autonomous vehicles. Extensive uncertainty research has been conducted on the required SD and similar research is lacking for the available SD determined using the Lidar.
5. Geographic information systems provide support for route planning and real-time and dynamic routing/navigation of AV using GNSS, localization techniques, and HP maps. The GIS navigation helps to guide the autonomous vehicle along the best route to its destination safely.
6. The intelligent car-following strategies are expected to show good performance in simulation environments. However, in the calibration using real-field conditions/experiments, the input to simulation models would not be beneficial as it is limited to the available parameters that correspond to specific conditions. The parameters should be based on the surrounding environmental factors for all possible combinations of autonomous and human-driven vehicles. This is a mammoth challenge due to the limited empirical data.
7. Another challenge of autonomous driving is related to the complexity of the software algorithms needed to process the large amount of information coming into the AV (estimated at 1 billion lines of code) which is then used to make decisions about the proper action. The software complexity is compounded by the fact that AV will operate on roads full of unpredictable drivers of human-driven vehicles.

## **Acknowledgements**

This chapter is financially supported by the Natural Sciences and Engineering Research Council of Canada (NSERC).

## **List of acronyms**

ACC	adaptive cruise control
AI	artificial intelligence
ALS	airborne laser scanning
AV	autonomous vehicles
CACC	cooperative adaptive cruise control
CAV	connected automated vehicles
DSD	decision sight distance
GIS	geographical information system
GNSS	global navigation satellite system
HP	high-precision
IDM	intelligent driver model
IoT	internet of things
ISD	intersection sight distance
Lidar	light detection and ranging

LOS	line of sight
MLS	mobile laser scanning
MSS	mobile mapping system
PSD	passing sight distance
SD	sight distance
SLAM	simultaneous localization and mapping
SSD	stopping sight distance
RSS	remote sensing sensor
VE	visibility estimation
V2I	vehicle-to-infrastructure
V2V	vehicle-to-vehicle
1D	one dimensional
3D	three dimensional

## **Author details**

Said Easa<sup>1\*</sup>, Yang Ma<sup>2</sup>, Ashraf Elshorbagy<sup>1</sup>, Ahmed Shaker<sup>1</sup>, Songnian Li<sup>1</sup> and Shriniwas Arkatkar<sup>3</sup>

<sup>1</sup> Ryerson University, Toronto, Canada

<sup>2</sup> Southeast University, Nanjing, China

<sup>3</sup> Sardar Vallabhbhai National Institute of Technology, Surat, Gujarat, India

\*Address all correspondence to: [seasa@ryerson.ca](mailto:seasa@ryerson.ca)

## **IntechOpen**

---

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] Shladover S. Connected & automated vehicle systems: Review. *Intelligent Transp. Sys. J.*, 22(3), 2018.
- [2] Allied Market Research. *Autonomous Vehicle Market by Level of Automation, Component, and Application*. 2019. <https://www.alliedmarketresearch.com/autonomous-vehicle-market>.
- [3] NCHRP *Connected & autonomous vehicles and transportation infrastructure readiness*. Project 20–24(111), TRB, Washington, DC, 2017.
- [4] Mohamed A. Literature survey for autonomous vehicles: Sensor fusion, computer vision, system identification, and fault tolerance. *Inter. J. Auto. Control*, 12(4), 2018.
- [5] Al-Qaysi Q, Easa S M, Ali N. Proposed Canadian automated highway system architecture: object-oriented approach. *Can. J. Civ. Eng.*, 30(6), 2003.
- [6] Shladover S. Automated vehicles for highway operations (automated highway systems). *IME Proceedings, Part I: Journal of Systems and Control Engineering*, 219(1), 53-75, 2005.
- [7] Easa S M. *Automated highways*. In Encyclopedia of Electrical and Electronics Engineering, John Wiley & Sons, New York, N.Y., 2007.
- [8] American Association of State Highway and Transportation Officials. *A policy on geometric design of highway and streets*. AASHTO, Washington DC, 2008.
- [9] Easa S M. *Geometric design*. In Civil Engineering Handbook, W.F. Chen and J.Y. Liew eds., CRC Press, Boca Raton, FL, Chapter 63, 2002, 1-39.
- [10] Transportation Association of Canada. *Geometric design guide for Canadian roads*. TAC, Ottawa, Ontario, 2017.
- [11] Khoury J et al. Initial investigation of the effects of AV on geometric design. *Journal of Advanced Transportation*, Vol. 2019, 2019, 1–10
- [12] Ma Y, Zheng Y, Cheng J, Zhang Y, and Han W. A convolutional neural network method to improve efficiency and visualization in modeling driver's visual field on roads using MLS data. *Transportation Research Part C: Emerging Technologies*, 106, 2019, 317–344.
- [13] Ma Y, Zheng Y, Cheng J, and Easa S M. Analysis of dynamic available passing sight distance near right-turn horizontal curves during overtaking using Lidar Data. *Can J Civ. Eng.*, 2019.
- [14] Ma Y, Zheng Y, Cheng J, and Easa S M. Real-time visualization method for estimating 3D highway sight distance using Lidar data. *J. Transp. Eng. - Part A: Systems*, 2019.
- [15] Gargoum S A, El-Basyouny K, and Sabbagh J. Assessing Stopping and Passing Sight Distance on Highways Using Mobile Lidar Data. *J. Comput. Civ. Eng.*, 32(4), 04018025, 2018.
- [16] Jung J, Olsen M J, Hurwitz D S, Kashani A G, and Buker K. 3D virtual intersection sight distance analysis using Lidar data. *Transportation Research Part C: Emerging Technologies*, 86(JAN.), 563–579, 2018.
- [17] Shalkamy A, El-Basyouny K, and Xu H Y. Voxel-Based Methodology for Automated 3D Sight Distance Assessment on Highways using Mobile Light Detection and Ranging Data. *Transportation Research Record*, 2674(5), 587–599, 2020. doi:10.1177/0361198120917376.
- [18] Zhang S X, Wang C, Lin L L, Wen C L, Yang C H, Zhang Z M, and Li J. Automated visual recognizability evaluation of traffic sign based on 3D

- Lidar point clouds. *Remote Sensing*, 11 (12), 2019. doi:10.3390/rs11121453.
- [19] MATLAB. 2020. sub2ind: Convert subscripts to linear indices. Retrieved from <https://www.mathworks.cn/help/matlab/ref/sub2ind.html?lang=en> (accessed on November 1, 2020).
- [20] OpenDSA. 2020. CS3 Data Structures and Algorithms. 2020, <https://opensa-server.cs.vt.edu/ODSA/Books/CS3/html/LinearIndexing.html> (accessed on November 1, 2020).
- [21] Ma Y, Zheng Y, Easa S M, and Cheng J. Semi-automated framework for generating cycling lane centerlines on roads with roadside barriers from noisy MLS data. *ISPRS J. Photo. Remote Sensing*, 167, 2020, 396–417.
- [22] Shan J, and Toth C K. *Topographic laser ranging and scanning: Principles and processing*. CRC Press, Boca Raton, FL, 2018.
- [23] Elshorbagy A. *A Crosscutting Three-modes-of-operation Unique LiDAR-based 3D Mapping System, Generic Framework Architecture, Uncertainty Predictive Model and SfM Augmentation*. Doctoral Dissertation, Department of Civil Engineering, Ryerson University, Toronto, Canada, 2020.
- [24] Shaker A and Elshorbagy A. Systems and methods for multi-sensor mapping. Application Number 62/889,845–24440–P58819US00, 2019.
- [25] Shaker A and Elshorbagy A. Systems and methods for multi-sensor mapping using a single device that can operate in multiple modes. PCT Patent Application No. PCT/CA2020/051133, 2019.
- [26] Seif H G and Hu X. Autonomous driving in the iCity—HD maps as a key challenge of the automotive industry. *Engineering*, 2(2), 2016, 159–162.
- [27] Vardhan H. HD Maps: New age maps powering autonomous vehicles. *Geospatial World*, 2017 <https://www.geospatialworld.net/article/hd-maps-a-utonomous-vehicles> (accessed on November 1, 2020).
- [28] Chou F C, Lin T H, Cui H, Radosavljevic V, Nguyen T, Huang T K, and Djuric N. Predicting motion of vulnerable road users using high-definition maps and efficient convnets. 2019, arXiv preprint arXiv:1906.08469.
- [29] Morales Y, Even J, Kallakuri N, Ikeda T, Shinozawa K, Kondo T, and Hagita N. Visibility analysis for autonomous vehicle comfortable navigation. *IEEE International Conference on Robotics and Automation*, 2014, 2197–2202.
- [30] Ahmed S. Pedestrian/cyclist detection and intent estimation for AV: A survey. *App. Sc.*, 9, 2019.
- [31] Yang B S, Liu Y, Dong Z, Liang F X, Li B J, and Peng X Y. 3D local feature BKD to extract road information from mobile laser scanning point clouds. *Journal of Photogrammetry and Remote Sensing*, 130, 2017, 329–343, doi: 10.1016/j.isprsjprs.2017.06.007.
- [32] Easa S M and Diachuk M. Optimal speed plan for overtaking of a autonomous vehicles on two-lane highways. *J. Infrastructures*, 5(44), 2020, 1–25.
- [33] Ryabchuk P. How Does Path Planning for Autonomous Vehicles Work? 2020, <https://dzone.com/users/3246906/paulryabchuk.html> (accessed on November 1, 2020).
- [34] Lee H, Chun J, and Jeon K. Autonomous back-in parking based on occupancy grid map and EKF SLAM with W-band radar. *Proc., International Conference on Radar*, Brisbane, Australia, 2018, 1–4.
- [35] Lin L and Zhu J J. Path planning for autonomous car parking. In *Proceedings of the ASME Dynamic Systems and*

Control Conference, Vol. 3, Atlanta, GA, USA, 2018.

*Intelligent Transportation Systems*, 7(4), 2006, 429–436.

[36] Kiss D and Tevesz G. Autonomous path planning for road vehicles in narrow environments: An efficient continuous curvature approach. *J. Adv. Transp.*, 2017.

[44] Kesting A, Treiber M, Schönhof M, and Helbing D. Adaptive cruise control design for active congestion avoidance. *Transportation Research Part C: Emerging Technologies*, 16(6), 2008, 668–683.

[37] Wang Y, Jha D K, and Akemi Y. A two-stage RRT path planner for automated parking. Proc., *13th IEEE Conference on Automation Science and Engineering*, Xi'anChina, 2017, 496–502.

[45] Talebpour A and Mahmassani H S. Influence of connected and autonomous vehicles on traffic flow stability and throughput. *Transportation Research Part C: Emerging Technologies*, 71, 2016, 143–163.

[38] Ballinas E, Montiel O, Castillo O, Rubio Y, and Aguilar L T. Automatic parallel parking algorithm for a car-like robot using fuzzy PD+1. *Control. Eng. Lett.*, 26, 2018, 447–454.

[39] Diachuk M, Easa S M, and Bannis J. Path and control planning for autonomous vehicles in restricted space and low speed. *J. Infrastructures*, 5(4), 2020, 1–27.

[40] Greer H, Fraser L, Hicks D, Mercer M, and Thompson K. *Intelligent transportation systems benefits, costs, and lessons learned*. U.S. Dept. of Transportation, ITS Joint Program Office, 2018.

[41] Wooseok D, Omid M R, Luis M-M. Simulation-based connected and automated vehicle models on highway sections: A Literature review. *Journal of Advanced Transportation*, 2019. <https://doi.org/10.1155/2019/9343705>.

[42] Treiber M, Hennecke A, and Helbing D. Congested traffic states in empirical observations and microscopic simulations. *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, 62(2), 1805–1824, 2000.

[43] Van Arem B, Van Driel C J, and Visser R. The impact of cooperative adaptive cruise control on traffic-flow characteristics. *IEEE Transactions on*



# Selected Issues and Constraints of Image Matching in Terrain-Aided Navigation: A Comparative Study

*Piotr Turek, Stanisław Grzywiński and Witold Buzantowicz*

## Abstract

The sensitivity of global navigation satellite systems to disruptions precludes their use in conditions of armed conflict with an opponent possessing comparable technical capabilities. In military unmanned aerial vehicles (UAVs) the aim is to obtain navigational data to determine the location and correction of flight routes by means of other types of navigational systems. To correct the position of an UAV relative to a given trajectory, the systems that associate reference terrain maps with image information can be used. Over the last dozen or so years, new, effective algorithms for matching digital images have been developed. The results of their performance effectiveness are based on images that are fragments taken from source files, and therefore their qualitatively identical counterparts exist in the reference images. However, the differences between the reference image stored in the memory of navigation system and the image recorded by the sensor can be significant. In this paper modern methods of image registration and matching to UAV position refinement are compared, and adaptation of available methods to the operating conditions of the UAV navigation system is discussed.

**Keywords:** digital image processing, image matching, terrain-aided navigation, unmanned aerial vehicle, cruise missile

## 1. Introduction

Global navigation satellite systems are widely used in both civil and military technology areas. The advantage of such systems is very high accuracy in determining the coordinates, however, the possibility of easy interference precludes their use in conditions of armed conflict with an opponent equipped with comparable technical capabilities. In the case of military autonomous unmanned aerial vehicles (UAVs), in particular cruise missiles (CM), the aim is therefore to determine navigation data for specifying the position and correcting the flight paths by means of other types of navigation and self-guidance systems.

Such systems are usually based on inertial navigation systems (INS) which use accelerometers, angular rate gyroscopes and magnetometers to provide relatively accurate tracking of an object's position and orientation in space. However, they are exposed to drift and systematic errors of sensors, hence the divergence between the actual and the measured position of the object is constantly increasing with time. This results in a significant navigational error.

Therefore, two types of systems designed to correct the position of an object in relation to a given trajectory are normally used in the solutions of the UAV/CM navigation and self-guidance systems. The first group contains systems whose task is to determine the position on the basis of data obtained from radio altimeters, related to reference height maps. Such systems include, for example: TERCOM (terrain contour matching), used in Tomahawk cruise missiles, SITAN (Sandia inertial terrain-aided navigation), using terrain gradients as input for the modified extended Kalman filter (EKF) estimating the position of the object, and VATAN (Viterbi-algorithm terrain-aided navigation), a version of the system based on the Viterbi algorithm and characterised – in relation to the SITAN system – with lower mean square error of position estimation [1–5]. The main disadvantage of these solutions is the active operation of measuring devices, which reveals the position of the object in space and eliminates the advantages associated with the use of a passive (and therefore undetectable) inertial navigation system. The second group consists of systems associating reference terrain maps with image information obtained by means of visible, residual or infrared light cameras [6, 7]. Such systems include the American DSMAC (digital scene matching area correlator), also used in Tomahawk missiles [8, 9], and its Russian counterpart used in Kalibr (aka Club) missiles. Their advantage is both the accuracy of positioning and the secrecy (understood as passivity) of operation.

Due to the dynamic development of UAVs/CMs equipped with navigation systems operating independently of satellite systems and a number of problems associated with the implementation of the discussed issue, the assessment on the sensitivity of the selected methods to environmental conditions and constraints in the measurement systems, which often negatively affect the results obtained, has been carried out. The essence of the work is to consider issues related to the processing of image information obtained from optical sensors carried by UAV/CM and its association with terrain reference images. In particular, issues of the correctness of image data matching and the limitations of the possibilities of their similarities' assessment are considered. The article compares modern image matching methods assuming real conditions for obtaining information. The main goal set by the authors is to verify selected algorithms, identify the key aspects determining the effectiveness of their operation and indicate potential directions of their development.

## **2. State of the art**

The operation of classic object identification algorithms, indicating the similarities between the recorded and reference images (the so-called *patterns*), is mainly based on the use of correlation methods. These algorithms, although effectively implemented in solutions to typical technical problems, are insufficiently effective in the case of topographic navigation. It is related to, inter alia, the limitations and conditions in the measurement system, environmental conditions and characteristics of the detected objects, which have a strong negative impact on the obtained correlation results. This disqualifies the possibility of their direct use in the tasks of matching reference terrain maps with the acquired image information.

A particularly significant obstacle is the fact that the sensory elements of navigation systems installed on UAV/CM record image data in various environmental and lighting conditions [10]. Frequently, reference data of high informative value, due to various conditions, constitute a pattern of little use or even lead to incorrect results. This is the case, for example, when the reconnaissance is



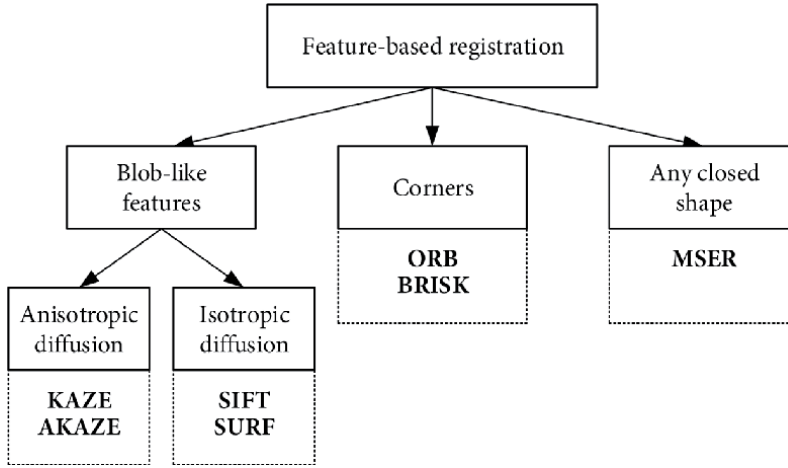


**Figure 1.**  
*Images of the same fragment of the Earth's surface taken under different weather and lighting conditions.*

conducted in different weather conditions than those in which the UAV/CM mission takes place (**Figure 1**). Therefore, image feature matching becomes a complex issue. The conditions related to the image recording parameters, e.g. variable view angle, maintaining scale or using various types of sensors, turn out to be equally important.

Image matching methods began to be strongly developed with the dissemination of digital image in technology. Initially, the classical Fourier and correlation methods were used. However, these methods did not allow for successful multi-modal, multi-temporal, and multi-perspective matching of different images. The taxonomy of the classical methods used in the image matching process was presented in the early 1990s [11]. The image feature space, considered as a source of information necessary for image matching, was defined and local variations in the image were identified as the greatest difficulty in the matching process. In the 21st century, further development of methods based on the features of the image continued [12]. It should be emphasised that most image matching methods based on image features include four consecutive basic steps: feature detection, feature matching, model estimation of mutual image transformation and final transformation of the analysed image. These methods became an alternative to the correlation and Fourier methods. For over a dozen years, new, effective algorithms for processing and matching digital images have been developed, using statistical methods based on matching local features in images [11–14], cf. **Figure 2**. Their authors point to the greater invariance of the proposed algorithms to perspective distortions, rotation, translation, scaling and lighting changes. Given their high reliability under static conditions, as well as their low sensitivity to changes in the optical system's position, including translation, orientation and scale, it is justified to conduct studies in order to verify their usefulness and effectiveness. The paper focuses on modern image matching algorithms, which can potentially be used in topographic navigation issues. It should be stressed that the problem is completely different in the indicated context. This is due to the fact that although the matched images represent the same area of the terrain, the manner and time of the recording differ significantly from each other. This is not a typical application of these algorithms, hence a limited effectiveness of their operation can be expected.

The common feature of all methods is the use of the so-called *scale space* described in [15], allowing the decimation of image data and examination of



**Figure 2.** Classification of the selected methods of image feature matching.

similarities between images of different scales. A significant step in the development of image matching methods based on local features was the development of the Scale-Invariant Feature Transform (SIFT) algorithm [16]. In this algorithm, the characteristics are selected locally and their position does not change while the image is scaled. Their indication is done by determining the local extremes of function  $D(\hat{x})$ , that is the difference between the results of image  $I(x, y)$  convolution with Gaussian functions  $G(x, y, \sigma)$  with different values of scale parameter  $\sigma$ :

$$D(\hat{x}) = D + \frac{1}{2} \frac{\partial D^T}{\partial x} \hat{x} \quad (1)$$

where

$$\begin{aligned} D &= D(x, y, \sigma) = [G_{\sigma_1}(x, y) - G_{\sigma_2}(x, y)] \star I(x, y) = \\ &= \frac{1}{\sqrt{2\pi}} \left[ \frac{1}{\sigma_1} e\left(-\frac{x^2 y^2}{2\sigma^2}\right) - \frac{1}{\sigma_2} e\left(-\frac{x^2 y^2}{2\sigma^2}\right) \right] \star I(x, y) \end{aligned} \quad (2)$$

A more numerically efficient version of the SIFT algorithm, called Speeded-Up Robust Features (SURF) is based on the so-called *integral images* [17]. Both methods use the basic processing steps described in [12]. Additionally, in order to ensure the effectiveness of feature detection in images of different resolutions, a scale space, consisting of octaves which represent the series of responses of a convolutional filter with a variable size, was introduced.

Simply put, the detection of the characteristic point is based on the use of the determinant of a Hessian matrix  $\det(H)$ . In the case of SURF, the second-order derivatives of the Gaussian function  $G$  approximated by the box filters  $B_{xx}$ ,  $B_{xy}$ ,  $B_{yy}$ , and the integral image are also used [18]. The Hessian matrix in these methods takes the form

$$H = \begin{bmatrix} L_{xx}(x, y, \sigma) & L_{xy}(x, y, \sigma) \\ L_{xy}(x, y, \sigma) & L_{yy}(x, y, \sigma) \end{bmatrix} \quad (3)$$

where

$$\begin{aligned}
 L_{xx}(x, y, \sigma) &= \frac{\partial^2}{\partial x^2} G(x, y, \sigma) \star I(x, y) \approx B_{xx} \star I(x, y) \\
 L_{xy}(x, y, \sigma) &= \frac{\partial^2}{\partial x \partial y} G(x, y, \sigma) \star I(x, y) \approx B_{xy} \star I(x, y) \\
 L_{yy}(x, y, \sigma) &= \frac{\partial^2}{\partial y^2} G(x, y, \sigma) \star I(x, y) \approx B_{yy} \star I(x, y)
 \end{aligned} \tag{4}$$

The determinant of the Hessian matrix after approximation using box filters and the Frobenius norm is given as

$$\det(H) \approx B_{xx}B_{yy} - \left(\frac{9}{10}B_{xy}\right)^2 \tag{5}$$

After detecting the local extremes of  $\det(H)$ , similarly to  $D(\hat{x})$ , the location of characteristic points representing local features, called *blob-like* for the SIFT and SURF methods, is determined. In this step, the SIFT method also rejects features whose contrast is lower than the assumed threshold  $t$  by comparing  $|D(\hat{x})| < t$  as well as points lying on isolated edges. This is done by comparing the value of the Hessian matrix  $H$  trace quotient, and its determinant with the curvature coefficient  $r$ :

$$\frac{\text{tr}(H)}{\det(H)} < \frac{(r+1)^2}{r} \tag{6}$$

In 2011, an alternative method to SIFT and SURF, Oriented FAST and Rotated BRIEF (ORB), was proposed [19]. The method was based on the modified Features from Accelerated Segment Test (FAST) detector [20, 21], enabling corner and edge detection, and a modified Binary Robust Independent Elementary Features (BRIEF) descriptor [22]. This approach involves changing the scale of the image on the basis of blurring with an increasing value of Gaussian filter. Despite the noise reduction and enhancing the uniformity of areas interpreted by human beings as unique (e.g. surface of the lake, wall of a building, shape of a vehicle, etc.), it causes blurring of their edges. This often leads to the inability to indicate the boundaries between areas and to define characteristic points in their neighbourhood.

The solution to this problem was proposed in the KAZE method (Japanese for “wind”) [23]. Unlike the SIFT and SURF methods, which use the Gaussian function causing isotropic diffusion of luminance to generalise the image, in the KAZE method the generalisation is based on nonlinear diffusion in consecutive octaves of the scale [24]. The anisotropic image blurring in this method depends on the local luminance distribution. Nonlinear diffusion can be presented in the following equation:

$$\frac{\partial I}{\partial t} = \text{div}[c(x, y, t)\nabla I] \tag{7}$$

The blur intensity can be adapted by the introduced conductivity function  $c$ , which is usually related to time. However, using the approach proposed in [15], parameter  $t$  is related to the image scale. Various forms of the conductivity function  $c$  were proposed in related works developing the use of nonlinear diffusion in the context of image filtration [24–26]. One of the functions used for nonlinear diffusion can be:

$$c = \exp\left(-\frac{|\nabla L_\sigma|^2}{k^2}\right) \quad (8)$$

where  $\nabla L_\sigma$  is the gradient of the Gaussian blurred function of the original image  $I$  on the scale  $\sigma$ , and  $k$  is the contrast ratio.

This function allows for blurring the image while maintaining the edges of structures. As a result, more features can be detected at different image scales. However, it involves the use of a gradient, which in the case of intense image disturbance, e.g. in the form of a shadow, may cause an unfavourable (due to the subsequent detection of features) distribution of diffusion in the image.

An important stage of the considered methods is the description of a characteristic point by means of a vector containing information about its surroundings. The SIFT method uses a luminance gradient and in the SURF method the image response to horizontally and vertically oriented Haar wavelet is applied. In general, around the characteristic point in the area with a defined radius dependent on the  $\sigma$  scale, a certain number of cells are created and dominant values of the gradient or the responses to Haar wavelets are determined. These are the basis for calculating the so-called *feature metrics*. Finally, the dominant orientation is established. Characteristic features in the SIFT method are determined by

$$\begin{aligned} m(x, y) &= \sqrt{[L_\sigma(x+1, y) - L_\sigma(x-1, y)]^2 + [L_\sigma(x, y+1) - L_\sigma(x, y-1)]^2} \\ \theta(x, y) &= \tan^{-1} \left[ \frac{L_\sigma(x, y+1) - L_\sigma(x, y-1)}{L_\sigma(x+1, y) - L_\sigma(x-1, y)} \right] \end{aligned} \quad (9)$$

where  $m(x, y)$  is the gradient size,  $\theta(x, y)$  is the orientation, and  $L_\sigma$  is the blurred image discussed above.

The SIFT method creates thereunder a gradient histogram that sums up the determined values in four cells. In analogous cells, according to the SURF method, the responses to Haar wavelets distributed along the radii in the neighbourhood of the point with an interval of  $\pi/3$  are summed up. In each SURF subregion, a vector  $\mathbf{v}$  is determined:

$$\mathbf{v} = \left[ \sum d_x \quad \sum d_y \quad \sum |d_x| \quad \sum |d_y| \right] \quad (10)$$

where  $d_x$  and  $d_y$  are the characteristic point's neighbourhood responses to the horizontally and vertically oriented Haar wavelets, respectively.

In the KAZE method the procedure is similar as for the SURF method with the difference that the first order derivatives from the image function are used. The point description operation is performed for all levels in the adopted scale space, thereby creating a pyramid of vectors assigned to subsequent levels containing an increasingly generalised image.

The Maximally Stable Extremal Regions (MSER) method introduced in [27] has a different approach to the detection and description of local features. In this method, regions (shapes), referred to as *maximally stable*, are selected as the characteristics of the image. The image in this method is treated as a function  $I$  which transforms

$$I : \mathcal{D} \subset \mathbb{Z}^2 \rightarrow \mathbf{S} \quad (11)$$

where  $\mathcal{D}$  is the domain of  $I$  and  $\mathbf{S}$  is its set of values, usually  $\mathbf{S} = \{0, 1, \dots, 255\}$ .

Regions (areas, shapes) with a specific (typically average) luminance level can be determined in the image. Region  $Q$  is understood as a subset of pixels in the

image which is a continuous subset of  $\mathcal{D}$  such that  $\forall p, q \in Q$  exist sequences  $p, a_1, a_2, \dots, q$  and  $pAa_1, \dots, a_iAa_{i+1}, \dots, a_nAq$ , where  $A \subset \mathcal{D} \times \mathcal{D}$  is the neighbourhood relation and the formula  $a_iAa_{i+1}$  is the neighbourhood between pixels  $a_i$  and  $a_{i+1}$ . The extremal region is  $Q \subset \mathcal{D}$ , such that  $\forall p \in Q, \forall q \in \partial Q : I(p) > I(q)$  (maximum intensity region) or  $I(p) < I(q)$  (minimum intensity region). The desired maximally stable extremal region (MSER) is the region  $R = Q_{i^*}$ , which for the sequence of extremal regions  $Q_1, \dots, Q_{i-1}, Q_i$ , nested i.e.  $Q_i \subset Q_{i+1}$  and for  $q(i) = |Q_{i+\Delta} \setminus Q_{i-\Delta}| / |Q_i|$  has a local minimum at  $i^*$ . Whereas  $\Delta \in \mathbb{S}$  is the stability parameter and the luminance threshold. The procedure of determining MSER regions is repeated throughout the assumed  $\sigma$  scale space.

In the feature description stage, a vector using image moments is determined for each region. Based on the moments  $m_{00}, m_{01}, \dots, m_{20}$ , the centre of gravity of each MSER region and the ellipse approximating the region are determined according to the procedure described in [28]. The ellipse equation is given as

$$\frac{\left[ x - x_g + \theta(y - y_g) \right]^2}{a_1(1 + \theta^2)} + \frac{\left[ y - y_g + \theta(x - x_g) \right]^2}{a_2(1 + \theta^2)} - 1 = 0 \quad (12)$$

The orientation  $\theta$  and the size of the ellipse defined by its  $a_1$  and  $a_2$  axes allow to describe the features of the region taken for comparison in the matching step. The moment  $m$  of the order  $(p + q)$  of the MSER region used to determine the centre of gravity of the  $C = I(x_g, y_g)$  region can be represented as follows:

$$m_{pq} = \sum_{\{x,y\} \in \mathbb{R}} x^p y^q \quad (13)$$

The use of moments and the centre of gravity is also a feature of ORB method which uses machine learning approach for corner detection. After their detection, based on the image moments, the centre of gravity  $C$  is determined for each corner according to the formula

$$C = \left( \frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right) \quad (14)$$

where

$$m_{pq} = \sum x^p y^q I(x, y) \quad (15)$$

On the basis of the corner's position and centre of gravity, the orientation of the feature is determined as shown in the equation:

$$\theta(x, y) = \text{atan2}(m_{01}, m_{10}) \quad (16)$$

The feature description step uses the assigned orientation to complete the binary BRIEF descriptor [22], with the condition of verifying the belonging of the point  $L_\sigma(x, y)$  to the matrix  $\mathbf{W}_\theta$ . It is based on the simple comparison of the pixel luminance values in the neighbourhood of the feature:

$$\tau(L_\sigma; x, y) = \begin{cases} 1 & \Leftrightarrow L_\sigma(x, y) < L_\sigma(x_1, y_1) \\ 0 & \Leftrightarrow L_\sigma(x, y) \geq L_\sigma(x_1, y_1) \end{cases} \quad (17)$$

The matrix  $\mathbf{W}_\theta$  is the product of the original matrix  $\mathbf{W}$ , containing the locations of the points, which are subject to the tests, and the rotation matrix based on the determined angles  $\theta(x, y)$ . In such case the ORB vector describing the feature takes the form:

$$\mathbf{v}_n = \sum_{1 \leq i \leq n} 2^{i-1} \tau(L_\sigma; x_i, y_i) | (x_i, y_i) \in \mathbf{W}_\theta \quad (18)$$

The common element for the described methods is the stage of comparing the distinguished features detected on the reference and registered images. It is of fundamental importance in the field of absolute terrain position designation, because the location of the matched features is the source of determining the matrix of mutual image transformation. In this comparison, vectors describing the features in a given method, e.g. feature metric and its orientation, are taken into account.

The determination of the similarity between the feature description vectors  $\mathbf{v}_a$  and  $\mathbf{v}_b$  is based on various measures. The most commonly used are the distances defined as follows:

$$d_1(\mathbf{v}_a, \mathbf{v}_b) = \sum |\mathbf{v}_a - \mathbf{v}_b| \quad \text{and} \quad d_2(\mathbf{v}_a, \mathbf{v}_b) = \sum (\mathbf{v}_a - \mathbf{v}_b)^2 \quad (19)$$

The third frequently used norm for binary vectors is the Hamming distance given as:

$$d_3(\mathbf{v}_a, \mathbf{v}_b) = \sum \text{XOR}(\mathbf{v}_a, \mathbf{v}_b) \quad (20)$$

Another approach for matching two features is the nearest neighbour algorithm based on the ratio of the distances  $d_1$  and  $d_2$ . However, it should be remembered that the matching result for the described distances may vary, hence the importance of the features' detection and description step.

The final step in all the discussed methods is the statistical verification of a set of matched local features. It happens that, as a result of the initial comparison of the vectors which describe the features, mismatches resulting from the acquisition conditions described above are indicated. Therefore, after the pre-processing step, additional criteria are applied to distinguish matches from mismatches, e.g. based on the Random Sample Consensus (RANSAC) method [29]. This method allows for the estimation of a mathematical model describing the location of local features in the image provided that most of the matched points fit into this model (with the assumed maximum error). Then those points that do not fit into the estimated model are discarded in the step of determining the image transformation matrix.

### 3. Problem formulation

The following set is considered:

$$\mathbf{I} = \{I_i, i \in \mathbb{N}\} \quad (21)$$

Elements of  $\mathbf{I}$  are two-dimensional discrete signals (digital images) and describe the same part of the Earth's surface, but recorded at different times and therefore under different environmental and lighting conditions. Image  $I_j$  chosen from the set  $\mathbf{I}$  is treated as a reference signal, i.e. characterised by excellent structural similarity to itself. In order to compare any  $I_k$  image selected from the set  $\mathbf{I}$  with reference

image  $I_j$ , the following similarity measures were used: mean square error  $J_{\text{MSE}}$  and the related  $J_{\text{PSNR}}$  (peak-signal-to-noise ratio) and  $J_{\text{SSIM}}$  (structural similarity index measure).

The mean square error is determined by the formula:

$$J_{\text{MSE}}(I_j, I_k) = \frac{1}{MN} \sum_{x=1}^M \sum_{y=1}^N [I_j(x, y) - I_k(x, y)]^2 \quad (22)$$

in which  $M$  and  $N$  are the width and height of images in pixels. Index  $J_{\text{PSNR}}$  is defined as:

$$J_{\text{PSNR}}(I_j, I_k) = 10 \log_{10} \left\{ \frac{\ell^2}{J_{\text{MSE}}(I_j, I_k)} \right\} \quad (23)$$

where  $\ell$  is the range of changes of the luminance value, while the index  $J_{\text{SSIM}}$  can be described as:

$$J_{\text{SSIM}}(I_j, I_k) = \frac{(2\mu_{I_j}\mu_{I_k} + \xi_1)(2\sigma_{I_j I_k} + \xi_2)}{(\mu_{I_j}^2 + \mu_{I_k}^2 + \xi_1)(\sigma_{I_j}^2 + \sigma_{I_k}^2 + \xi_2)} \quad (24)$$

in which  $\mu_{I_j}$  and  $\mu_{I_k}$  are the mean values of the  $I_j$  and  $I_k$  image luminance,  $\sigma_{I_j}$  and  $\sigma_{I_k}$  are the variances of  $I_j$  and  $I_k$ ,  $\sigma_{I_j I_k}$  is covariance of a pair  $(I_j, I_k)$ , and  $\xi_1 = (0.01\ell)^2$  and  $\xi_2 = (0.03\ell)^2$  are positive constants avoiding instability when the denominator of the Eq. (24) is very close to zero [30–32].

For such defined initial conditions, the best match of the subsequent elements of the set  $\mathbf{I}$  in relation to the reference element  $I_j$  is sought, assuming that the similarity index measures of the examined pairs  $(I_j, I_k)$  are strongly undesirable, i.e.

$$J_{\text{PSNR}}(I_j, I_k) \rightarrow 0 \quad \text{and} \quad J_{\text{SSIM}}(I_j, I_k) \ll 1 \quad (25)$$

The term *best match* is understood as defining certain vectors  $\mathbf{v}_j$  and  $\mathbf{v}_k$  of values, which characterise the considered signals  $I_j$  and  $I_k$ , and then linking them in a way that makes it possible to explicitly state that the selected pair  $(I_j, I_k)$  describes the same fragment of the Earth's surface.

#### 4. Performance analysis

In order to verify the sensitivity of the selected methods to limitations in the measurement system and environmental changes, a number of studies taking into account the actual conditions of obtaining information were conducted. Due to their difficult nature, they were performed with the use of computer simulation methods. The research was carried out in three stages. In the first stage, a detailed analysis of the test sets, using the values of the similarity indexes of the elements defined in the article, was completed. On the basis of the performed tests, special cases were selected and subjected to detailed analysis. In the further part of the study, the methods and verification of the correctness of image data matching in the scope of mutual matching of the sets presented to the analysis were compared. Finally, the influence of changes in the contrast of the acquired image on the number of features detected and the subsequent matching results was examined.

#### 4.1 Analysis of test set elements

For the initial numerical tests, the test set **I** consisting of four elements was adopted, whereby  $I_0$  is treated as reference. Each element of the set **I** is a 24-bit digital image with a size of  $1080 \times 1080$  pixels, representing the same fragment of the Earth's surface, with a centrally located characteristic terrain object (**Figure 3**).

The object is located in a natural environment characteristic for tundra and therefore distinguished by a rocky ground with a very low plant cover, dominated by mosses and lichens. Image  $I_0$  (reference) was taken in the autumn and mostly brown colours, associated with the tundra soils and rock formations in this area, prevail in it. The image  $I_1$  shows the environment in spring–summer conditions, i.e. during the growing season. Images  $I_2$  and  $I_3$  were taken in winter, with snow cover, whereby in the case of  $I_3$ , there is also a strong cloud cover. Test set **I** elements' similarity index measures, determined on the basis of the Eqs. (22)–(24) with the reference image  $I_0$ , are presented in **Table 1** (columns 2–5).

Based on the obtained results, it can be shown that the elements constituting the test set **I** were selected so that only one of them ( $I_1$ ) has a relatively high degree of

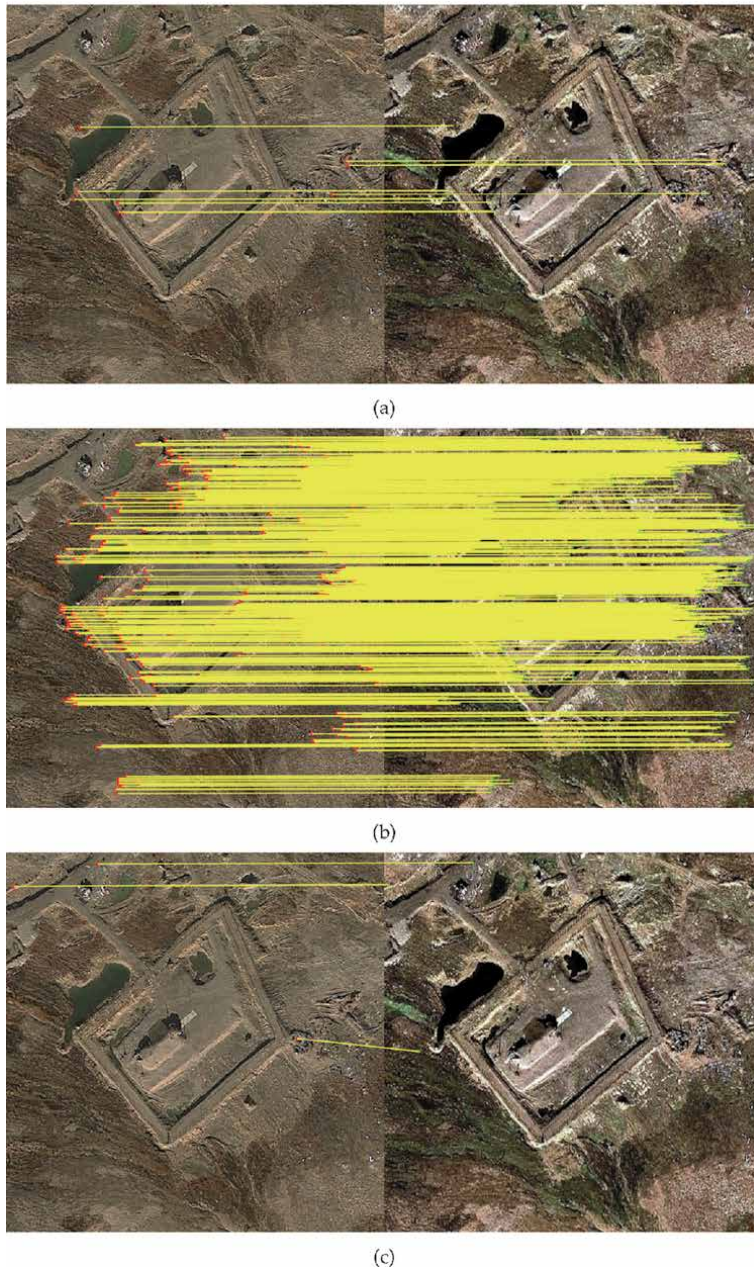


**Figure 3.** Test image set **I**: (a) reference image  $I_0$ , (b)–(d) test images  $I_1, I_2, I_3$  (source: Google Earth).



	$(I_0, I_0)$	$(I_0, I_1)$	$(I_0, I_2)$	$(I_0, I_3)$	$(I_2, I_2)$	$(I_2, I_3)$
$J_{MSE}$	0	2.23E03	1.46E04	1.19E04	0	2.99E03
$J_{PSNR}$	$\infty$	14.65	6.47	7.36	$\infty$	13.37
$J_{SSIM}$	1	0.4373	0.0742	0.0755	1	0.3378

**Table 1.**  
 Similarity index measures for selected pairs of the set I.



**Figure 4.**  
 Image pair  $(I_0, I_1)$  matching result for: (a) SURF, (b) KAZE, and (c) MSER method.



**Figure 5.** Image pair  $(I_0, I_1)$  matching result for ORB method.

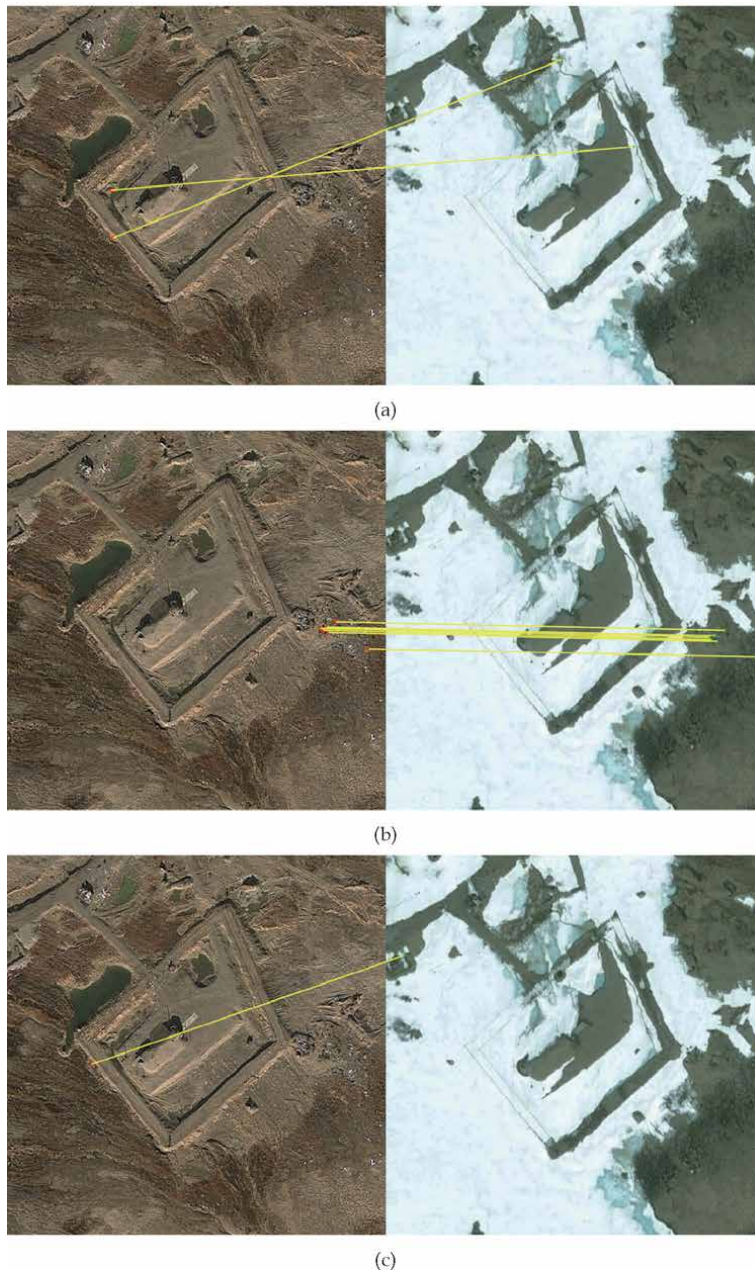
similarity to the reference image. The remaining items ( $I_2$  and  $I_3$ ) have unfavourable similarity index measures  $J_{\text{PSNR}}$  and  $J_{\text{SSIM}}$ , which enables the assumptions of the Eq. (25) to be met. Although in a subjective manner the image  $I_2$  is more similar to the reference image  $I_0$ ,  $I_3$  is characterised by more favourable  $J_{\text{PSNR}}$  and  $J_{\text{SSIM}}$  index measures.

It should be noted that  $I_2$  and  $I_3$  were also carefully selected. Both images show a similar arrangement of snow cover, which is reflected in the determined values of the pair similarity indexes  $(I_2, I_3)$ , cf. **Table 1**, columns 6 and 7. This is justified in practice: it may happen that the data from the reconnaissance are accurate (e.g. they take into account the snow in the area concerned, the lack of leaves on the trees in late autumn or the high water level in spring), but strong fogging or rainfall make the image  $I_3$  obtained from UAV/CM recording systems several hours later significantly different from the reference image (in this case  $I_2$ ).

#### 4.2 Comparison of the selected methods of image feature matching

The test set **I** presented in SubSection 4.1 was examined in order to compare the effectiveness of image matching performed by algorithms using local features. The SURF, MSER, ORB and KAZE methods were taken into account. Image  $I_0$  is a pattern, and  $I_1, I_2, I_3$  are matched images. In the algorithms, the values of parameters proposed by their authors were used with the exception of the features' similarity threshold, which was lowered to the level of 50% due to large differences between individual elements of the set **I**. The best matching of individual features in the compared images was assumed, using the similarity measures proposed for these methods. The RANSAC method was used for the final correction of the matched features, for which an affine transformation model between images was adopted. In order to verify the effectiveness of the considered methods and the correctness of the parameters adopted in the last study, the pattern was replaced. It was assumed that  $I_2$  is a reference image and  $I_3$  is a matched image. The matching results of the individual test pairs of **I** are shown in **Figures 4–8** and in **Tables 2–5**.

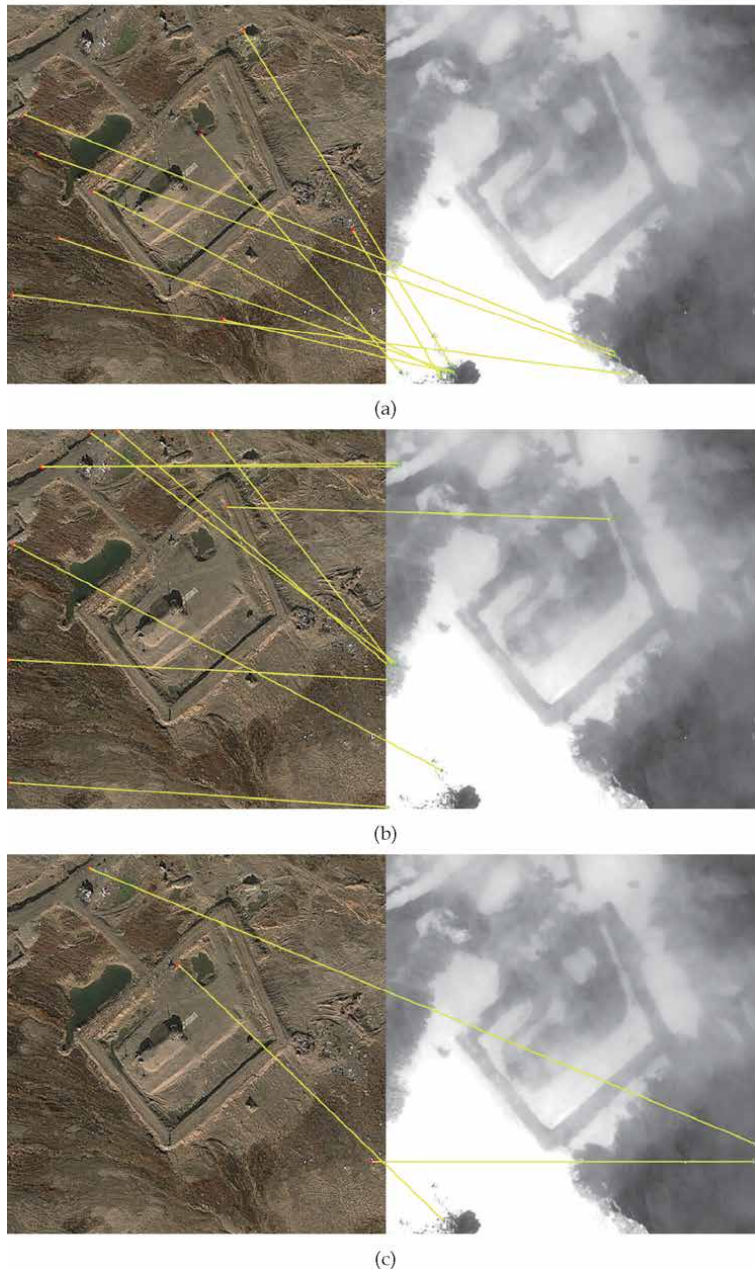
Analysis of the matching results has shown that the selected algorithms are not effective when the matched images, despite the same content, differ significantly, cf. pair  $(I_0, I_3)$ . All methods were most effective in matching the pair  $(I_0, I_1)$ . While the SURF and MSER methods indicated mismatches for matching pairs  $(I_0, I_2)$  and  $(I_0, I_3)$ , the ORB method did not (cf. **Table 3** and **Table 4**). The KAZE method identified correctly the fragment of the image on which the corresponding features



**Figure 6.**  
*Image pair  $(I_0, I_2)$  matching result for: (a) SURF, (b) KAZE, and (c) MSER method.*

of the pair  $(I_0, I_2)$  were located. When comparing a relatively similar pair  $(I_2, I_3)$ , it appeared that all algorithms indicated mismatches or lack thereof, with KAZE and MSER indicating two correct matches each (**Table 5**).

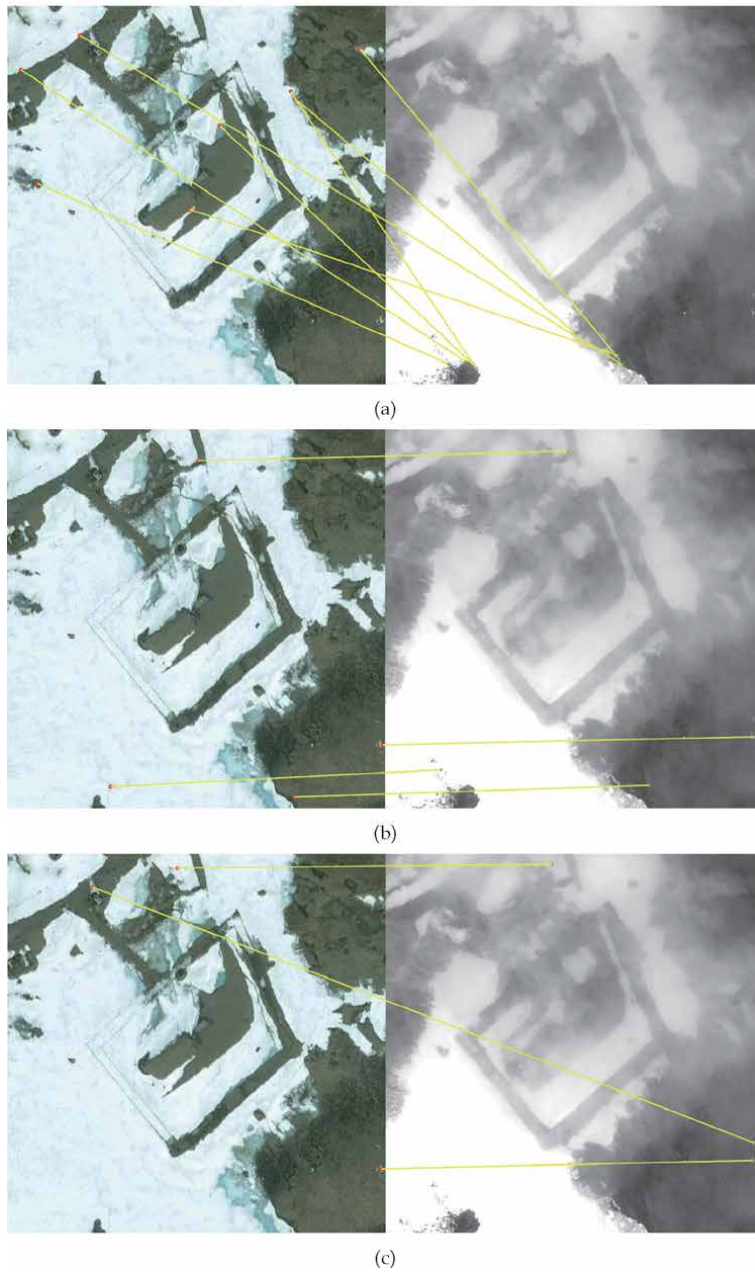
In general, the KAZE method proved to be the most effective, while the ORB method showed the least processing efficiency of the set **I**. Due to the lack of any pair  $(I_0, I_2)$ ,  $(I_0, I_3)$  and  $(I_2, I_3)$  matches, no graphical results are presented for the ORB method. A potential cause of a lack of pair  $(I_2, I_3)$  matches is a large contrast change, characteristic of the occurrence of acquisition interference, such as the fog visible in the image  $I_3$ .



**Figure 7.** Image pair  $(I_0, I_3)$  matching result for: (a) SURF, (b) KAZE, and (c) MSER method.

### 4.3 Effect of contrast change on the number of the detected features

The research focused on the analysis of the effect of contrast change on the number of features detected in the image. For this purpose, the contrast of the image  $I_0$  had been gradually reduced until a uniform colour throughout the image was obtained. Afterwards, the transformed set was further analysed. SURF, MSER, ORB and KAZE methods were used again. **Figure 9** shows the cumulative results of this study.



**Figure 8.**  
 Image pair  $(I_2, I_3)$  matching result for: (a) SURF, (b) KAZE, and (c) MSER method.

	SURF	KAZE	MSER	ORB
Correct matches	9	489	2	11
Mismatches	0	0	1	0
Percentage of correct matches	100%	100%	67%	100%

**Table 2.**  
 Image pair  $(I_0, I_1)$  matching results.

	SURF	KAZE	MSER	ORB
Correct matches	0	6	0	0
Mismatches	2	0	1	0
Percentage of correct matches	0%	100%	0%	0%

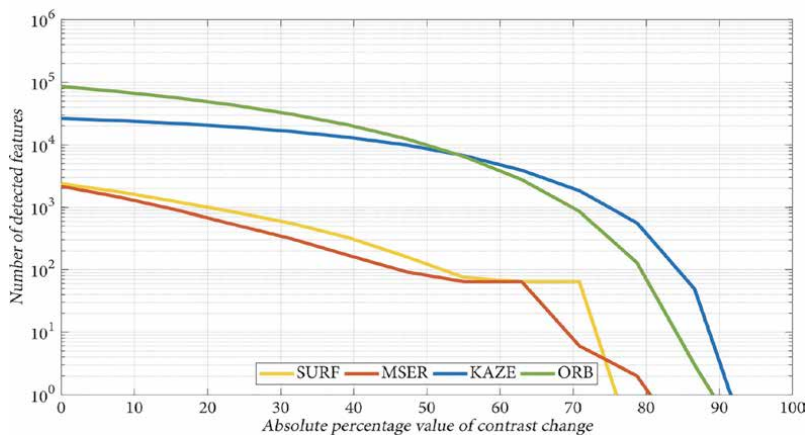
**Table 3.**  
Image pair  $(I_0, I_2)$  matching results.

	SURF	KAZE	MSER	ORB
Correct matches	0	2	1	0
Mismatches	9	7	2	0
Percentage of correct matches	0%	29%	33%	0%

**Table 4.**  
Image pair  $(I_0, I_3)$  matching results.

	SURF	KAZE	MSER	ORB
Correct matches	0	3	2	0
Mismatches	8	1	1	0
Percentage of correct matches	0%	75%	67%	0%

**Table 5.**  
Image pair  $(I_2, I_3)$  matching results.



**Figure 9.**  
Effect of contrast change on the number of the detected image features.

On the basis of the results obtained, it can be concluded that the number of features detected by the examined methods decreases with image contrast reducing, which results in a smaller statistical sample processed in each subsequent step of these methods. This may be the cause of the lower matching efficiency of the methods considered for images that are significantly different from each other.

## 5. Conclusions and final remarks

The results of the algorithms presented in the literature are usually related to images that are fragments of source images, i.e. have their qualitatively identical counterparts in Ref. images. In the analysed cases, the differences between the reference image stored in the memory of the navigation system and that recorded by the sensor are significant. As a result, there are certain consequences that often prevent the image representing the same field object from being effectively matched. This is due to real environmental conditions and restrictions on obtaining information. The measurement system parameters and the quality of the images taken have a direct impact on the number of detected features. For example, the lack of complete information about the accuracy of field object's image mapping makes it impossible to properly select the size of the filters. This results in the detection of objects that are completely irrelevant to the issue considered, such as bushes, leaves or grass blades, which are highly variable over time. Consequently, it has a significant impact on the performance of individual algorithms.

The study concluded that the use of statistical algorithms such as RANSAC improves the effectiveness of the selected methods. However, the results obtained strongly depend on the size of the set taken into consideration and the match/mismatch ratio. Therefore, in the terrain image processing, it is necessary to conduct an analysis of the informational characteristics of the examined objects and the conditions of acquisition. This allows for extracting characteristic points whose description does not significantly change due to atmospheric conditions.

The results of the simulation tests enable a general conclusion that the methods considered are often insufficient to determine the coordinates of a UAV/CM flying under unfavourable environmental conditions. The greatest development potential, in the context of the implementations examined in this work, is characterised by methods based on anisotropic diffusion, which in the course of simulation studies showed the highest effectiveness. Therefore, it seems justified to focus the research effort on further development of new image processing methods within the group of anisotropic diffusion methods. In particular, it is proposed to take the informative character of terrain images as determinants of the input parameters of the designed processing methods into account, to apply pre-processing methods aimed at decimation of the input data, their segmentation and determination of the main components, and to extend the definition of the designed methods with additional criteria increasing the effectiveness of detection and image feature matching. The newly developed methods should be aimed at the improvement of feature detection efficiency in terrain images and the selection of processing parameters taking into account environmental conditions as well as limitations and conditions in the measurement system.

## Acknowledgements

This work is financed by the National Centre of Research and Development of the Republic of Poland as part of the scientific research program for the defence and security named *Future Technologies for Defence – Young Scientist Contest* (Grant No. DOB-2P/03/06/2018).

## Conflict of interest

The authors declare no conflict of interest.


### **Author details**

Piotr Turek, Stanisław Grzywiński and Witold Bużantowicz\*  
Military University of Technology, Warsaw, Poland

\*Address all correspondence to: [witold.buzantowicz@wat.edu.pl](mailto:witold.buzantowicz@wat.edu.pl)

### **IntechOpen**

---

© 2020 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 



## References

- [1] Boozer DD, Fellerhoff JR. Terrain-Aided Navigation Test Results in the AFTI/F-16 Aircraft. *Navigation – Journal of The Institute of Navigation*. 1988;35(2):161–175. DOI: 10.1002/j.2161-4296.1988.tb00949.x
- [2] Enns R. Terrain-aided navigation using the Viterbi algorithm. *Journal of Guidance, Control, and Dynamics*. 1995; 18(6):1444–1449. DOI: 10.2514/3.21566
- [3] Han Y, Wang B, Deng Z, Fu M. An improved TERCOM-based algorithm for gravity-aided navigation. *IEEE Sensors Journal*. 2016;16(8):2537–2544. DOI: 10.1109/JSEN.2016.2518686
- [4] Hua Z, Xiulin H. A height-measuring algorithm applied to TERCOM radar altimeter. In: *Proc. of the 3rd International Conference on Advanced Computer Theory and Engineering (ICACTE)*; 20–22 August 2010; Chengdu (China). New York: IEEE, 2010. p. (V5–43)-(V5–46). DOI: 10.1109/ICACTE.2010.5579215
- [5] Wei E, Dong C, Liu J, Tang S. An improved TERCOM algorithm for gravity-aided inertial navigation system. *Journal of Geomatics*. 2017;42(6):29–31. DOI: 10.14188/j.2095-6045.2016190
- [6] Naimark L, Webb H, Wang T. Vision-Aided Navigation for Aerial Platforms. In: *Proc. of the ION 2017 Pacific PNT Meeting*; 1–4 Mai 2017; Honolulu (USA). Manassas: ION, 2017. p. 70–76. DOI: 10.33012/2017.15051
- [7] Yang C, Vadlamani A, Soloviev A, Veth M, Taylor C. Feature matching error analysis and modeling for consistent estimation in vision-aided navigation. *Navigation*. 2018;65:609–628. DOI: 10.1002/navi.265628
- [8] Carr JR, Sobek JS. Digital Scene Matching Area Correlator (DSMAC). In: *Proc. of 24th Annual Technical Symposium, SPIE 0238, Image Processing for Missile Guidance*; 23 December 1980; San Diego (USA). Bellingham: SPIE, 1980. DOI: 10.1117/12.959130
- [9] Irani GB, Christ JP. Image processing for Tomahawk scene matching. *Johns Hopkins APL Technical Digest*. 1994;15 (3):250–264.
- [10] Turek P, Bużantowicz W. Image matching constraints in unmanned aerial vehicle terrain-aided navigation. In: *Proc. of the 2nd Aviation and Space Congress*; 18–20 September 2019; Cedzyna (Poland). p. 206–208.
- [11] Brown LG. A survey of image registration techniques. *ACM Computing Surveys*. 1992;24(4):325–376. DOI: 10.1145/146370.146374
- [12] Zitová B, Flusser J. Image registration methods: A survey. *Image and Vision Computing*. 2003;21(11): 977–1000. DOI: 10.1016/S0262-8856 (03)00137-9
- [13] Bouchiha R, Besbes K. Automatic Remote-Sensing Image Registration Using SURF. *International Journal of Computer Theory and Engineering*. 2013;5(1):88–92. DOI: 10.7763/IJCTE.2013.V5.653
- [14] Kashif M, Deserno TM, Haak D, Jonas S. Feature description with SIFT, SURF, BRIEF, BRISK, or FREAK? A general question answered for bone age assessment. *Computers in Biology and Medicine*. 2016;68:67–75. DOI: 10.1016/j.compbiomed.2015.11.006
- [15] Lindeberg T. Scale-space theory: A basic tool for analysing structures at different scales. *Journal of Applied Statistics*. 1994;21(2):224–270. DOI: 10.1080/757582976
- [16] Löwe DG. Distinctive Image Features from Scale-Invariant

- Keypoints. *International Journal of Computer Vision*. 2004;60:91–110. DOI: 10.1023/B:VISI.0000029664.99615.94
- [17] Bay H, Ess A, Tuytelaars T, Van Gool L. Speeded-Up Robust Features (SURF). *Computer Vision and Image Understanding*. 2008;110(3):346–359. DOI: 10.1016/j.cviu.2007.09.014
- [18] Viola P, Jones M. Rapid object detection using a boosted cascade of simple features. In: *Proc. of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*; 8–14 December 2001; Kauai (USA). p. (I-511)-(I-518). DOI: 10.1109/CVPR.2001.990517
- [19] Rublee E, Rabaud V, Konolige K, Bradski G. ORB: An efficient alternative to SIFT or SURF. In: *Proc. of the 13th International Conference on Computer Vision*; 6–13 November 2011; Barcelona (Spain). p. 2564–2571. DOI: 10.1109/ICCV.2011.6126544
- [20] Rosten E, Drummond T. Machine Learning for High-Speed Corner Detection. In: Leonardis A, Bischof H, Pinz A, editors. *Proc. of the 9th European Conference on Computer Vision 2006 – Lecture Notes in Computer Science*, vol. 3951. Berlin-Heidelberg: Springer; 2006. p. 430–443. DOI: 10.1007/11744023\_34
- [21] McIlroy P, Rosten E, Taylor S, Drummond T. Deterministic sample consensus with multiple match hypotheses. In: *Proc. of the 21st British Machine Vision Conference*; 31 August – 3 September 2010; Aberystwyth (UK). pp. 111.1–111.11. DOI: 10.5244/C.24.111
- [22] Calonder M, Lepetit V, Strecha C, Fua P. BRIEF: Binary Robust Independent Elementary Features. In: Daniilidis K, Maragos P, Paragios N, editors. *Proc. of the 11th European Conference on Computer Vision 2010 – Lecture Notes in Computer Science*, vol. 6314. Berlin-Heidelberg: Springer; 2010. p. 778–792. DOI: 10.1007/978-3-642-15561-1\_56
- [23] Alcantarilla PF, Bartoli A, Davison AJ. KAZE Features. In: Fitzgibbon A, Lazebnik S, Perona P, Sato Y, Schmid C, editors. *Proc. of the 13th European Conference on Computer Vision 2010 – Lecture Notes in Computer Science*, vol. 7577. Berlin-Heidelberg: Springer; 2012. p. 214–227. DOI: 10.1007/978-3-642-33783-3\_16
- [24] Perona P, Malik J. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1990;12(7):629–639. DOI: 10.1109/34.56205
- [25] Weickert J. Efficient image segmentation using partial differential equations and morphology. *Pattern Recognition*. 2001;34:1813–1824. DOI: 10.1016/S0031-3203(00)00109-6
- [26] Charbonnier P, Blanc-Feraud L, Aubert G, Barlaud M. Deterministic edge-preserving regularization in computed imaging. *IEEE Transactions on Image Processing*. 1997;6(2): 298–311. DOI: 10.1109/83.551699
- [27] Matas J, Chum O, Urban M, Pajdla T. Robust wide baseline stereo from maximally stable extremal regions. In: *Proc. of the 13th British Machine Vision Conference*; 2–5 September 2002; Cardiff (UK). pp. 384–396.
- [28] Chaumette F. Image moments: a general and useful set of features for visual servoing. *IEEE Transactions on Robotics*. 2004;20(4):713–723. DOI: 10.1109/TRO.2004.829463
- [29] Fischler MA, Bolles RC. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*. 1981;24(6):381–395.

[30] Wang Z, Bovik AC. Mean squared error: love it or leave it?. *IEEE Signal Processing Magazine*. 2009;26(1):98–117. DOI: 10.1109/MSP.2008.930649

[31] Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. *IEEE Transactions on Image Processing*. 2004;13(4): 600–612.

[32] Horé A, Ziou D. Image quality metrics: PSNR vs. SSIM. In: *Proc. of the 20th IAPR International Conference on Pattern Recognition*; 23–26 August 2010; Istanbul (Turkey). p. 2366–2369. DOI: 10.1109/ICPR.2010.579



---

Section 4

# Future Mobility

---



# Future Mobility Advances and Trends

*Michela Longo, Wahiba Yaïci and Federica Foiadelli*

## Abstract

The trends of main interest on a global scale are those that can influence the development of humanity in the long term and are sometimes referred to as megatrends. The changes they bring with them can span several generations, profoundly changing society and, consequently, the competitive landscape of companies. The megatrends are numerous and each one involves the development of entire areas of activity. It is important to identify the megatrends of interest for strategic mobility planning and follow their developments, in order to consider them in the planning processes and correctly pilot investments. Megatrends are made possible and also influenced by the offer of new technologies, and lead to changes in cultural models. This chapter shows an era characterized by major technological innovations that are changing people's ways of thinking and acting, with the establishment of new mobility models in order to meet new emerging needs.

**Keywords:** transportation system, smart mobility, electric vehicle, mobility as a service (MaaS)

## 1. Introduction

The transport sector acquires a key role in promoting a correct balance between the different components of sustainable development. On the one hand, in fact, the mobility of people and goods and the conditions with which it is met (times, price, safety, reliability) decisively influence the present and future competitiveness of production and territorial systems, and, jointly, the accessibility to a series of basic functions within modern societies: work activities, educational services, social and health services, leisure and recreational activities, etc. [1, 2]. It therefore represents a fundamental component of the economic and social dimension of sustainable development [3, 4]. On the other hand, the quantitative evolution of demand volumes and the relative modal shares is at the basis of important critical issues from the point of view of eco-compatibility and security of supply, or two essential determinants of the concept of intergenerational equity that is at the basis of sustainable development [5, 6]. The inability to completely dissociate the demand for transport from the evolution of economic indicators and the almost absolute preponderance of fossil fuels in satisfying it have in fact led to a continuous increase in the contribution of transport to climate-altering gas emissions in recent decades and, at the same time, to increase the vulnerability of present consumption trends with respect to the exhaustion of non-renewable resources and dependence on foreign countries [7, 8]. In this scenario, transport policy becomes a decisive hub for achieving global environmental commitments, including those envisaged by the Kyoto Protocol and

subsequent developments, as well as the objectives of the Community energy policy. The sectoral dynamics also determine effects characterized by a particular territorial connotation, which assumes specific importance at the level of urban areas, where a preponderant share of movements are carried out and where, consequently, a series of characterizing problems are concentrated: delays due to congestion, employment of soil and competition with other uses (homes, commercial activities, non-motorized vehicles, green spaces), local pollution emissions and greater exposure of targets (people and things), visual and landscape intrusion [9, 10]. The management of choices and the ability to change the trends experienced up to now become in this scenario one of the essential components of urban sustainability policies, decisively influencing the quality of life and the overall level of attractiveness of cities. The need to attribute a specific value to transport, both in Ref. to global issues (climate change, energy dependence) and to those related to the local dimension (congestion, atmospheric pollution, noise, etc.), finds recognition in the European Development Strategy Sustainable [11, 12], which identifies the ability to promote a model of “sustainable transport” as one of the seven key challenges that the European system must face in the future.

A challenge based as for the other economic and social sectors, on the affirmation and diffusion of new technological solutions in the production/consumption patterns, but also, if not above all, on the recognition of the need to assign a transversal value to the mobility issue and related choices of satisfaction within the various sectoral policies (trade, industrial logistics, tourism, planning and management of the territory, etc.) in order to pursue the first (and functional to all the others) operational objective, or the dissociation of volumes of demand from economic growth.

## **2. New paradigms of society**

In a constantly changing society there are three fundamental aspects that need to be considered:

- Sharing (sharing economy): resources, especially if in excess, will be shared with others (“prosumer”, from consumer to producer of services). The demand will increasingly be oriented towards the use of shared services (in the United States car owners have drastically decreased: from 74% of Generation X, born between 1960 and 1975, it has gone to 48% of Millennials).
- Information (big data and data analytics): there will be large amounts of data available from which to extract information, also to offer new services. Those who can use the data will enjoy enormous competitive advantages.
- Supply of customizable and integrated services: the services will be customizable on the basis of demand, integrating those also provided by different subjects (providers). These trends will change the characteristics of the demand and supply of services also in the field of mobility.

According to [13, 14] it emerged that in 2050 two thirds of the world population will live in urban areas (over six billion people); the total amount of urban kilometers traveled will triple compared to the current situation; the costs for urban mobility will amount to over 800 billion per year; over 17% of the planet’s biocapacity will be used for urban mobility. In addition, regarding the urban distribution of goods [15, 16] between 2006 and 2014 the number of commercial vehicles in the world went from



250 to 330 million, mainly due to e-commerce; e-commerce turnover volumes are estimated to increase by 85% between 2015 and 2020 [17]. The demand for mobility of people and goods in urban areas has grown and is destined to increase further and it will not be possible to satisfy it by increasing the infrastructure. It will be necessary to switch to a disruptive technology that is the type of innovation that is considered when it quickly and radically changes a market or the ways in which to operate in it.

Some of interest for mobility may be:

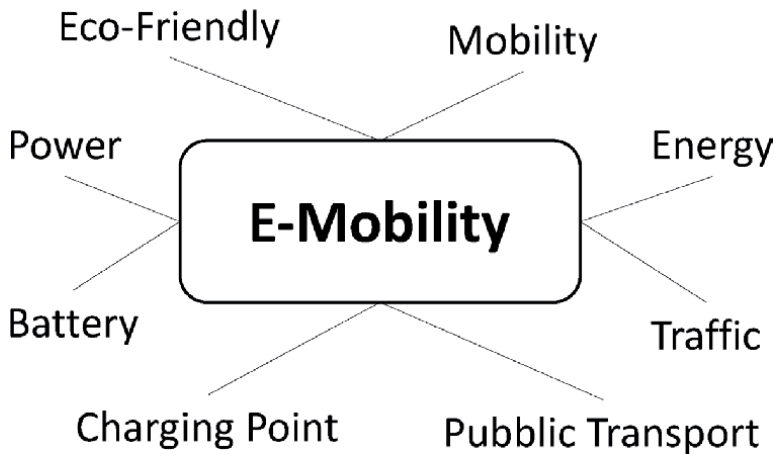
- Artificial intelligence;
- Autonomous vehicles;
- Electric mobility;
- Big data and data analytics;
- Internet of Things (IoT) and Internet of Everything (IoE);
- 5G, connected vehicles (V2X);
- Technologies for Blockchain transactions.

Widely available over the next decade, these technologies will have a profound impact on mobility services.

### 3. E-mobility

The transport sector, with particular reference to the passenger car segment, being one of the main contributors to CO<sub>2</sub> emissions, must undergo substantial improvements in environmental efficiency. Vehicle electrification is often seen as the primary option to help achieve this goal. Although electrification is a recurring theme in the history of the automotive industry, in recent years some changes in the reference context have opened up new development opportunities for electric vehicles: the phenomenon of climate change, the increase in oil prices and the long-term oil shortages, major technological innovations in sectors relevant to the automotive industry (e.g. in the battery industry), pressures to introduce innovations in the automotive sector and the response of manufacturers to the requirements contained in European legislation for reduction of carbon emissions [18, 19].

E-Mobility has become a keyword. Refers to vehicles that use electricity as their main source of energy, with the possibility of recharging the battery by connecting with an outlet to the electrical network, regardless of whether the vehicles are equipped with an auxiliary internal combustion engine to be used in long journeys distances or to keep the battery charged (battery electric vehicles, plug-in hybrid electric vehicles and extended range electric vehicles). This system is not limited only to passenger cars, but also covers motorized two-wheeled vehicles, quadricycles, vans, etc. E-Mobility currently dominates the debate on the future of transport and is becoming popular with policy makers, research institutes and industry. National and local authorities are already providing support for the introduction of these low-carbon vehicles, granting them special tax treatments or favoring their use, compared to conventional cars, with other measures (parking facilities, access to traffic areas limited use of preferential lanes, etc.). **Figure 1** illustrates the concept of E-mobility.



**Figure 1.**  
*Concept of E-mobility.*

The transition from a conventional to an electric car is not automatic, especially if users are not actively involved in the process and if they are not assisted in understanding the meaning and advantages of these new technologies [20, 21]. It is still necessary to overcome not only some major uncertainties in the market, which affect the propensity to buy and consumer behavior (regarding costs, autonomy and viability of electric mobility), but also delicate political issues. The Authorities should favor the development of e-Mobility without creating market distortions, adopting a principle of technological neutrality: the incentives should be linked to performance in carbon emissions (“from well-to-wheels”) and not to a specific technology. Furthermore, the incentives should not further aggravate overall energy taxation; the spread of electric vehicles should be linked to the use of renewable energy (with a positive environmental impact “from the well to the wheels”); Standardization Bodies and the industrial sector should agree, adopting common standards and protocols regarding the systems and devices for recharging batteries and the communication and information systems associated with them.

However, also with the aim of reducing CO<sub>2</sub> emissions by improving the efficiency of internal combustion engines (e.g. by reducing vehicle weight and engine power) and by increasing the use of alternative fuels (methane, biodiesel, etc.), electric vehicles could be an important way to improve individual mobility while minimizing emissions, representing a major challenge for European industries. The development of electric mobility, in fact, will depend not only on the adoption of specific technologies, but also on the ability to organize and manage the activities of different actors: automotive industry, battery manufacturers, mobility service providers, energy suppliers and distributors [22, 23].

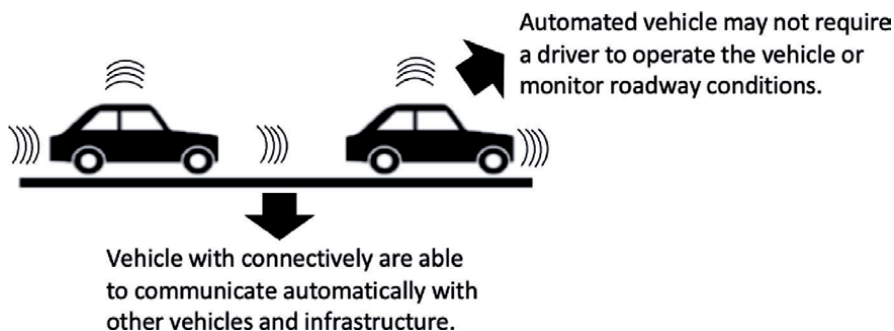
It is important to understand the mobility needs and the types of demand that electric vehicles can meet and the performance they can offer to families compared to cars with internal combustion engines. Excluding future advances that may allow the capacity of electric vehicles to be increased, the latter currently offer a limited range compared to traditional cars, with the possibility of quick recharging only by replacing the battery. Therefore, electric vehicles are better suited for travel in urban areas and over short distances. This need not necessarily be a handicap once consumers understand the difference between battery electric vehicles and conventional vehicles, and the benefits the former can offer: zero emissions “from tank to wheels”, affordable charging costs, flexibility in urban areas, etc. Electric vehicles offer consumers a wider choice to meet their mobility needs.

The important challenge is to be able to meet the different demand needs even if using different technologies. Therefore, the policies should guarantee the presence on the market of a differentiated mix of technologies. Meanwhile, hybrid vehicles currently on the market - including plug-in hybrids - offer a range comparable to vehicles with traditional engines. Alongside a diversification of the demand for mobility, there are also some changes in the use of the car. The high costs of batteries, current and predictable, associated with a limited range will continue to represent barriers to the purchase of electric vehicles. These difficulties could be mitigated through different mechanisms, such as car-sharing systems, corporate fleets and leasing. A conceptual change in the use of cars is observed, especially among young people: from an owned asset to an asset that can be rented only when necessary, such as is happening with bike-sharing services. Electric cars, such as electric motorcycles and bicycles, could reinforce this new relationship between citizens and mobility.

As a guideline, between 2025 and 2030 we could reach the breakeven point between the prices of electric cars (BEVs) and those with internal combustion power trains. All these will depend on the technological evolution of the accumulators; the growth in demand for batteries; the methods of their reuse and/or disposal; from access to raw materials to make them. As experiences in other countries (Denmark, Norway) show, the spread of electric-only vehicles is strongly linked to the availability of economic incentives. A widespread diffusion of the recharging network (columns) is fundamental. Dynamic inductive charging is being tested and it is not currently possible to understand if and when it will contribute to the spread of BEV power trains. The tendency to promote electric mobility in urban areas is shared by regulators. Technology providers are gearing up to answer this question.

#### 4. Connected and cooperative vehicles

The technology for the highest levels of automation is already available, but it is nevertheless necessary to gradually prepare for the impact that the phenomenon will have on the way of using vehicles. Mobility understood as the ability of people to move, in the shortest possible time, with the least use of resources and at the same time reducing the environmental impact should be a strategic objective for the institutions, with a view to integrating public mobility systems and collective with those of private and individual mobility [24, 25]. Obviously, this integration is also related to the complex issue of traffic management, especially urban traffic and in this sense computer networks will tend to take on ever greater importance due



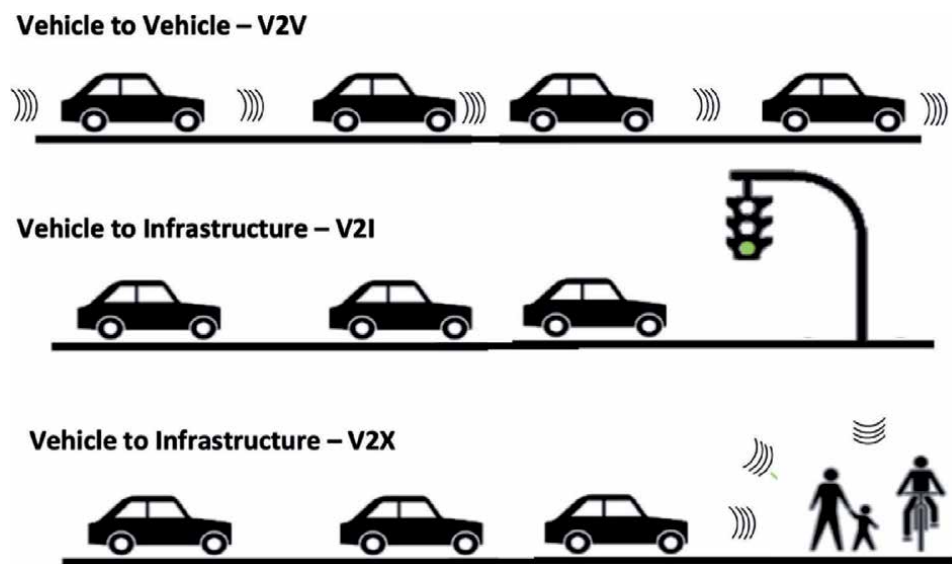
**Figure 2.**  
*Example of connected vehicles.*

to the amount of useful information that can be exchanged and the possibility of crossing millions of data also in a predictive function. **Figure 2** shows an example of connected vehicles.

Expressions such as smart mobility and smart city are used to indicate intelligent infrastructural and mobility systems. With these terms we mean that set of logistics and transport systems that are supported and integrated by ICT. In particular, smart mobility refers to a new mobility model that uses new technologies for road safety and integrates information and innovations on board the vehicle to increase transport efficiency [26, 27].

Smart city does not mean, of course, digital city, even if in the past the tendency was to essentially make the two expressions coincide. The goal of the smart city is not digitization, which is instead an effective and flexible tool for improving many aspects of the quality of life of citizens and promoting the country's economic growth. The approach to the issue of smart cities brings with its undoubted elements of difficulty, just think of identifying the interventions to be carried out, their alignment with the economic and social context of the city and the assessment of the impact on the community, without considering that the various projects, once conceived, must be able to be effectively carried out in that specific urban and social context. It is possible to argue that a city can be defined as "smart" that is to say "intelligent" when, according to a strategic, integrated and organic vision, using ICT tools to improve the lives of its citizens, it uses real-time information from various areas and exploits both tangible (e.g. infrastructure, energy and natural resources) and intangible (e.g. human capital, knowledge) resources, adapting from time to time to the needs of users with a view to sustainable development [28, 29].

It was estimated that in 2020 approximately 75% of new vehicles were able to connect to the internet, thus accessing different services and potentially allowing the exchange of information with the infrastructure (V2I), with other vehicles (V2V) and, generalizing, with anyone (V2X) (for example for updates of on-board software (SW) or the acquisition of travel information by various service providers). **Figure 3** presents the different types of connections.



**Figure 3.**  
*Different types of connections.*

The spread of connected vehicles supports the adoption of increasingly high levels of automation (autonomous driving) and access to articulated mobility services offered by different subjects. Connected travelers and vehicles will become Internet of Things (IoT) or Internet of Everything (IoE) nodes. The demand for mobility will therefore be satisfied through a multimodal, on-demand, and shared offer. The consumer will have multiple offers, more choice between different service levels. Public and private operators will coexist. Vehicles and travelers, as nodes in the network, will generate data that will allow, if shared, an optimization of the offer and resources. The trend to take advantage of shared services and integration into the web through the IoT will push to meet the needs of mobility by accessing a different mobility service according to needs, rather than buying vehicles or making medium-long term choices.

## 5. Autonomous vehicles

The term autonomous refers to the ability and faculty to govern or stand alone. According to the NHTSA (National Highway Traffic Safety Administration), or the US government agency for road safety, it defines an autonomous car “*a vehicle whose operation takes place without direct intervention by the driver to control steering, acceleration and braking and which is designed in such a way that do not expect to constantly check the road, when the automatic mode is running*” [30, 31].

In order to achieve a certain level of autonomy, the car exploits the ability to detect the surrounding environment through techniques such as radar, LIDAR, GPS and sensors. Therefore, the interaction between these components and the advanced control systems on board the car allows the latter to make decisions about the paths to follow and any obstacles and signals to monitor. To verify the degree of autonomy of the car, there are different classifications and standards coexisting with each other. The most adopted and followed by the scientific literature are the standard published by the NHTSA and the standard published by the SAE (a standardization body in the field of the automotive industry). In 2016, the NHTSA adopted the SAE J3016 standard, which therefore is configured as the reference standard. The latter has established six levels of autonomous driving that are based on the greater or lesser degree of automation of the vehicle, with the relative level of human participation in driving the car [32–34]:

- *Level 0*: No autonomy. The car does not have a driver assistance function and the driver is in full control.
- *Level 1*: Driving assistance. This level of automation requires the driver to make decisions as to when to accelerate, decelerate or steer but is informationally supported by other systems that may indicate the presence of hazards or adverse conditions. The car simply analyzes and represents situations in the form of visual or acoustic alerts. The driver has full responsibility for the vehicle.
- *Level 2*: Partial automation. In this degree of automation, the car is able to manage acceleration and deceleration through different types of systems such as assisted braking and emergency anti-collision braking. The direction and traffic control remain under the control of the driver.
- *Level 3*: Conditional automation. In this level the car begins to automate. It is able to manage acceleration, deceleration and steering, while the driver intervenes in problematic situations such as driving on dirt roads or where

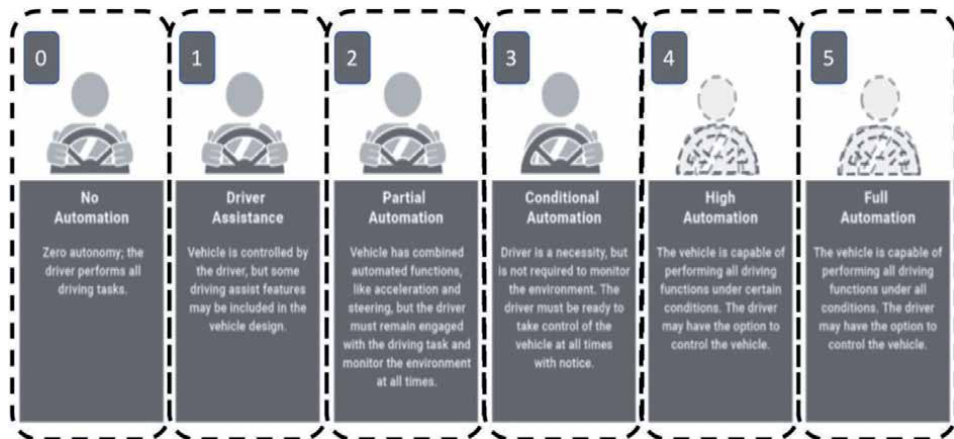
autonomous driving is not allowed or is too dangerous, for example in case of bad weather. The driver can, therefore, momentarily divert attention, but must quickly acquire control of the car if necessary.

- *Level 4*: High automation. This level provides for autonomous management of acceleration, deceleration, steering, and traffic control. The car handles the typical situations caused by traffic or traveling on urban or suburban roads. In this situation the car is able to drive in complete autonomy, but it is possible for the driver to regain full control of the car, if he so requests.
- *Level 5*: Complete automation. In this level, no intervention is required from the driver. The car drives exclusively autonomously, completely managing all the typical aspects of driving and based on the required tasks, it autonomously identifies the path to follow, take the right direction, accelerate or decelerate based on traffic conditions or upcoming situations. **Figure 4** shows a summary of automation levels as defined by the SAE.

There is significant consensus that, in the context of urban mobility, robotaxi fleets will be available between 2025 and 2030 (currently being tested). Autonomous vehicles and the services that can be activated thanks to their diffusion are expected to bring significant benefits to mobility [36, 37]. However, regulators will have to carefully govern the dissemination process. The advantages that can be found in the introduction of autonomous vehicles can be:

- Increase of road safety;
- Optimization of traffic flows with consequent reduction of urban congestion and better environmental impact;
- Mobility guaranteed to the entire population (elderly, disabled, minors) and complete territorial accessibility;
- Reduction of the “driver costs” currently incurred with the use of Local Public Transport (LPT) or Taxi;
- Reduction of parking areas;
- Transformation of time spent driving from unproductive to productive;
- Diffusion of a new shared mobility model based on car sharing and ridesharing.

The spread of shared and autonomous on-demand services will offer solutions for Local Public Transport in the first and last mile, as feeder services of the power lines. They must be part of a flexible and integrated public transport service. Autonomous vehicles should be shared as much as possible, and not merely replace the current private vehicles. However, the driverless, without driver is the future of motoring, and more. In a first phase, an authorized driver must in any case be present in the driving seat: he will be able to carry out work or play activities but must always be available to regain control of the vehicle if requested by the computer system. In a second phase, less distant than one might imagine, there will be no driver, but only passengers in a vehicle entirely managed by technology. Every year around 1,400,000 people in the world die from being involved in



**Figure 4.**  
*Automation levels defined by the SAE [35].*

road accidents: a massacre. The most accredited statistics on the causes of these accidents attribute them, in about 90% of cases, to inappropriate behavior or distraction of the driver. In only about 2% of cases, the responsibility is attributed to technological defects [38, 39].

With the widespread use of driverless, the decrease in the number and severity of accidents will be drastic: this is indicated by all independent scientific forecasts currently available.

However, technology still poses challenges. For example, environmental perception can be made more robust through the fusion of information from different sensors, a research area in which further development is expected in order to be able to make full use of all the information available. In addition, new deep learning algorithms for object detection have shown significant performance gains, but still need to be extended in order to operate with fused data from different sensors. Still, despite recent advances in solving the localization problem, there are problems with long-term mapping. Updating the maps with static, topometric, activity and semantic data as time changes in order to ensure the vehicle can be located precisely and consistently with respect to the environment is an open research topic with many challenges to be solved.

Despite the significant advances demonstrated in the field of planning algorithms, further improvements are anticipated in the field of real-time planning in dynamic environments. The field of control has also shown important progress in recent years, however, many of the fundamental results obtained have only been validated in simulation. Ensuring that the autonomous system pursues the intentions of higher-level decision making is crucial. Finally, it has been demonstrated how vehicle cooperation (V2V) can increase the performance of the perception and planning process, but there is still much to be achieved to offer greater scalability of multi-vehicle cooperation algorithms and despite the fact that the hardware has been standardized, there is currently no standard that defines what types of information vehicles should exchange [40].

But technological issues are only one, probably minor, aspect of a problem whose solution involves evaluating several issues to consider. One of these concerns the regulatory and ethical problems. The first refers to the legislative question. It is necessary to have a regulation that modifies the highway code in order to allow the circulation of autonomous cars. At the moment only a few states have opened road sections dedicated to the transit of autonomous cars. In the United

States it is possible to test certain cars without a driver on board. In Japan, the test of autonomous cars without humans on board was allowed, as long as they were controlled and monitored remotely. In Germany, the presence of a human being is still required, but it is allowed to carry out technological tests while the driver can take care of other things, without having to keep their hands on the wheel. France is preparing a regulatory evolution to facilitate and expand the opportunities for experimenting with autonomous cars, as long as there is a human being on board.

Furthermore, the legislative problem is intertwined with the ethical-moral question, for which a definitive solution has not yet been found. This refers to who to attribute responsibility in the event of an accident. Who to blame in case of damage, if the manufacturing company or the owner/passenger and what decision to make the car make about who to save for example in a situation where the car has a school group in front of it and has to choose to avoid a collision with another vehicle. These issues are crucial in carrying out this technological diffusion.

## **6. Mobility as a service**

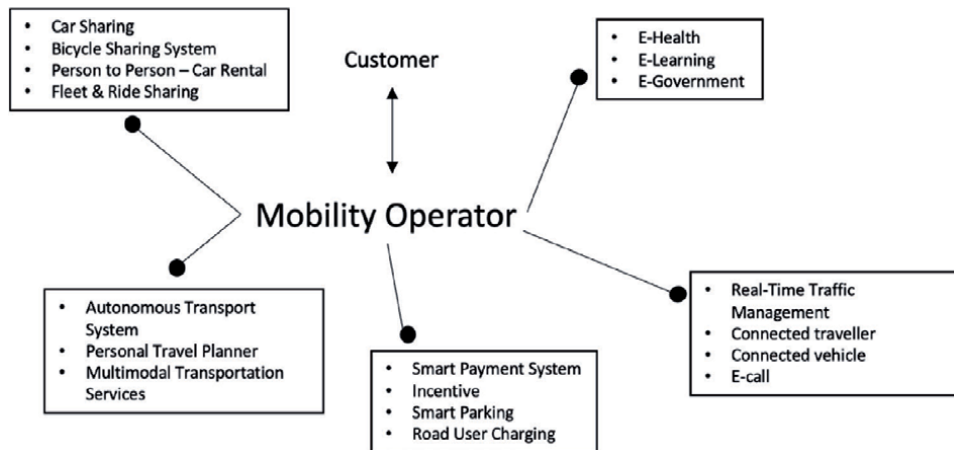
The acronym MaaS (Mobility as a Service) describes a new way of moving which, to the concept of personal ownership of the vehicle, replaces the concept of shared mobility understood as a service to be used according to need [41].

Moving from a lifestyle based on the possession of the means of transport, in particular the car, to a lifestyle based on the concept of Mobility as a Service, is not easy but considering mobility as a shared service offers many advantages for the individual citizen, for society and for the environment. MaaS is an ICT platform to manage the supply–demand meeting of transport and services offered by different subjects through a single information system interoperable with the proprietary systems of the individual operators. Service providers will also be able to operate on larger scales than the local one [42]. A successful MaaS service also brings new business models and ways to organize and manage various transport options, with benefits for transport operators including access to improved user and demand information, and new opportunities to meet unmet demand. MaaS's goal is to provide an alternative to private car use that can be cheaper, more sustainable, help reduce congestion and transport capacity constraints. For the user, MaaS can offer added value through the use of a single application to provide access to mobility, with a single payment channel instead of multiple ticketing and payment operations.

Mobility as a service is a relatively new concept that, in addition to changing the business model for the provision of transport services, promises a change in the means and methods of providing the service. This concept was created to be applied above all in large cities, where traffic congestion and levels of atmospheric and environmental pollution have reached their peak [43, 44].

Technology plays a fundamental role in making possible the spread of this business model, which has as its main feature the possibility for the citizen to choose the most suitable means of transport based on the route to be taken, passing from car to train, up to get to busses, trams, scooters and bicycles. In perspective, in fact, the user, through a single application, will have a service available on his smartphone that will allow him to plan the trip and to choose which means of transport to use for each journey to be made, paying for the single trip or taking advantage of monthly passes or unified rates for several different means of transport. The main feature of MaaS lies in offering travelers solutions based on their real travel needs. To do this, it is essential to combine public transport service providers (such as busses, trams and trains) with private services such as car sharing, bike sharing or





**Figure 5.**  
 Mobility as a service framework.

car rental services. In this way, through a single platform, users can plan their trip and pay using a single account. The most advanced platforms will have to be able to show the user the different travel options with relative prices and travel times, to allow him to choose the best solution according to his needs.

Once the trip has been planned, the natural evolution of the service lies in allowing the user to book the means of transport directly in the App (taxi, car sharing, scooter, train) to be sure to arrive at their destination in the manner and on schedule without unnecessary waste of time [45].

In the long term, with a view to increasingly shared and sustainable mobility, Mobility as a Service should also allow roaming: a single application that can be used by the user to move around different cities without having to sign up for different services each time. Customization and flexibility in the transport system is an increasingly requested feature that has generated in recent decades a market space and growing interest in MaaS by both the public and the private sector. In the private sector, many services for sharing cars, bicycles, scooters and busses were born, for citizens and companies. But it is in the public sector that Mobility as a Service can be considered as a real revolution, able to connect trains, planes, trams and busses, to car sharing and bike sharing services that complete the range of customizable travel possibilities by the user. In Europe, the state that has made the most progress towards the concept of mobility as a service is Finland, where there are already pilot cases of MaaS. In Italy, on the other hand, the city that most of all believed in shared and sustainable mobility is Milan. **Figure 5** illustrates the framework of MaaS.

## 7. Sharing mobility

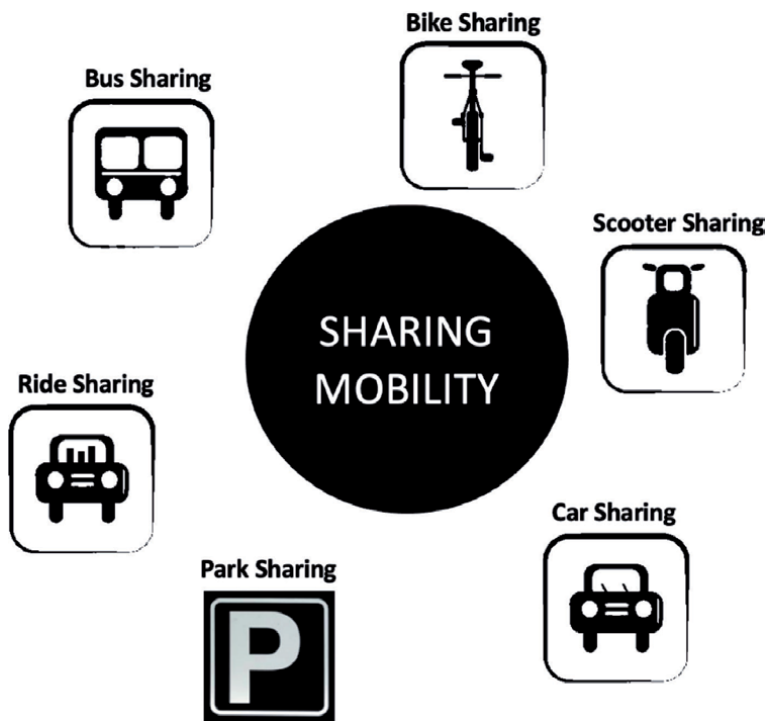
Shared mobility is a topic of great depth and importance as it is going to revolutionize the traditional essence of transport itself, and which takes the name of Sharing Mobility (SM) or also called Shared Mobility. SM is a particular mobility system, which allows people to move from one place to another, through shared vehicles [46, 47]. Users therefore do not only use proprietary vehicles for travel, but use rental services, which leverage on digital platforms for the provision of the service. This system describes a transport service that includes, public transport and taxis, bike sharing, car sharing, carpooling, scooter sharing, shuttle services

and others (**Figure 6**). This shared mobilization system relies not only on private users who make their own vehicle available to others, allowing access to it, but this has also been possible thanks to the birth of companies that make their services available (as for example Car2Go and Enjoy in Italy).

The SM aims to respond to new travel needs, trying to offer new options and solutions for transport. This system is able to provide more mobility choices to the user, put the last mile in contact with the first mile and reduce traffic congestion through the shared use of vehicles. In addition, it helps reduce air pollution, reduce transport costs, increase efficiency, and last but not least, it offers travel options for those who are unable to economically maintain a vehicle they own. This system also seeks to solve some historical problems inherent in traditional mobility, or to facilitate the sharing of vehicles and journeys between individuals, creating “tailor-made” services for each user of the platform, and maximizing the use of latent resources. Recent technological innovations have allowed the sharing of vehicles at lower transaction costs than in the past, thus allowing the sharing of vehicles that were normally designed for personal use (see for example Uber or Auting).

To classify a transport service under the “Sharing Mobility” label, certain characteristics must be present. In the first analysis, there is a need for the sharing of a mobility service, or that this service is shared between two users. This is possible in two different ways: there can be the use of the service simultaneously, as for example with BlaBlaCar, when the service is used simultaneously with other passengers, or differently the service can be offered in succession, when for example it is used a Car Sharing or Car-Pooling service (such as Uber) [48, 49].

Another important feature of shared mobility is the use of digital platforms, these are a necessary support for the creation of an original collaborative service. These platforms are based on the use of websites accessible from desktops, apps for smartphones and other mobile devices. Digital platforms allow levels of



**Figure 6.**  
*Sharing mobility.*

interactions unimaginable compared to the past, in fact they are able to easily put demand and supply in contact in real time, create relationships with service users, are easy and immediate to use and finally are more effective allowing to reduce transaction costs.

The use of the internet and the development of the “Information Technology System” (ITS) has proved to be of vital importance for these platforms; these tools have in fact allowed niche practices to impose themselves as forms of mass consumption, thus gaining visibility in the minds of consumers and at the same time market share.

The car in the model of traditional mobility is a “Status Symbol”, which indicates the acquisition of full freedom by the individual in travel and the ability to satisfy their interests: it is the quintessential symbol of emancipation. This has given rise to a model in which almost every individual owns one or more owned cars to make their trips around the city, with a consequent increase in the number of cars in circulation over time. Today, this increase is causing major problems both in the field of mobility and in the sustainable and environmental sphere, and this is how measures have been developed to reduce and discourage the use of individual vehicles, to induce users to more sustainable forms of transport.

SM can be a useful model for encouraging these behaviors. It is unthinkable to pass from a model that contemplates owned vehicles as a vehicle for travel, to a model that instead contemplates only shared means, without there being a period of experimentation and adaptation. The SM allows you to use your individual vehicle in a shared way, making your vehicle accessible to other users. Access to these services gives users new methods to evaluate which is the most favorable option regardless of the means of transport that you own or otherwise do not have at all. The main objective but also the most difficult to achieve since it must go to break already consolidated paradigms, is to encourage these practices of sustainable mobility, and make sure that the first choice of the individual is not to use his own vehicle, but rather the use of systems headed by the SM. Only in this way will it be possible to have decisive repercussions in the field of sustainability with the reduction of consumption of traditional mobility.

Technology plays a fundamental role in the orientation of users to choose the best vehicle to use for their journeys. There is a need to replace vehicle keys with smartphones, able to search, through mobility service aggregators, what is the best way to move to the preset destination.

SM could also solve traffic congestion problems. Often, in fact, the cars do not travel with their full load capacity, and many drivers are found on the streets who drive their vehicle without any passengers. This is even more relevant at peak times, when users are on their way to work or on their way home. The roads are invaded by thousands of vehicles at the same time, and this causes major traffic problems with consequent queues and delays. MS could partially solve this problem by reducing the number of vehicles on the roads. As previously mentioned, digital platforms are able to connect supply and demand in real time, and to bring users with the same travel needs together. With these tools, it would be relatively easy to organize the sorts of shuttle vehicles, traveling fully loaded to transport users who have a similar destination, thus reducing the number of vehicles needed to move on the roads. These forms of displacement are already present today but are still underused. There is a need to make users understand what the advantages are of sharing a vehicle.

A reduction of vehicles on the roads as well as advantages to the road mobility system would also lead to a reduction in polluting emissions from vehicles. MS is known not only for the use of innovative digital platforms, but also for the use of new forms of energy. This model is increasingly pushing towards an

eco-sustainable approach, through the use of vehicles that no longer rely on traditional fuels, but on new electrical technologies. Electric vehicles are known to be non-polluting vehicles as they use electric propulsion for driving and are silent. In the main cities there are today Car Sharing services that use these vehicles, which are parked in special parking areas equipped with an electric charging column. In addition to being vehicles that respect the environment more, they are characterized by low cost of refueling compared to fossil fuel vehicles, they also arouse great interest in people, since electric is still a technology that is little used for traditional travel.

Today we have reached a point where the traditional transport system is no longer able to support the needs of users, everyone wants to go everywhere in the most efficient way, but the road network is now congested by too many vehicles that travel it, creating big problems in moving. The car is the vehicle that guarantees the greatest versatility, it can be used for short urban journeys or for long journeys from one city to another: therefore, it allows the driver a degree of freedom and autonomy that no other public transport can guarantee. People today want to feel free, and they do not want to feel constrained in their movements, so they are looking for the type of mobility that can meet these needs. For these reasons, MS can be the solution to all these problems: it promotes access to mobility services with respect to vehicle ownership and uses a digital platform capable of representing the best travel solutions both from the point of view of the child, travel time and lower cost, both from the point of view of environmental impacts and efficiency. It is a model that goes against the traditional paradigms of mobility, but at the same time wants to satisfy the same needs: freedom and versatility of movement.

Owning a car is a significant cost in families' assets, in fact it can weigh up to 20% on family income. On the other hand, car sharing users are freed from these ownership costs, from the fixed costs of maintaining the vehicle, from insurance, and pay only what they consume. With traditional mobility based on ownership, we have reached a point where a vicious circle has been triggered, in which as the ownership of a private car increases, congestion on the road network and the need for new road infrastructure increases. With SM, on the other hand, we are experiencing the birth of a virtuous circle, in which the decrease in ownership and use of the private car follows a propensity to reduce ownership, which leads to a regeneration of the urban area and better accessibility within cities. Ownership by young people is perceived as something ancient, thanks to the internet they are now used to sharing, exchanging, reusing goods, and services. They are no longer willing to pay to own something for one-time use. They are inclined to pay for the actual use and are not interested in mere possession.

Innovations, especially radical ones, are capable of changing the game rules of a market. An innovation is said to be radical when it gives rise to new technological paradigms thanks to the Research and Development (R&D) of industrial or government laboratories, with the aim of combining product, process and organizational innovations to develop new markets. SM can be part of this context of radical innovation, as it is a model that is challenging traditional mobility and the most common transport methodologies.

Sharing Mobility can also be thought of as a disruptive type of innovation, an innovation that radically changes habits and the way consumers use a good or service, bringing about changes that can affect an entire ecosystem. As we have been able to analyze in the previous paragraphs, MS has led to new ways of conceiving mobility thanks to the support of new technologies, but many wonders if this model is truly capable of overwhelming and replacing the previous one based on ownership and possession of the vehicle.

## 8. Conclusions

In a decade, as a result of technologies, mobility management will be significantly more complex and business models will change. New mobility services (e.g., robotaxi) will be offered by more and more operators, public and private: carpooling, car sharing, ride sharing. It will be important to provide mobility services based on the integration, including multimodal, of Local Public Transport, private mobility, light mobility, shared transport services, etc. Public Transport operators will have to reposition their offer and services, forge alliances, review the value chain. A dynamic and integrated allocation of resources managed by various different entities (public and private transport services, physical network infrastructures, etc.) will be needed. An “intelligent” transport infrastructure will have to be developed, able to communicate with users and vehicles through multiple standards and control and regulation centers for the transport infrastructure offer will have to be implemented. It will be necessary to spread a greater culture, awareness of the economic value of the data generated as a result of access to shared services and autonomous driving. As a result of the high levels of vehicle automation, the spread of vehicle connections with the web, the cybersecurity aspects will assume absolute importance. As a result of its complexity and integration, mobility could be more vulnerable to malfunctions of its components (communication networks, power grid, control centers, etc.). It is therefore a field of work that is still open, in constant evolution.

## Conflict of interest

The authors declare no conflict of interest.

## Author details

Michela Longo<sup>1\*</sup>, Wahiba Yaïci<sup>2</sup> and Federica Foiadelli<sup>1</sup>


1 Federica Foiadelli, Politecnico di Milano, Department of Energy, Milan, Italy

2 CanmetENERGY Research Centre, Natural Resources Canada, Ottawa, Ontario, Canada

\*Address all correspondence to: [michela.longo@polimi.it](mailto:michela.longo@polimi.it)

## IntechOpen

---

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] “Mobility 2030: Transforming the mobility landscape”, KPMG, 2019: <https://assets.kpmg/content/dam/kpmg/xx/pdf/2019/02/mobility-2030-transforming-the-mobility-landscape.pdf> (Accessed by March 2021)
- [2] “McKinsey Center for Future Mobility®”, McKinsey & company: <https://www.mckinsey.com/features/mckinsey-center-for-future-mobility/overview#> (Accessed by March 2021)
- [3] “Autonomous Mobility and Energy Service Management in Future Smart Cities: An Overview”, Xiaoqi Tan, Alberto Leon-Garcia, 2018: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8642141> (Accessed by March 2021)
- [4] “Race 2050 – a vision for the european automotive industry”, McKinsey & company, 2019: [https://www.mckinsey.com/~/\\_/media/mckinsey/industries/automotive%20and%20assembly/our%20insights/a%20long%20term%20vision%20for%20the%20european%20automotive%20industry/race-2050-a-vision-for-the-european-automotive-industry.pdf](https://www.mckinsey.com/~/_/media/mckinsey/industries/automotive%20and%20assembly/our%20insights/a%20long%20term%20vision%20for%20the%20european%20automotive%20industry/race-2050-a-vision-for-the-european-automotive-industry.pdf) (Accessed by March 2021)
- [5] “Future Networks 2030: Challenges in Intelligent Transportation Systems”, Mădălin-Dorin Pop, Jitendra Pandey, Velmani Ramasamy, 2020: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9197951>
- [6] “Future Networks 2030: Architecture & Requirements”, Anastasia Yastrebova, Ruslan Kirichek, Yevgeni Koucheryavy, Aleksey Borodin, Andrey Koucheryavy, 2018: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8631208>
- [7] “Smart Transportation System: Mobility solution for Smart Cities”, Samir Maqbool Al-Shariff, M. Saad Alam, Zaureh Ahmad, Furkan Ahmad, 2019: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9124988>
- [8] “Big Data Analytics for Electric Vehicle Integration in Green Smart Cities”, Boyang Li, Mithat C. Kisacikoglu, Chen Liu, Navjot Singh, and Melike Erol-Kantarci, 2017: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8114543>
- [9] “Highlighting the future of Autonomous vehicle technology in 2020-2050”, Nedaa Baker Al Barghuti, Deepa Pavithran, Huwida E. Said, 2018: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8649510>
- [10] “How to Conceive Future Mobility Services in Smart Cities according to the FIWARE frontierCities Experience”, Lorenzo Carnevale, Antonio Celesti, Maria Di Pietro, Antonino Galletta, 2018: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8497005>
- [11] “Smart city drivers and challenges in urban mobility, health-care, and interdependent infrastructure systems”, Amro M. Farid, Muhannad Alshareef, Parupkar Singh Badhessa, Chiara Boccaletti, Nelio Alessandro Azevedo Cacho, Claire-Isabelle Carlier, Amy Corriveau, Inas Khayal, Barry Liner, Joberto S.B. Martins, Farokh Rahimi, Rosaldo Rossetti, Wester C.H. Schoonenberg, Ashlynn Stillwell, and Yinhai Wang, 2020: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9307293>
- [12] “Planning the Second Generation of Smart Cities”, Itai Dadon, 2019: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8818647> (Accessed by March 2021)
- [13] S K. Kaur, “A Survey on Internet of Things – Architecture, Applications, and Future Trends,” in 2018 First International Conference on Secure

Cyber Computing and Communication (ICSCCC), Jalandhar, India, 2018, pp. 581-583.

[14] "The Internet of Things for Intelligent Transportation Systems in Real Smart Cities Scenarios", Alberto Attilio Brincat, Federico Pacifici, Stefano Martinaglia, Francesco Mazzola, 2019: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8767247>

[15] "A Closer Look at the IoT's "Things"", Jeffrey Voas, Bill Agresti, Phillip A. Laplante, 2018: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8378976>

[16] "Application of Internet of Things and Big Data towards", Preeti Yadav, Sandeep Vishwakarma, 2018: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8519920>

[17] "A Study on a Routing-Based Mobility Management Architecture for IoT Devices," M. Ishino, Y. Koizumi and T. Hasegawa, 2014 IEEE 22nd International Conference on Network Protocols, Raleigh, NC, 2014, pp. 498-500, doi: 10.1109/ICNP.2014.78.: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6980417>

[18] "From electric mobility to hydrogen mobility: current state and possible future expansions", Guido Ala, Vincenzo Castiglia, Gabriella Di Filippo, Rosario Miceli, Pietro Romano and Fabio Viola, 2020: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9140482>

[19] "Battery Based Last-Mile Module for Freight Electric Locomotives," M. Brenna, F. Foadelli and J. Stocco, 2019 IEEE Vehicle Power and Propulsion Conference (VPPC), Hanoi, Vietnam, 2019, pp. 1-6, doi: 10.1109/VPPC46532.2019.8952376: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8952376>

[20] "E-Mobility — Advancements and Challenges", Aswad Adib, Khurram K. Afridi, Mahshid Amirabadi, Fariba Fateh, Mehdi Ferdows, Brad Lehman, Laura H. Lewis, Behrooz Mirafzal, Maryam Saeedifard, Mohammad B. Shadmand, Pourya Shamsi, 2019: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8895956>

[21] "Dynamic Wireless Charging of Autonomous Vehicles: Small-scale demonstration of inductive power transfer as an enabling technology for self-sufficient energy supply.", Giuseppe Guidi, Anastasios M. Lekkas, Jon Eivind Stranden, and Jon Are Suul: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9024243>

[22] "Study of Wireless Charging Lane for Electric Vehicles", Jiongran Xiao, Eric Cheng, Norbert Cheung, Bo Zhang, J. F. Pan, 2016: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7845989>

[23] "A review on the key issues for lithium-ion battery management in electric vehicles", Languang Lu, Xuebing Han, Jianqiu Li, Jianfeng Hua, Minggao Ouyang, 2013: [https://www.researchgate.net/profile/Languang\\_Lu/publication/257225400\\_A\\_review\\_on\\_the\\_key\\_issues\\_for\\_lithium-ion\\_battery\\_management\\_in\\_electric\\_vehicles/links/5c42fe6ba6fdccd6b5b84a94/A-review-on-the-key-issues-for-lithium-ion-battery-management-in-electric-vehicles.pdf](https://www.researchgate.net/profile/Languang_Lu/publication/257225400_A_review_on_the_key_issues_for_lithium-ion_battery_management_in_electric_vehicles/links/5c42fe6ba6fdccd6b5b84a94/A-review-on-the-key-issues-for-lithium-ion-battery-management-in-electric-vehicles.pdf) (Accessed by March 2021)

[24] "E-Mobility & Microgrid Laboratory at the Savona Campus of Genova University", Stefano Bracco, Federico Delfino, Giorgio Piazza, 2020: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9241138>

[25] "How Electric Vehicles and the Grid Work Together", 2018: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8501603> (Accessed by March 2021)

- [26] “Power Interchange Analysis for Reliable Vehicle-to-Grid Connectivity”, Saba Al-Rubaye, Anwer Al-Dulaimi, and Qiang Ni, 2019: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8808171>
- [27] “Optimized power flow control of smart grids with electric vehicles and DER”, Metody Georgiev EORGIEV, Rad Stanev, Anastassia Krusteva, 2019: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8771575>
- [28] “Analysis of Electrical Vehicle behavior from real world data: a V2I Architecture”, Luca Bascetta, Giambattista Gruosso, Giancarlo Storti Gajani, 2018: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8493203>
- [29] “Reliability Verification Procedure of Secured V2X Communication for Autonomous Cooperation Driving”, Han-Gyun Jung, Dae-Kyo Shin, Ki-Taeg Lim, Sang-Hun Yoon, Seong-Keun Jin, Soo-Hyun Jang, Jae-Min Kwak, 2018: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8539617>
- [30] “Functional architecture for autonomous driving and its implementation”, Rihards Novickis, Aleksandrs Levinskis, Roberts Kadiis, Vitalijs Fescenko, Kaspars Ozols, 2020: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9276943>
- [31] “The autonomous mobility innovation lifecycle”, Evangelos Simoudis, 2019: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8943257>
- [32] “Autonomous Vehicle Ethics Stock or Custom?”, Sally Applin, 2017: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7948873>
- [33] “Incorporating Ethical Considerations into Automated Vehicle Control”, Sarah M. Thornton, Selina Pan, Stephen M. Erlien, and J. Christian Gerdes, 2017: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7588150>
- [34] “Public Health, Ethics, and Autonomous Vehicles”, Janet Fleetwood, 2017: <https://ajph.aphapublications.org/doi/pdfplus/10.2105/AJPH.2016.303628>
- [35] “Smart Car Road Testing 101”, 2020: <https://www.acmwillowrun.org/smart-car-road-testing-101/> (Accessed by March 2021)
- [36] “Understanding autonomous vehicles: A systematic literature review on capability, impact, planning and policy”, Asif Faisal, Tan Yigitcanlar, Md Kamruzzaman, Graham Currie, 2018: [https://conservancy.umn.edu/bitstream/handle/11299/209218/JTLU\\_vol-12\\_pp45-72.pdf?sequence=1](https://conservancy.umn.edu/bitstream/handle/11299/209218/JTLU_vol-12_pp45-72.pdf?sequence=1)
- [37] “Driving Information Logger with In-Vehicle Communication for Autonomous Vehicle Research”, Kyungbok Sung, Kyoungwook Min, and Jeongdan Choi, 2018: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8323732>
- [38] “New frontiers in driverless vehicles”, Brad Pietras, 2015: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7088509>
- [39] “A Preliminary Investigation of an Autonomous Vehicle Validation Infrastructure for Smart Cities”, Kyriakos M. Deliparaschos, Gergely Santha, Luca Zanotti Fragonara, Ivan Petrunin, Argyrios C. Zolotas, Antonios Tsourdos, 2020: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9201644>
- [40] “Development of Key Technologies for Autonomous Driving Vehicles”, Jason Sheng-Hong Tsai, Jyh-Ching Juang, Chia-Heng Tu, Tzong-Yow Tsai, Pau-Choo Chung, Chih-Chung Hsu, Chao-Yang Lee, Ching-Fu Lin, 2019:



<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9024730>

[41] “Urban Mobility Digitalization: Towards Mobility as a Service (MaaS)”, Luís Barreto, Antonio Amaral, Sara Baltazar, 2018: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8710457>

[42] “A Generic Future Mobility Sensing System for Travel Data Collection, Management, Fusion, and Visualization”, Linlin You, Fang Zhao, Lynette Cheah, Kyungsoo Jeong, Pericles Christopher Zegras, and Moshe Ben-Akiva, 2020: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8833515>

[43] “Evaluating citizens' willingness to uptake a MaaS tool for metropolitan multimodal trips”, Andres Monzon, Iria Lopez-Carreiro, Elena Lopez, 2019: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=9071696>

[44] “Mobility as a Service (MaaS) in rural regions: An overview”, Luís Barreto, Antonio Amaral, Sara Baltazar, 2018: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8710455>

[45] “Want to Ride My Bicycle: a Microservice-Based Use Case for a MaaS Architecture”, Franco Callegati, Giovanni Delnevo, Andrea Melis, Silvia Mirri, Marco Prandini, Paola Salomoni, 2017: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8024498>

[46] “Concept of interlinking mobility services for urban transport towards intermodal mobility including private and shared electromobility”, Daniel Breuer, Philipp Spichartz and Constantinos Sourkounis, 2019: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8813511>

[47] “Innovative approach of the sharing E-Mobility”, Mariacristina Roscia, Luigi Mingrone, Gianni Pignataro,

George Cristian Lazaroiu, 2016: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7526011>

[48] “The Future of Mobility—Electric, Autonomous, and Shared Vehicles”, Paul R. Donnellan, 2018: <https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=8625919>

[49] “Mobility Sharing as a Preference Matching Problem”, Hongmou Zhang and Jinhua Zhao, 2019: <https://ieeexploriee.org/stamp/stamp.jsp?tp=&arnumber=8478802>



# Trends in Next Generation Intelligent Transportation Systems

*Prem Chand Jain*

## Abstract

The objective of Intelligent transportation system (ITS) and related National highway traffic safety administration (NHTSA) is to improve vehicle safety and reduce accidents, injuries, and deaths. Advanced driver assistance system (ADAS) is making a difference in vehicle safety. The objective of ADAS is to provide a continuous picture environment surrounding the vehicle. This vision around the vehicle is seen by the driver to take the decision. Vehicular communication is a part of Intelligent Transport System which provides an intelligent way of transport to avoid accidents. As the transportation moves towards environment of connected and autonomous vehicles, the role of communication and data transfer becomes important. Connected vehicles can be used for both infotainment and navigation for vehicle safety. Vehicle-to-vehicle (V2V) communication allows vehicles to talk to each other and exchange data about location, direction of travel, speed, brake, accelerator status, and other facts. This information is analyzed and used to avoid collision. C-V2X (Cellular-Vehicle-to-Everything) can provide better quality of service support, large coverage, and high data rate for moving vehicles. Device-to-device (D2D) communication in C-V2X provides high reliability and low latency. In 5G Rel.16 C-V2X will become an integral part of 5G cellular network providing higher capacity, coverage, etc. Today old aged/disabled person look for driving technology that is convenient and easy to use. V2X technology will offset some of the concerns about old aged/disabled driver's abilities to respond quickly to challenge by driving environment as they no longer be required to handle most of the decisions.

**Keywords:** ITS, ADAS, DSRC, C-V2X, LiDAR, Vehicle, Autonomous

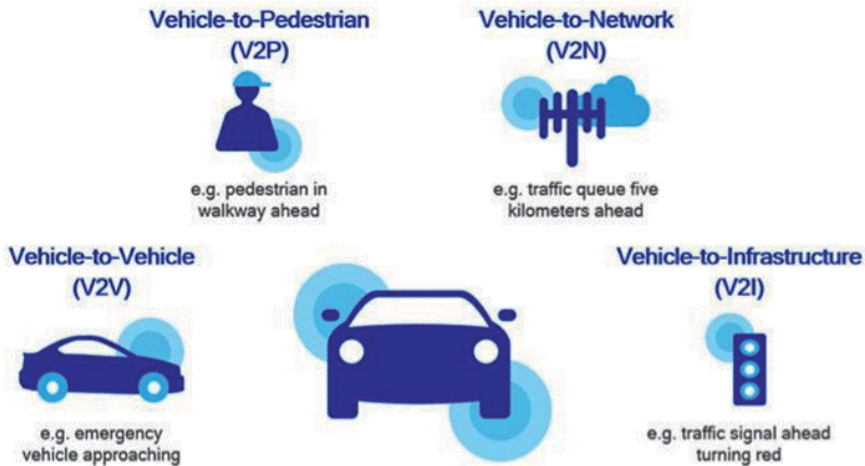
## 1. Introduction

There are many accidents occurring on roads due to negligence and lack of proper Intelligent transport system (ITS). Around 1.25 million people worldwide die from traffic accidents each year and between 20 to 50 million suffer from non-fatal injuries (WHO, May 2013). Studies proved that 60% of accidents can be avoided if a warning message is sent to that vehicle at least half a second before accident. The primary objective of ITS and related National highway traffic safety administration (NHTSA) to improve vehicle safety and reduce accidents, injuries, and deaths. Advanced driver assistance system (ADAS) is making a difference in vehicle safety. Manufacturers are adding ADAS to achieve greater safety in vehicles. ADAS additional cost will become low when ADAS is made part of vehicle cost. ADAS includes backup camera, lane-keeping steering, blind-spots detection,

automatic braking, parking assistance, pedestrian detection, traffic sign detection, night vision, etc. In ADAS, a cluster of sensors like video cameras, 77GHz millimeter (mm) wave radars like LiDAR (Light detection and ranging), and ultrasonic transducers feed their digital data via multiple serial interfaces to processors that store the data and process it. The vision processors prepare data for fusion processing, which combines the inputs from multiple sensors to produce feature recognition of nearby objects as well as their distance, motion, and status. They in turn make decisions and initiate the control of selected driving functions or provides driver notification by visual, audible signals. The objective of ADAS is to provide a continuous picture environment surrounding the vehicle. This vision around the vehicle is seen by driver to take decision. ADAS development involves massive software which includes special algorithms as well as Artificial intelligence (AI) technique such as machine vision and deep learning. It requires popular processors, interfaces, programing, and AI tools to speed up and simplify development.

As the transportation moves towards environment of connected and autonomous vehicles, the role of communication and data transfer becomes essential. Automakers are looking to offload critical, time-sensitive decisions-making from passengers to autonomous vehicles through a complex set of sensors and communication with everything around the vehicle. Connected vehicles can be used for infotainment and navigation to vehicle safety. There is large potential for connecting other road users such as heavy duty vehicles, pedestrians, cyclists, etc. The introduction of a communication interface into vehicles will also provide a pathway for state and local government that operate transportation system to communicate safety information from the Road side units (RSU) to vehicles. For the same the RSUs are connected to the Traffic management center (TMC). Given the volume of data to be transmitted, the RSU are connected to TMC through some short of back-haul. Wide spread audio/video (AV) technology deployment for infotainment will require robust communication system and communication infrastructure capable of moving the Megabit data generated in AV system. The integrated In-Vehicle Infotainment (IVI) helps to improve automobiles safety by providing the driver high quality information from audio, video, radar, and other sensors. It generates a seamless 360 degree view of vehicle surroundings from multiple independent HD video streams from front, rear, and side-mounted cameras. IVI also provides high quality audio, video entertainment. Imagine 10 to 20 cameras providing 360 degree view, all sending 4 k (3840X2160) resolution HD video streams with pixel depth increasing from 16 to 20 or even 24 bits. A single 4 k video camera with 24bits per pixel will produce around 200Mbit per frame at 10 to 30 frames per sec. This will range overall data rate around 6Gbps.

One forthcoming technology to be adopted is Vehicle-to-Everything (V2X) radio communication. V2X improves ADAS safety by providing radio communication between vehicles, and between vehicles and nearby roadside units that supply valuable information. V2X is a broaden term which includes V2V, V2I (Infrastructure), in addition to objects such as pedestrian crossing road (V2P), detecting bicycles, motor cycles on a car lane, detecting traffic light conflicts, and Internet based networks (V2N) which distribute software and firmware updates to vehicles including HD maps shown in **Figure 1** [1]. V2X technology has been mandated by NHTSA. Vehicle-to-vehicle (V2V) communication allows vehicles to talk to each other and exchange data about location, direction of travel, speed, brake, accelerator status, and other facts. This information is analyzed and used to avoid collision. The V2I will connect vehicles to Infrastructure that can inform about traffic congestion, road conditions, weather alerts, and construction going on to provide safety as well as convenience. The V2V communication is to avoid accidents by sending data about position and speed of vehicle in transit to another



**Figure 1.**  
V2X communication [1].

vehicle over an adhoc mesh network. It provides 360-degree awareness of surrounding threats. It supersedes techniques like blind spot detection, rear parking sonar, and backup camera. The main purpose of vehicle-to-vehicle communication is to provide an intelligent means of transport service. This is introduced in order to avoid accidents between vehicles by sending warning messages to each other. These warning messages consist of information regarding vehicles speed, emergency stopping, brake status, etc. V2V communication is like an additional step for warning the drivers. With the growth of mobile data, a cellular network has great potential to support various vehicular communication services for safety applications. Cellular system such as 4G-LTE can become the useful vehicle communication. Cellular V2X (C-V2X) can provide better quality of service (QoS) support, large coverage, and high data rate for moving vehicles. Device-to-device (D2D) communication in C-V2X provides high reliability and low latency, range, scalability, number of devices supported, security, and reduced cost of ownership. In 5G Rel. 16 C-V2X will be an integral part of 5G cellular network providing higher capacity, coverage. In this chapter Section 2 discusses V2X communication while Section 3 discusses Dedicated short range communication (DSRC). Section 4 discusses about Cellular vehicle-to-everything (C-V2X) while Section 5 discusses level of automation and Section 6 about self driving vehicles, and finally Section 7 conclude the chapter.

## 2. Internet of vehicles

Internet of vehicles (IoV) can be named as Internet of Things (IoT) on wheels. It will allow vehicles to communicate with their drivers, with other vehicles, with traffic signals, and city infrastructure.

### 2.1 Vehicle-to-vehicle communication

Inter-Vehicle communication uses multi-hop multicast/broadcast to communicate between each vehicle. Collision warning messages broadcast from V2V across multi-hops. This is suitable for short range, a vehicle communicates with another vehicle by using different protocols. In such cases receiver takes appropriate

decision on the basis of emergency messages received and accordingly takes appropriate action. To avoid collision in V2V communication, location based multicast and broadcasting is used [2]. Multi-hop communication propagates the message in the absence of RSU infrastructure. However, in low density vehicular network V2V communication is not very good solution due to large range. V2V could eliminate 80% of crashes that do not involve alcohol or drugs.

## **2.2 Vehicle-to-infrastructure communication**

Vehicle-to-Infrastructure (V2I) provides awareness of traffic light status, a road that is closed, to guide the vehicle around obstacles, traffic, road condition, weather, construction, etc. Vehicles connected to stationary infrastructure is known as Road side unit (RSU). Communication between vehicles and RSU are supported by V2R protocol. In V2R warning messages are sent first to RSU, and then RSU broadcast to all vehicles in the range. The Road side infrastructure involves additional installation costs. V2I provides large BW link between vehicles and RSU. The RSU can be deployed after every km to obtain high data rates required during heavy traffic. When a vehicle has mechanical failure or detects road hazards, vehicle generates an EWM (Emergency warning message) and keeps one copy with him for retransmission, if required. Vehicle broadcasts to neighboring vehicles and it transmits periodically to RSU also through transceivers with different frequency bands till it receives the message with the same ID from vehicle behind and RSU respectively. When RSU receives EWM from source vehicle, it replaces with own ID and forwards to all vehicles within the range [3].

## **2.3 Vehicle-to-network communication**

Vehicles connected to Application server. It transforms connected transportation around the globe. It allows one way broadcast to multiple vehicles thereby facilitating V2N functions. The network based communication (V2N) operates over licensed spectrum to support telematics, connected infotainment, and growing variety of advanced informational safety use cases.

## **2.4 Vehicle-to-pedestrian communication**

Vehicle is able to detect in advance pedestrians including cyclist, motor cycles in cross walks, blind spots, or other dangerous locations.

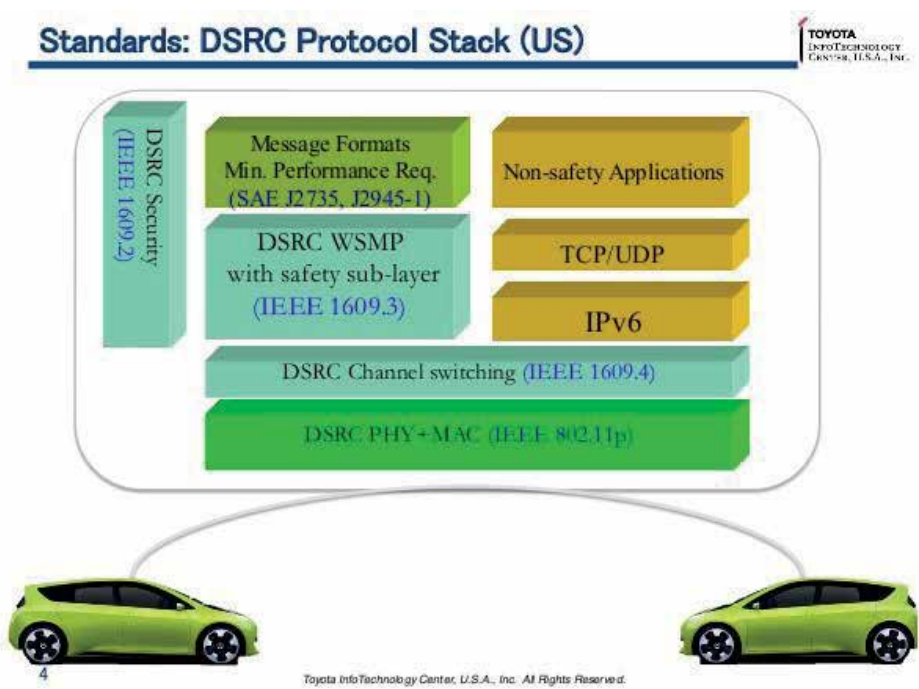
## **3. Dedicated short range communication**

NHTSA proposed V2V technology called Dedicated short range communication (DSRC) for all new light vehicles. As a result IEEE approved an amendment to IEEE802.11 standard named as 802.11p for Wireless access in vehicle environment (WAVE). The 802.11p is enhancement of 802.11a Wireless local area network (WLAN) required supporting ITS applications [4]. This includes data exchange between high speed vehicles and between vehicles and Road side infrastructure (V2R) in ITS 5.9GHz band. The DSRC combines Wi-Fi and GPS positioning to get a 360-degree awareness of all vehicles around them. It helps to provide the driver with warning to avoid collision using on-board computer. The cars share their positions very rapidly even though two cars could be around a corner. DSRC allows Cooperative collision avoidance (CCA) in which cars warns each other about changing conditions

to significantly improve safety. DSRC is a two way short-to-medium range wireless communication that permits very high data transmission in V2V communication. The main reason for using DSRC in V2V communication is to detect hazards in vehicle's path even though the driver is not in such a position to see. DSRC PHY (Physical) layer is targeted to operate in 5.9 GHz band (5.85–5.925GHz) with 75 MHz BW as opposed to IEEE 802.11a that allows only the unlicensed frequency band 5GHz. In MAC (Media access control) layer the DSRC band consists of seven channels each with 10 MHz BW and that include one control channel and 6 service channels. It can support a large family of vehicular safety and non-safety applications. Some requirements of MAC are low latency and high reliability. The PHY Layer provides better radio propagation with respect to multi-path reflection delay and Doppler effects occur due to high speed vehicles and road environment. The end-to-end delay is low (25 ms) and packet delivery ratio is high compared to 802.11a [5].

The 802.11p standard adds wireless access to vehicular networks and implements OSI layers stack. Upper layers of OSI follows IEEE1609 family standard as shown in **Figure 2**. Wireless Protocol works at licensed band of 5.9GHz with 300 m range, and data rate of 6 to 27Mbps. DSRC PHY layer adopts same OFDM and digital modulation types BPSK, QPSK, 16-QAM, 64-QAM as used in 802.11a.

The DSRC transmits Basic safety message (BSM) between vehicles. It includes information like exact vehicle location, direction of travel, speed, braking status, and some other data. The BSM is updated and transmitted 10 times per sec. DSRC is proven standard, mature technology, and cost effective but with little growth potential. Latency is around 25 ms which is not fast enough for collision avoidance action. Another application of DSRC is V2I but V2I depends on new development and deployment of infrastructure. This new infrastructure is a major investment.



**Figure 2.**  
 DSRC protocol stack.

#### **4. Cellular vehicle-to-everything (C-V2X) communication**

The Automation industry was agreed upon DSRC but suddenly C-V2X popped up disrupting the DSRC plans. The DSRC will require new deployment of thousands of roadside units (RSU) and associated infrastructure network such as fiber backhaul along the roads. This is a challenge for rural areas considering vast distances involved. The progress of DSRC has been delayed because of new alternative radio technology based on cellular mobile communication called C-V2X. C-V2X utilizes cellular mobile technology to provide the link between the vehicles and rest of the world including other vehicles and traffic control system. C-V2X provide high performance for capacity, coverage, range, scalability, number of devices supported, and security. C-V2X can achieve longer range which can directly translates into earlier alert and better visibility of unexpected dangerous situation. It also allows vehicles to travel at higher speeds while still being able to stop in time to avoid hazardous conditions compared to DSRC. It was designed initially to use 4G-LTE cellular mobile standard defined in 3GPP Rel.14 [6]. The new version 5G (Rel. 15) includes device-to-device (D2D) communication (V2V, V2I, V2P) and device-to-network (V2N) which will transform connected transportation around the globe [7]. The 4G-LTE and 5G can provide RSU functions eliminating the need of additional RSUs in DSRC. The V2V and V2I will connect and interact with ADAS providing intelligence beyond the short range environment covered by ADAS sensors. V2V and V2I will make ADAS an attractive alternative to full automation. Finally security is a major factor, if automobiles are connected to the Internet where they may be exposed to hack.

C-V2X also uses 5.9GHz band but instead of OFDM, it uses Single carrier-Frequency division multiple access (SC-FDMA) to lower down the Peak to average power ratio (PAPR) to reduce power consumption. It uses Turbo coding and Hybrid automatic repeat request (HARQ) protocol to enhance reliability in data transfer at high vehicle speeds, and lower latency (< 5 ms). C-V2X is designed to be globally compatible with 5G. The 5G will provide very high throughput, high reliability, low latency, and accurate position determination. 5G can offer multi-gigabit speeds for infotainment, telematics, and teleoperation. C-V2X is designed to offer low latency communication to V2V, V2I, and V2P. C-V2X encompasses two transmission modes, direct communication (V2V), and network based communication (V2N). The network based communication utilizes 4G and emerging 5G cellular network for V2N, and operates over licensed spectrum to support telematics, connected infotainment, and growing variety of advanced informational safety use cases. C-V2X is the gateway to the connecting vehicles and in long run to the self-driving vehicles. Full automated driving will be in 5G environment and C-V2X will be the bridge towards 5G [8]. C-V2X will be able to take 5G network advantages namely enhanced mobile broadband (eMBB), ultra low-latency communication (uRLLC), and massive scale machine-to-machine (M2M) communication (mMTC). This will enable more V2X services by providing long range, higher density, very high throughput, reliability, high precision positioning, and ultra low latency (1 ms). LTE variant LTE-m (Machine) and Narrow band (NB)-IoT Low power WAN (LPWAN) are not fast enough. LTE-b (Broadcast) is another variant which allows one way broadcast to multiple vehicles thereby facilitating V2I and V2N functions. 5G Automation association (5GAA) has 50+ automobile manufacturers, and mobile network industry including Audi, BMW, Daimler, Ericsson, Huawei, Intel, Nokia, Qualcomm to collaborate between automotive and mobile communication industry. 5GAA has recommended C-V2X to NHTSA. After taking the decision, 2 years period



would occur to accommodate manufacturers product development cycle, and full compliance would require additional 2 years.

#### **4.1 LTE-m**

Some potential candidates for V2V and V2I in 4G LTE cellular network are LTE-m and NB-IoT (Rel 13). LTE-m (Machine) is a versatile technology, supporting high data rates, mobility, and voice facility. It is a stripped version of LTE, uses 1.4 MHz instead of 20 MHz BW. Reduction in BW to 1.4 MHz will reduce size of FFT blocks results in cost reduction by 28%. Power consumption will also reduce as fewer number of subcarriers needs to be processed at RF level. This will further reduce the cost by 20 to 30% by means of RF transceiver design including low noise amplifier, mixer, and local oscillator. Overall average 39% cost saving and also modem complexity by 50% observed [9]. It effectively provides down link and uplink peak rates of 1Mbps covering 1.08 MHz bandwidth in half duplex mode. LTE-m supports full voice functionality via Voice over LTE (VoLTE) along with full mobility and in-vehicle hand-over. The major reason is that it can serve in automotive sector because of its extended range, deep penetration in buildings and basement, and low latency. The maximum uplink power transmitted by device is 23 dBm, and 46 dBm for downlink with 10–15 ms latency. It also supports power saving mode— sends an acknowledgement before going to sleep and then on waking up sends check along with data (if any) to the network.

#### **4.2 NB-IoT**

NB-IoT uses different technology (DSSS modulation in place of OFDM), but operates in LTE band [10]. With NB-IoT gateways are not required, and thus sensor data is directly transmitted to main server. NB-IoT has an advantage of low cost and lesser power requirements. The maximum transmitted power is 20 dBm over 200 kHz bandwidth (180 kHz one resource block). The complexity is reduced by 75% as compared to the LTE-m. Down link and uplink peak data rates supported by NB-IoT is around 250 kbps with 1.6 to 10sec latency. It works best for applications that requires moderate latency and throughput. It uses licensed band which eliminates interference and provides high security. It can provide 50 k to 100k vehicles connections per cell. It can operate by uploading a new software on LTE infrastructure. When transmitting 200 bytes in day on an average, one can achieve 10 years battery life time. NB-IoT Rel. 13 lacks mobile support and high power consumption while NB-IoT Rel. 14 support 160 kbps data rate and lower transmit power level.

#### **4.3 Automotive Ethernet**

Autonomous vehicles users will expect that their vehicle should provide seamless Internet connectivity to their laptop, mobile phone as provided in living room. It will be possible by using automotive Ethernet [11]. Ethernet is well known ubiquitous solution to traditional LAN (Local area network). The advantage of Ethernet is multipoint connections, higher BW, and low latency. IEEE introduced 802.3bu for automotive applications called 100Base-T1 supporting 100Mbps data on a single balanced twisted pair cable CAT5 or CAT6 of 15 m length [12]. To achieve 100Mbps it uses 3 bit per symbol (PAM3). It supports full duplex. Power over Ethernet (PoE) is being investigated and standardized using IEEE802.3bu with one pair power over data line group. The Gigabit ethernet (GbE) for 2.5/5/10GBase-T1 on single pair of wires standardized using IEEE802.3ch, is also being investigated for infotainment. Audi, BMW Mercedes have began implementation of Ethernet based connectivity.

## **5. Level of automation**

There are five levels of automation defined by the Society of automation engineers (SAE) [13]. Level 0 has no automation, driver performs all the functions. In level 1 driver performs all the functions but ADAS system provides alerts and partial control. Level 2 defines partial automation. Driver must still monitor actions but automated system controls braking, steering, and acceleration. Level 3 called automated driving systems performs all driving activities. Driver must still be available to take control in special circumstances. Level 4 automated driving systems perform all driving activities. Driver may still control the vehicle if needed or desired. Finally in level 5 automation no driver is needed.

## **6. Self driving cars**

Self driving cars also referred as Autonomous vehicles. It improves automotive safety and provides help to old age and handicapped citizen who may not be able to drive in difficult traffic conditions. Uber and Ola my also like to reduce the cost of personnel in driving activities. Currently automation level is 1 and 2. Testing is being carried out for level 4 and 5 but final product on road may take long time. Autonomous vehicles need ADAS with full sensor inputs from Radar, LiDAR, Video camera (360 degree field of view), and Ultrasound sensor unit. GNSS/GPS navigation mapping is also essential. Most modern vehicles use simultaneous location and mapping algorithm. It combines information from multiple sensors and an off-line map to calculate current vehicle location and generate more real-time map of the current environment. Real time HD mapping is critical ingredient for automated driving. HD maps play a critical role in path planning at cm level distance. All the data received from sensors are processed in real time so that steering, braking, and accelerator setting outputs are generated. This requires massive computing. Artificial intelligence (AI) has been responsible to implement because of improved performance of processors like GPU (Graphics processing unit). Some applications uses GPU cluster for training neural network. The challenge in neural networks is to determine weights, and number of nodes in each layer, and number of layers (or depth) of system. The weights are determined by training the network via set of inputs to recognize the object. Training uses feedback system when an input is matched with outputs and internal hidden layer weights are adjusted. Self driving cars do not sleep, do not drink, do not get phone calls, will certainly reduce car accidents caused by human error.

Auto Industries already ahead with manufacturers like Tesla, Audi, Mercedes, BMW produce level 2 automobile where driver not have to physically operate the vehicle, and can have his hands off the steering wheel and foot-off the accelerator/brake pedals at the same time. In level 3 human drivers fail to properly take over when necessary, or drivers of other cars are at fault. The Google automotive spin-off Waymo has recently started to test level 4 vehicle on public roads in Arizona, USA. According to GSMA 2013, 500 million cars will be connected by 2025, and 75% of cars on the road will be autonomous by 2035 as per Navigation Research 2013.

### **6.1 Autonomous motor bikes**

A new self driving motor bikes developed by M/S AB Dynamics promises to allow ADAS and autonomous system to be tested under much challenging

conditions. Motor bikes stability control utilizes advanced gyroscopes and accelerometers to detect parameters such as speed, lean angles, and braking force, and can quickly adjust electronic braking and throttle settings to help prevent crash. The system provides assistance by continuously monitoring a comprehensive set of key vehicle data including torque, lean angle, and acceleration to detect critical situations. This improves both riding stability and braking performance. According to Bosch Accident Research motor bike-to-car communication could prevent nearly 1/3 of the motor cycle accidents. The system uses DSRC IEEE802.11p. The motor bike will exchange information up to 10 times per sec. with other vehicles on the road within radius of several hundred meters. The information consists of vehicle type, speed, position, and directional travel. When a motor bike ends up in a car's blind spot or changes lanes to pass, this technology informs the car that motor bike is approaching. If system identifies a potential dangerous situation, it can warn the rider by sounding an alarm and flashing a warning notice on the dash board. The blind spot warning system works similar to those implemented in a car. A radar sensor serves as the blind spot recognition system's electronic eye, registering objects in hard-to-see areas. The small radar sensor will help to detect vehicles approaching and offer a warning ideally by illuminating light in the appropriate side mirror. This system keeps a lookout in all directions to add motorcyclist change lanes safely. The system is active as soon as the vehicle starts and it supports the rider in all relevant speed ranges. If the system detects another vehicle is dangerously close and the rider does not react to the situation, it warn rider by way of an acoustic or optical signal.

The streets of China, Asia, and India are filled with millions of motor bikes, manufacturers feel that autonomous technology built in to motor bikes can be a good solution to manage the movements of vehicles. Kawasaki, BMW, Honda, Yamaha, and Ford are making a lot of headway to bring the concept to fruition much sooner.

## **7. Conclusions**

V2X communication benefits the environment by reducing traffic jams that increases pollution. The coordination between V2I will reduce unnecessary braking further reduce fuel consumption and emission. C-V2X will save millions of lives but it will take time to equip into vehicles on road. The mobile network operators (MNO) need to offer C-V2X services by extending their networks to accommodate the C-V2X applications. For the same, service agreement would be required for each vehicle. Embedded SIM (eSIM) can be soldered in to cellular device of vehicles. GSMA has developed such eSIM profile specifications. ADAS and V2X form full autonomous vehicles. From level of automation it can be seen that level 5 automation may take long time to achieve with above, but level 4 can be achieved in shorter time. Although full automation is a goal but perhaps a better solution is simply let AI complement the driver rather than replace the driver. While some learning and adaptation will be required by the human, the powerful combination of human/AI should be able to provide the best safety improvement everyone needs. With C-V2X, 5G cellular network, and AI the intelligent connected vehicles will stay at the forefront of the automotive industry. Growing manufacturers have come to the conclusion that C-V2X is superior to DSRC IEEE802.11p technology. Global shifting is towards C-V2X as C-V2X is superior in a sense it enables not only V2V services, but in addition V2N services are also possible.

## **Acknowledgements**

The author is thankful to Prof. Dinkar Prasad, Head, EE Dept., and Associate Dean, and Prof. Sandeep Sen, Dean of School of Engineering, Shiv Nadar University, G. Noida (UP) for his encouragement, and permission to publish this paper.

## **Author details**

Prem Chand Jain  
EE Department, School of Engineering, Shiv Nadar University,  
Greater Noida, UP, India

\*Address all correspondence to: [premchand.jain@snu.edu.in](mailto:premchand.jain@snu.edu.in)

## **IntechOpen**

---

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

## References

- [1] A.F. Molisch, F. Tufvesson, J. Karedal, C.F. Mecklenbräuker, "A survey on vehicle-to-vehicle propagation channels", *Proceedings of the IEEE Wireless Communications*, vol. 16, no. 6, 2009, pp 12-22.
- [2] S. Biswas, R. Tatchikou, F. Dion, "Vehicle-to-vehicle wireless communication protocols for enhancing highway traffic safety", *IEEE Communications Magazine*, vol. 44, no. 1, 2006, pp 74-82.
- [3] B. Hu, H. Gharavi, "A joint vehicle-vehicle and vehicle-roadside communication for highway traffic safety", *Hindawi Publishing Corporation, International Journal of Vehicular Technology*, 2011
- [4] ---, "Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications", *IEEE Standard for Information technology- Telecommunications and information exchange between systems-Local and metropolitan area networks- Specific requirements-Part 11, IEEE Std. 802.11-2012*, 6 Feb., 2012
- [5] Y.J. Li, "An overview of the DSRC/WAVE technology", *NICTA, NSW, Australia Report*, 2015, pp 1-15.
- [6] ---, "Initial cellular V2X standard completed", *3GPP*, Sept., 2016
- [7] A. Papathanassion, A. Khoryacu, "Cellular V2X as essential enabler of superior global connected transportation services", *IEEE Future Networks Enabling 5G and Beyond*, June, 2017
- [8] ---, "The case for cellular V2X for safety and cooperative driving-5GAA", *5GAA*, 23 Nov., 2016
- [9] ---, "LTE-M: Optimizing LTE for Internet of things", *Nokia Network White Paper*, 2015, pp 1-16.
- [10] ---, "NB-IoT deployment guide to basic feature set requirements", *GSMA, White Paper*, 02 August 2017.
- [11] D. Porter, "100Base-T1 Ethernet: The evolution of automotive networking", *Texas Instruments White Paper*, April 2018, pp 1-10.
- [12] ---, *IEEE standard for Ethernet amendment 1: Physical layer specifications and management parameters for 100Mbps operation over a single balanced twisted pair cable (100Base-T1)*, *IEEE Standard 802.3bw*, 2015, pp 39-63.
- [13] ---, "SAE Standards", *SAE International*, March, 2017



*Edited by Marian Găiceanu*

This book examines the development and technical progress of self-driving vehicles in the context of the Vision Zero project from the European Union, which aims to eliminate highway system fatalities and serious accidents by 2050. It presents the concept of Autonomous Driving (AD) and discusses its applications in transportation, logistics, space, agriculture, and industrial and home automation.

Published in London, UK

© 2021 IntechOpen

© metamorworks / iStock

**IntechOpen**

ISSN 2633-1403

ISBN 978-1-83881-076-4



9 781838 810764

