THESIS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY IN MACHINE AND VEHICLE SYSTEMS

Bionic navigation

From emergent retinal optic flow to locomotor ballistic corrections in humans, for robotics

BJÖRNBORG NGUYEN

Department of Mechanics and Maritime Sciences
Division of Vehicle Engineering and Autonomous Systems
CHALMERS UNIVERSITY OF TECHNOLOGY

Göteborg, Sweden 2025

Bionic navigation
From emergent retinal optic flow to locomotor ballistic corrections in humans, for robotics
BJÖRNBORG NGUYEN
ISBN 978-91-8103-110-2

Doktorsavhandlingar vid Chalmers tekniska högskola
Ny serie nr. 5568
ISSN 0346-718X
Department of Mechanics and Maritime Sciences
Division of Vehicle Engineering and Autonomous Systems
Chalmers University of Technology
SE-412 96 Göteborg
Sweden
Telephone: +46 (0)31-772 1000

Cover:
Three image panels illustrating retinal optic flow fields, what a human may visually perceive during a right curve bend in virtual reality.

Chalmers digitaltryck
Göteborg, Sweden 2025

Bionic navigation
From emergent retinal optic flow to locomotor ballistic corrections in humans, for robotics

BJÖRNBORG NGUYEN
Department of Mechanics and Maritime Sciences
Division of Vehicle Engineering and Autonomous Systems
Chalmers University of Technology

Abstract

This thesis explores the development of bio-inspired methods for addressing complex navigation tasks in cyber-physical robotic systems. In human locomotion, a sensation of a visual flow cue is created by continuously registering moving visual features on the human retina. An interpretation of visual flow cues forms a low-level motion perception more commonly known as *retinal optic flow*. It is often mentioned and credited in recent human locomotor research, but is limited as a theory or concept. A computational method of reconstructing retinal optic flow fields in humans is introduced and studied along with its effect on *intermittent control* to initiate informed *ballistic corrections*. Unifying the theories of human visual perception, intermittent control, muscular control, and coordination, the analysis from experiments reveals that human response time is approximately 0.14 s.

To achieve retinal optic flow reconstruction for robotics, a set of optic flow estimators is fairly and systematically evaluated on the run-time performance, reliability, and accuracy criteria. This work developed a formalized methodology utilizing a microservice paradigm and containerization technology to perform benchmarking and generate results. Doing so enables continuous integration, continuous deployment, and continuous experimentation, which are beneficial for research and development. Furthermore, this thesis found that the readiness of vehicles for adopting modern robotic software, with special emphasis on real-time computing, has matured and essentially turned vehicles into mobile data centers and capable robots. To show this partially, two optic flow-based local localization methods are demonstrated for marine vehicles in a littoral environment.

Keywords: Bionics, retinal optic flow, active gazing, active fixation, smooth pursuit, visual perception, optic flow, locomotor control, neuromuscular behavior, human locomotor behavior, robotic navigation

Sammanfattning

Människans förmåga att kunna navigera i sin dynamiska omvärld är ett komplext ämne med många involverade processer från synintryck till rörelse. Med inspiration från människan kan man lösa liknande utmaningar inom robotiken, för navigering. När en människa befinner sig i rörelse skapas ett optiskt flödes-intryck på ögats retina, vilket kallas för retinalt optiskt flöde. Det har sedan länge funnits en hypotes om att människan nyttjar retinalt optiskt flöde för att lokalt navigera sig själv med kort tidshorisont och i hög hastighet, t.ex. vid vardaglig bilkörning. Detta tros ha gett upphov till varför man i körskolor intuitivt lär ut "titta dit man vill"-strategin.

Genom att internt uppskatta behovet av att agera via externa stimuli, såsom retinalt optiskt flöde, avfyras ballistiska styrkorrektioner för att tillfälligt minska behovet. Hastiga koordinerade rörelsekorrektioner i extremiteter, även kallat *reaching*-rörelser, karakteriseras rörelseavståndet som en logistik funktion (S-kurva) och dess hastighet som en klockformad kurva (Gauss-kurva). I detta arbete har det visat sig att dessa korrektioner inte kan avbrytas abrupt, men istället kan överlagras för att förstärka eller dämpa den befintliga rörelsen. På grund av sin grundläggande karaktär, uppvisas även dessa rörelser vid styrning av exempelvis ett fordon (via ratten) då människor utför koordinerade rörelsekorrektioner. I detta arbete, unifierades av flera teorier inom mänsklig perception, rörelsemönster och beteende, framkommer genom experiment att mänsklig responstid är ca 0,14 s mätt från synintryck till initiering av rörelse.

Arbetet undersöker även hur dessa forskningsresultat kan överföras till robotiken, någon som visar sig möjligt. För att kunna effektivt återskapa retinalt optiskt flöde utvärderades ett urval av lämpliga beräkningsalgoritmer på ett systematiskt och rättvist sätt. Utvärderingskriterier inkluderade exekveringstid, pålitlighet och noggrannhet. En formaliserad metod, baseras på mikrotjänstarkitektur, arbetades fram för att kunna uppnå jämförbara prestandaresultat. En positiv biprodukt är kontinuerlig integration, kontinuerlig utrullning och kontinuerligt experimenterande vilket är önskvärt vid mjukvaruutveckling för forskning. Slutligen, fann detta arbete att mognadsgraden hos farkost för realtidsberäkning för moderna mjukvaruutveckling är tillräcklig vilket evolverat dem till mobila datacenter och kraftfulla robotar. För att delvis påvisa detta, koncepttestades och demonstrerades två optiskt flödes-baserade lokaliseringsalgoritmer för fartyg i kustnära vatten.

"My time has come. You must continue your journey — without me.
You must believe."

– Oogway

Acknowledgements

I naively embarked on this journey without fully understanding what path I had started walking on or what it really entailed. The journey turned out to be quite lengthy and arduous, accompanied by doubts and countless setbacks. However, with the support I received, I made it. Here I am.

Firstly, I would gratefully thank my supervisors, Prof. Ola Benderius and Prof. Christian Berger, for their incredible support in my journey and for shaping me into the researcher I have become. It also needs to be acknowledged that, during my moments of despair, Ola Benderius never fails to shed and instill in me the needed optimism and crazy perspectives to overcome obstacles. Christian Berger never ceases to amaze me when providing me with technical assistance. Furthermore, I extend my gratitude to my examiner, Prof. Mattias Wahde, for accepting me into his research group, which has dramatically shaped my academic path for the better.

My work would not have been possible without the research laboratory Revere. I am incredibly grateful for the resources and opportunities I received from Dr. Fredrik Von Corswant, Arpit Karsolia, and Daniel Poveda.

My direct peers, Krister Blanch and Tarun Kadri Sathiyan, must be thanked for making my work and journey substantially endurable and enjoyable, not to mention inheriting my technical debt. It has been a bliss to see you stepping and striding on my explored beaten paths, and your future will be interesting to follow once I finally leave. I also want to extend my thanks to my past project assistants, Liv Johansson, for her contributions to my research project and for keeping me company while implementing my research vision, and Mateo Vega, for making my experiences more eventful and memorable (not to mention the gout incident).

I want to acknowledge all former, current, and future colleagues at VEAS for contributing to a friendly, inviting, and ever-relaxing work environment. Thanks to you guys, the fikas and the heated discussions about random subjects were always something to look forward to.

A *special* thanks goes to the PhD student protector and hero, Sonja Laakso Gustafsson, for standing up and raising a voice for all of us students here at VEAS. Her never-ending essential support enables us, PhD students, to do what we must.

Finally, but not least, I want to thank my sweetheart, significant other Yao Wang, for enabling me to end my PhD journey in the best possible way. Not only did I end up with a thesis, but I also ended up with you. Let us embark on new journeys to come.

List of included papers

This thesis consists of an extended summary from the following papers:

Paper A B. Nguyen and O. Benderius. Intermittent control and retinal optic flow when maintaining a curvilinear path. *Scientific reports* **15**.1 (2025), 18926. DOI: 10.1038/s41598-025-02402-3. URL: https://doi.org/10.1038/s41598-025-02402-3

Paper B B. Nguyen et al. Application and evaluation of direct sparse visual odometry in marine vessels. *IFAC-PapersOnLine* **55**.31 (2022). 14th IFAC Conference on Control Applications in Marine Systems, Robotics, and Vehicles CAMS 2022, 235–242. ISSN: 2405-8963. DOI: 10.1016/j.ifacol.2022.10.437. URL: https://www.sciencedirect.com/science/article/pii/S2405896322024831

Paper C B. Nguyen, C. Berger, and O. Benderius. "Systematic benchmarking for reproducibility of computer vision algorithms for real-time systems: The example of optic flow estimation". *2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. 2019, pp. 5264–5269. DOI: 10.1109/IROS40897.2019.8968066

Paper D C. Berger, B. Nguyen, and O. Benderius. "Containerized Development and Microservices for Self-Driving Vehicles: Experiences & Best Practices". *2017 IEEE International Conference on Software Architecture Workshops (ICSAW)*. 2017, pp. 7–12. DOI: 10.1109/ICSAW.2017.56

Paper E K. Blanch, B. Nguyen, and O. Benderius. Topographic flow-based odometry. *Submitted to IEEE Journal of Oceanic Engineering* (2025)

Table of contents

Abstract i

Sammanfattning ii

Acknowledgements v

List of included papers vii

Table of contents ix

I Extended summary 1

1 Introduction 3
1.1 Understanding human locomotor behavior . . . 4
1.2 Implementing bionic navigation . . . 6
1.3 Research questions . . . 7
1.4 Limitations . . . 9
1.5 Author contributions . . . 9
1.6 Thesis outline . . . 10

2 Visual motion perception 11
2.1 Optic flow . . . 13
2.2 Retinal optic flow . . . 16
2.2.1 Retinal image stabilization and gazing behaviors . . . 16
2.2.2 Perception for curvilinear locomotor control . . . 20
2.2.3 Smooth pursuit during curvilinear locomotion . . . 24

3 Neuromuscular behavior 27
3.1 Motor control and muscle coordination . . . 27
3.2 Modeling human reaching movements . . . 30
3.3 Sufficient response through intermittent control . . . 33

4 Robotics for bionic navigation 37
4.1 Vision sensors . . . 37
4.2 Robotic software . . . 39
4.2.1 Microservice design, real-time computing, and embedded system deadlines . . . 40
4.3 Optic flow estimation methods . . . 42
4.3.1 Machine learning approach . . . 44

4.3.2 Benchmarking performance and evaluation 45
4.4 Optic flow-based odometry 47
4.5 Retinal optic flow estimation 49
4.6 Identifying ballistic reaching corrections 51
4.6.1 Particle swarm optimization 53

5 Discussion **61**
5.1 Retinal optic flow for locomotor control 61
5.1.1 On the assumption of exact smooth pursuits 63
5.1.2 Gaze placement on ground 67
5.2 Maintaining steering in curvilinear paths 69
5.2.1 Heavy tail property in reaching velocity profiles 71
5.3 Bionic navigation in robotics 73

6 Concluding remarks and future work **79**
6.1 Conclusions 79
6.2 Future work 81

7 Summary of included papers **83**
7.1 Paper A 83
7.2 Paper B 84
7.3 Paper C 84
7.4 Paper D 85
7.5 Paper E 85

Bibliography **87**

II Appended papers A–E **101**

Paper A **103**

Paper B **127**

Paper C **137**

Paper D **145**

Paper E **153**

Part I

Extended summary

Chapter 1

Introduction

Robotic systems are already an integral part of human society and are expected to become even more integrated, interactive, and interconnected in the near future. This imposes implicit requirements on them to have navigation capabilities in their respective contexts and robotic applications. Furthermore, this development will be progressive and gradually rolled out, creating a heterogeneous mix of coexisting robots and humans in society. Thus, designing a robotic navigation system is highly relevant in research and development in academia and industry alike to make it simultaneously performant, safe, and *human-acceptable* [112]. Already today, many modern robotic systems are capable of solving and navigating in a static environment using perfect knowledge or assumptions of their surroundings. However, general navigation with uncertainties is difficult for robotics to solve; therefore, new theories, methodologies, and further innovation are still relevant and needed.

Looking to biologically inspired engineering, or *bionics*, may provide an alternative way of solving the navigation task. Biology and evolution have been finding solutions to non-trivial tasks or problems since the inception of life. For example, it is not by coincidence that the shape of a modern-engineered aircraft has a striking resemblance to that of birds or echolocation technology to echolocating bats. Therefore, finding inspiration and studying biological solutions to navigation can be exploited and transferred to robotics and engineering to improve the performance whilst increasing the acceptability and adoption of technology.

Perhaps unsurprisingly, biological creatures can consciously navigate in order to search for food, find potential partners, or avoid dangerous situations. Among simpler microscopic lifeforms such as plankton, navigation without a set target or higher planning and simple surroundings awareness is sufficient for sustaining

life [119]. However, for larger animals, skillful target-based navigation capabilities are required or even essential e.g. for energy conservation or accomplishing predation (or *escaping* predation for that matter). Despite the seemingly simple task of moving towards a target, known as *pursuit*, it is not fully understood in detail from perception to action response due to the neurological complexity and the resulting emergent behavior.

1.1 Understanding human locomotor behavior

In the past decades, there has been an ongoing debate in the scientific community on how humans navigate and how the demonstrated behavior can be described and further modeled. This can range from *what* and *how* visual cues are used to what control strategies humans may employ for locomotor control. Furthermore, this includes all modes of transportation such as walking, driving, or flying, as humans seem to deploy slightly different control and perception behavior [83, 129, 75, 50]. In high-speed smooth curvilinear motion in a textured plane, e.g. driving, an established and identified heuristic strategy has emerged *"you should look where you are going,"* which loosely describes how humans may exploit the emerging patterns of *retinal optic flow fields* [138, 64].

Retinal optic flow has been identified and acknowledged as an integral part of perceiving relative motion and, in turn, used for ego-localization and motion control [143, 138, 137, 64]. Much of the previous work related to retinal optic flow has been done in an interdisciplinary fashion between experimental psychology, behavioral science, neuroscience, neurocognition, applied computer vision, and mathematics. It has been carried out on a qualitative level and within theoretical frameworks. Recent advances in eye-tracking technology have enabled further research in human locomotor control, complementing theoretical frameworks with high-quality experimental data. This has resulted in discoveries that simultaneously challenge previous hypotheses and archaic understanding of human behavior.

When humans intently gaze, or *look*, at objects, the formation of retinal images in the fovea of the eyes must be stable to increase the clarity and acuity of the visual details. This is achieved through retinal image stabilization through various developed biomechanical reflexes and processes. One such mechanism is the *smooth pursuit*, where humans actively fixate their gaze on a visual feature using fine, controlled, and accurate eye movements to track the visual target. Another is *vestibulo-ocular reflex*, which compensates for chaotic head movements through vestibular sensing organs, resulting in counteracting expressed eye movements. However, beyond solely retinal image stabilization, it is believed that skillful active fixation further aids human locomotor control in curvilinear motion [138, 64] and also on foot [83]. The hypothesis is that stabilized retinal optic flow is used for locomotor control [82, 84] and partly anticipatory trajectory planning [37, 83, 74].

Kim and Turvey [64]; Wann, Land, Swapp, and Wilkie [137, 138, 146]; and Lee and Cheng [78] and others advocate that retinal optical flow cues are the primary sources of locomotion. The latter even further questions the need to retrieve *extra-retinal* information such as heading [77, 137], implying retinal optical flow cues are sufficient for local navigation. Regardless, it is argued that the human steering control strategy mainly revolves around using the retinal optical flow field patterns by continuously *nulling flow curvature* (or controlling *locomotor flow lines*) through proper active fixation, combined with applying steering maneuvers [138, 73, 64]. By intelligently or skillfully placing the gazing point by simply *"looking where you want to go,"* the retinal optic flow of the intended path will be vertical at the correct instantaneous motion. This simple proposed heuristic online control is shown to be mathematically plausible using the retinal optic flow field to correctly make the judgment for curvilinear motion control [138, 73, 64].

Intermittent control has been found to better explain human muscular and limb control than optimal control or classical engineering control approaches. The control phenomenon is characterized by applying discrete ballistic corrections of given strength and duration, and may even be *superposed* (overlapping corrections). Both Craik and Tustin first independently introduced this in 1947 when they studied humans as an engineered control system and noticed the similarities to servo control applying intermittent corrections [29, 130]. Extending this work, Beggs and Howard, Georgopoulos et al. realized that these ballistic corrections had a sigmoid curve (logistic shaped) and by extension a bell-shaped velocity profile when studying the rapid limb movements in humans [8] and monkeys [48]. This phenomenon is known as *reaching* in the fields of neuroscience [30, 53, 31, 51, 15, 14, 104]. Plamondon derived a mathematical description of the bell-shaped velocity profiles using a stochastic model, further accounting for the lagging heavy tail often observed in experimental data, and provided a plausible explanation of the origin of the bell-shaped velocity profile for this [104]. Although the early introduction in human intermittent control research, it has only been properly considered and gained traction in recent years [81]. Furthermore, Benderius and Markkula, and Smith et al. found that human-operated controls, such as steering wheels, exhibit the same fundamental properties of intermittency and movement characteristics as reaching in human limbs [120, 11].

To summarize, there is a scientific consensus that when humans maintain locomotor control, the retinal optic flow is vital as a perception cue. However, to what extent it is used or how exactly it relates to humans intermittently applying locomotor ballistic corrections to maintain locomotion is still unknown. Previous work has proposed a simplified heuristic, "look where you are going" and "going where you look" to describe human locomotor behavior to correctly maintain their trajectory on a textured plane using proper active gaze control and emerging retinal optic flow.

1.2 Implementing bionic navigation

The topic of *optic flow* is a well-established domain in computer vision and image motion analysis. Optic flow estimation has proved useful in various applications such as video compression, motion analysis, image stabilization, motion detection, simultaneous localization and mapping (SLAM), and odometry. Retinal optic flow may be regarded as an extension of, but not *confused* with, optic flow. The main difference between the two is that the retinal optic flow captures the biomechanical dynamics of the complex movements of the head and the eyes into the perception of motion. This facilitates retinal image stabilization by computationally displacing the vector field to the biomechanical movements. Thus, retinal optic flow fields may be computationally reconstructed using gaze data via eye-tracking systems and optic flow fields in head-fixed video data.

Incorporating human perception and cognitive features in human locomotor control research is often heavily simplified or deliberately left out, e.g. human driver modeling [113, 124]. Understandably, many researchers make conscious simplifications to keep the model complexity low, or due to a lack of relevant high-quality data. However, doing so may result in inaccurate representation and loss of explainability and interpretability of the results, further losing its transferability and generalization. Retinal optic flow is understood and acknowledged as essential for human locomotor control and is thus often mentioned in modeling contexts. However, from a computational and engineering standpoint, the use of visual motion flow is frequently not fully described with sufficient detail or incorporated due to the lack of research and data, and further, because most prior research is carried out in experimental psychology. For example, Matthis and colleagues have investigated how to reconstruct retinal optic flow fields computationally, the emerging *sparse* visual flow lines, and further their potential use for human bipedal locomotor control [82, 84].

Further, the control behavior manifested by humans differs from the traditional or classical engineering control paradigms such as the *PID* controller (proportional–integral–derivative). Despite that, human locomotor control is commonly modeled with a PID control-loop scheme in research and engineering [113, 128]. The main reason is often simple; it is possible to do so in order to achieve satisfying results. Researchers then deduce how it could be useful and attempt to further explain the behavior and its implications. Contrary to this, this work incorporates the mathematical framework developed by Plamondon [104], and considers the sensorimotor models developed by Gawthrop et al. [43] and Markkula et al. [81], to produce the intermittency property in the control system. Thus, it is possible to emulate the human-like correction control to trigger a ballistic correction using an error accumulation, or *evidence accumulation*, avoiding the PID approach altogether.

Regarding implementing perception and control in a sensorimotor fashion, robotics software applications are typically monolithic in design. It is a *single* executable binary handling all functionality. This design has significant drawbacks, including code complexity, run-time robustness, reusability, software deployment, scalability, and flexibility. Monolithic software designs in robotics research and engineering pose a significant challenge since robotic platforms may have varying configurations of sensors, motors, data processing, and applications, especially in mission-critical tasks. Modern software practices often deploy a microservice design, mitigating these challenges, where each single executable application binary handles one specific designated task, such as querying information from a central database or visualizing video data. The appended papers explicitly explore the utility and feasibility of adopting and deploying a microservice design paradigm for research and robotics. These methodologies are widely adopted and used throughout this thesis.

To summarize, various numerical methods of estimating optic flow already exist for motion perception, which may be used to reconstruct retinal optic flow by incorporating human gazing components. This motion perception cue may be used in a computational framework, a sensorimotor control model with evidence accumulation, to trigger a ballistic correction mimicking human intermittent control behavior. Bionic implementation could be enabled and supported by adopting modern software development methods, for example, containerization and microservice design paradigms.

1.3 Research questions

This thesis aims to investigate and understand how humans visually perceive the environment to locomote and further apply this newly gained knowledge to solve navigation tasks in robotics. The retinal optic flow has been identified as an essential source of information when studying human locomotor behavior. Previous research has proposed that retinal optic flow is used during locomotion, and much of the current academic debate is on the details of *how* humans exploit it. Various methods have been developed for estimating optic flow within computer vision, especially with the introduction of deep learning. It slightly differs from the retinal counterpart as it lacks human gazing behavior and how motion perception may be exploited for locomotor control.

Furthermore, vehicles on roads, in air, on water, and under the water surface are continuously tasked with more complex navigation challenges. This has pushed traditional vehicles to be equipped with various sensors and advanced computational capabilities, effectively evolving them into highly mobile data centers and robots. Multiple aspects of bionic navigation are investigated in this thesis with the following posed research questions (**RQs**):

RQ1 How do humans locally navigate by looking where they are going?

As discussed in the previous section, the retinal optic flow field is credited as a quintessential part of human navigation in terms of localization and locomotor control tasks. Previous research has proposed a simple heuristic for human locomotor control using immediately available visual cues. However, the proposal does not consider human intermittent control and provides insufficient details on implementing perception and control that mimic human behavior. Furthermore, experienced and skilled drivers traveling on textured surfaces often demonstrate similar behavior of actively fixating their gaze in the direction they intend to go, implying there is a value in doing so.

RQ2 How can retinal optic flow be used as a perceptual input to maintain locomotor control?

Related to **RQ1**, early research has investigated and characterized human sensorimotor behavior. However, there is a lack of understanding of how humans intermittently, not continuously, maintain their locomotion using retinal optic flow fields. In particular, how can the locomotor control relate to the perceived retinal optic flow field using modern computer vision? Thus, understanding and systematically describing human sensorimotor behavior enables robotics to mimic human locomotor behavior.

RQ3 How can optic flow be used for odometry in various settings?

A common robotics approach is using a global navigation satellite system (GNSS) or similar technology for self-localization. GNSS technology is even more accurate and precise when complemented with real-time kinematic positioning (RTK). However, alternative or complementary self-localization methods are needed when GNSS is neither suitable nor applicable. Using visual or topographic data may provide sufficient information to estimate self-localization as a dead-reckoning or stand-in replacement when GNSS technology becomes unavailable.

RQ4 How can bio-inspired methods of perception and control be accommodated and implemented in robotics for mission-critical navigation?

This research question partly concerns how one can develop software that is agnostic to underlying hardware architecture and application context, and partly how to do so with regard to hard deadlines. With vehicles ever-evolving and becoming more computationally capable, it is interesting to investigate if they are suitable for adopting modern software development paradigms such as containerization or microservice design. Modern software development practices enable software modularity to develop, compare, and evaluate bio-inspired methodologies more easily.

1.4 Limitations

In this thesis, only local and short-term time horizon navigation has been considered, leaving out strategic long-term route planning and execution thereof. Local and short-term time horizon navigation here refers to the traversing ability of the agent in the immediate environment, for example, lane keeping during driving. Strategic long-term planning encompasses the mission goal of knowing the ends of the journey and how to accomplish and adapt the trajectory involving higher cognitive function. Long-term time horizon navigation falls outside the scope and resources of this work.

Artificial neural networks, particularly *deep learning*, have significantly impacted many research fields, including human behavioral research, thanks to their adaptive capabilities and performance on complex tasks. Due to the very nature of deep neural networks, it is by design a complex end-to-end system, i.e. a black box model. As a result, it lacks *any* meaningful interpretation of its inner workings. In substantial cases where it is applied, the resulting utility is often prioritized over the fact *how* it managed to solve the problem. As a reaction to this, a movement emerged urging to stop blindly using black box models for high-stakes decisions and use *interpretable* artificial intelligence instead, to increase transparency and accountability [111, 136]. Hence, for this thesis, deep learning will not be fully considered. However, its contributions to research and engineering are acknowledged.

1.5 Author contributions

The division of work for the included papers of this thesis is detailed below.

Paper A

Benderius and Nguyen conceived the experiment. Nguyen prepared, constructed, and implemented the experimental hardware and software setup. Nguyen conducted the experiment, which included recruiting and informing the research subjects. Nguyen procured, curated, and maintained the data sets [93, 91]. Nguyen conducted the formal analysis. Nguyen produced and visualized the results. Nguyen and Benderius analyzed and interpreted the results. Nguyen drafted the manuscript, and Benderius reviewed the manuscript and supervised the research project.

Paper B

Petersson implemented the initial work as their master's thesis under the supervision of Blanch. Blanch procured the data sets. Nguyen finalized the software implementation. Nguyen post-processed the data and generated the results. Blanch,

Nguyen, and Petersson created a draft of the manuscript. Blanch and Nguyen finalized the manuscript under the supervision of Benderius and Berger.

Paper C

Nguyen conceived the research idea, prepared the experimental setup, formalized and conducted the data collection, and procured the research results. Nguyen drafted the initial manuscript. Nguyen finalized it under the supervision of Benderius and Berger.

Paper D

Berger initiated the idea to formalize and publish the findings from recent research activities. Nguyen drafted the manuscript under the supervision of Berger and Benderius, and Berger and Benderius finalized it.

Paper E

Blanch conceived the research idea, prepared the experimental setup, and procured the data sets. Nguyen formalized the theory, implementation, and computational framework. Nguyen contributed to the analysis and interpretation of results and discussions. Blanch drafted and finalized the manuscript under the supervision of Benderius.

1.6 Thesis outline

Chapters 2 and 3 provide the biological background and insight into how humans locally navigate and their locomotor behavior, from visual motion perception to muscular response. Chapter 4 concerns robotics, visual sensors suited or adapted for optic flow, computational implementation and evaluation of optic flow, software architecture, software design patterns, and real-time computing constraints. Then, Chapter 5 discusses research findings in the thesis and their implications in the scientific community. Lastly, final concluding remarks and future works from this thesis are given in Chapter 6 followed by the summaries of the appended papers in Chapter 7.

Chapter 2

Visual motion perception

There is a famous thought experiment from the 17th century in philosophy called *Molyneux's problem* which poses an interesting question on the topics of *sensation* and *perception*. Suppose a congenitally blind person can feel the differences between geometric shapes via tactile touch. Could that very same person also distinguish and identify between the very same objects by only visually looking, if they were given the ability to see? This particular line of thinking emphasizes the fundamental differences between basic sensation and the higher cognitive perception of the same object. Held et al. later investigated this question via experimental trials by letting newly sighted patients attempt to pair haptically touched objects with seen ones [56]. The answer to the thought experiment was found to be simply *no* since the newly sighted could *not* establish the perception of the same object via different sensation modalities. However, the researchers quickly noted that they were learning and developing the cross-modal mapping after the trial. This confirms that perception has a significantly complex learning aspect compared to straightforward sensation.

Visual perception is thus a complex field to study, as prior experience has a non-trivial impact on perception capabilities. There is evidently a difference between visual sensing and visually perceiving something, where the latter requires a higher cognitive function to interpret the information. Moreover, various types of visual perception, such as the recognition of patterns, edges, objects, and motion, form an intricate and complex cognitive system that constructs the mental or internal representation of the environment. This creates an *awareness* which some researchers argue is an important, if not the most significant step, towards *consciousness* as we understand it [39].

On the other hand, the sensation does not arguably require higher cognitive

function but instead operates on a lower cognitive level with minimal interpretation. One example is when a photon hits the photoreceptors on the retina, converting the light into an electric action potential — a process called *signal transduction*. The signal transduction process, happening in the retina of animals, continuously conveys information and provides minimal interpretation as input to the neural perception system. The signal will propagate throughout clusters of neural networks for various purposes like knowledge and memory formation, recognition of patterns and objects, and awareness. These biological processes occur without voluntary introspection, and the formations are automated to a certain degree. This thesis proposes that low-level vision includes similar automatic properties to the high-level perception for the purpose of navigation. This could specifically relate to the phenomenon where one produces an initial response before the cognitive action planning is fully complete.

The visual motion flow emerges from the relative movement of the observer and its visual surroundings. Subconsciously registering visual features such as light, shape, color, or patterns in the surrounding scenery creates the projected optic flow on the retina, *retinal optic flow*. It is apparent that such a stimulus provides crucial information for the observer about the ever-dynamic and changing environment, and further their locomotion. One indication of its importance can be seen in the documented condition where the ability to process visual motion is impaired, called *akinetopsia*. It has been reported [28, 2, 59] that the affected have significant challenges in perceiving the moving environment and anecdotally describe it as viewing the world in a series of still images. To compensate for the inability to perceive motion visually, some have reportedly relied on other sensations to judge motion, for example, relying on hearing to judge approaching vehicles. Interestingly, those affected by the impairment can still perceive the surrounding scene in a way many would consider perfect scene awareness. Despite that, the absence of properly functioning vision motion perception negatively affects their daily lives.

Thus, it may be argued that visual motion perception may at least have a distinct pathway in our cognitive system, a vital function in our daily lives. When studying visual motion perception, it is intuitive and natural to consider the relative visual movement input cue, generally referred to as an optic flow field. The optic flow field may be regarded as a three-dimensional visual motion vector field at every visual point, projected to a two-dimensional space. This does not consider the higher cognitive or psychological functions of the observer, such as gazing, stabilization of the retinal image through biomechanical reflexes, or the tendency to focus the gaze on visual details. Each of the phenomena mentioned above is important to visual motion perception. Humans can quickly change the direction of their eyes, turn their heads, and intelligently guide their gaze through the scene using knowledge-driven gaze control, as described by Henderson [58].

To summarize, human visual motion perception is crucial for cognitively recognizing and understanding the surrounding dynamic world and moving objects.

Cognitive motion perception involves both low-level and high-level information processing and is aided by retinal image stabilization, which is achieved by various biomechanical reflexive processes. Retinal image stabilization is believed to play a quintessential role in locomotor control and is highly interconnected with knowledge-driven gaze control and intermittent control in humans.

2.1 Optic flow

In the 1950s, Gibson published their influential work [50] where the optic flow concept was formally conceived in the academic world. The author gave the example of avian pilots guiding their aircraft to the runways during level flight using optic flow as a primary source of information for locomotor control. In addition to the theories of optic flow, the resulting phenomenon *focus of expansion* (FOE), also known as *vanishing point* and *point of aim*, was introduced and developed, and its hypothetical use was discussed. The focus of expansion is the point of divergence in the optic flow field, analogous to the divergent source point in vector field analysis. Early interpretation and purpose of FOE were that the direction of locomotion, or heading, could be detected and used for locomotor control. A large part of the early research focused on the optic flow field in general and the perception of heading. Optic flow and retinal optic flow have unfortunately been used interchangeably since inception. In contrast, the latter considers biomechanical gazing as mentioned previously, resulting in the ambiguity of the term optic flow in psychology and computer vision.

In the development and explanation of the retinal optic flow phenomenon, Prazdny and Longuet-Higgins made significant contributions in modeling and interpreting moving retinal images in a theoretical mathematical framework [79]. Their work describes a projected image plane of the world scene texture, similar to the pinhole camera model and the anatomy of the human eye, see Fig. 2.2. The emergent retinal optic flow can be mathematically described by considering a relative motion between the observer and the texture. The optic flow may then be modeled as

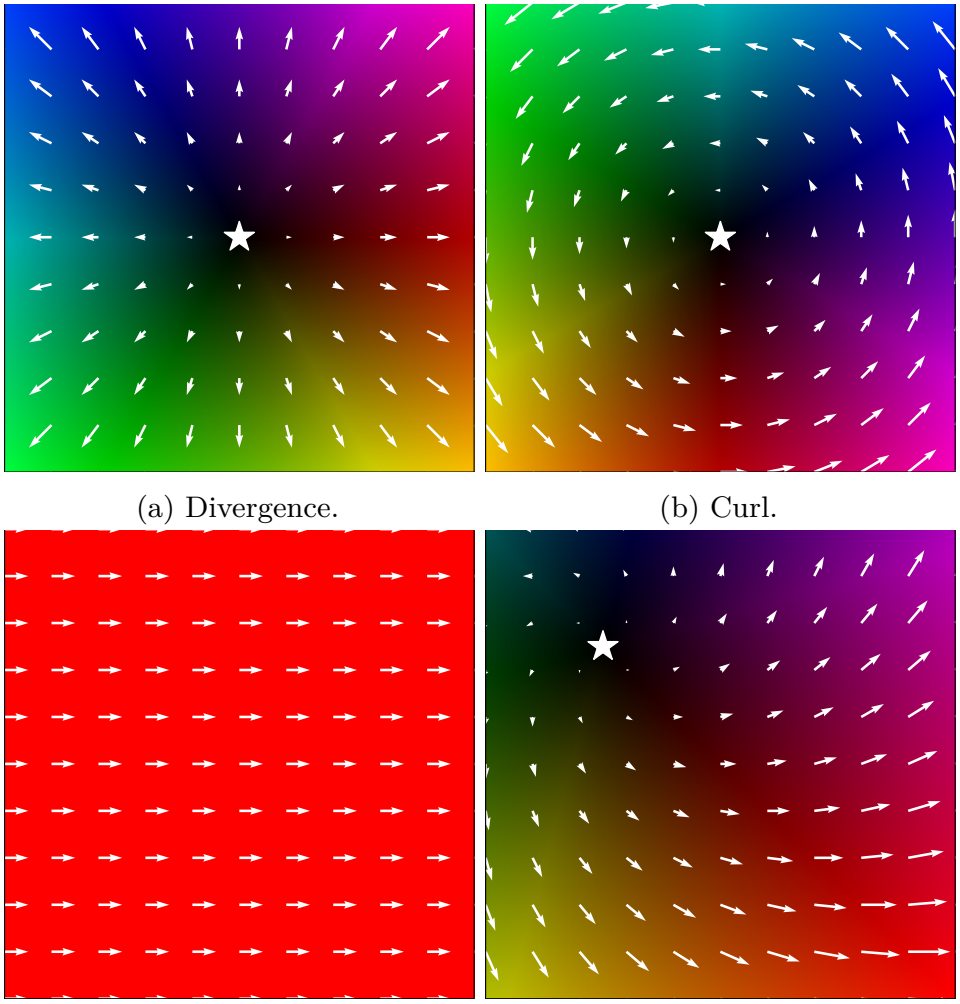
$$\begin{aligned}\vec{Q}(y,z) = \begin{pmatrix}\dot{y}\\ \dot{z}\end{pmatrix} = &\underbrace{\frac{1}{X}\begin{pmatrix} y & -f & 0\\ z & 0 & -f\end{pmatrix}\begin{pmatrix}v_x\\ v_y\\ v_z\end{pmatrix}}_{\text{translational component}} + \\ &\underbrace{+\frac{1}{f}\begin{pmatrix} fz & yz & -(f^2+y^2)\\ -fy & (f^2+z^2) & -yz\end{pmatrix}\begin{pmatrix}\omega_x\\ \omega_y\\ \omega_z\end{pmatrix}}_{\text{rotational component}}\end{aligned} \tag{2.1}$$

where y, z represent the projection pixel coordinates in the image space, $(\dot{y}, \dot{z})$ the pixel velocity i.e. the optic flow, f the focal length, $\vec{v}$ the relative velocity, $\vec{\omega}$ the relative angular velocity, and X the depth to the arbitrary visual point in world space. The optic flow described in Eq. (2.1) emerges from any relative motion of any linear combination of translation and rotation to the observer. Examples of the emergent optic flow from the different motions using Eq. (2.1) are illustrated in Fig. 2.1. Note that the scene information is required for computing optic flow in the translational component, as it depends on the X scene depth to the visual point. In contrast, the rotational component is independent of the scene geometry. This model is only suitable when perfect knowledge of the scene and the observer is fully accessible, for example, in a simulation or a fully controlled environment. However, the theoretical optic flow model would most likely not suffice with inaccurate or insufficient information in real-world settings. To cope with this, various developed numeric methods use a temporal image sequence to estimate optic flow, which will be further discussed in Section 4.3.

In contrast to the straightforward and naive manner of modeling optic flow, as shown Eq. (2.1), Prazdny made further attempts to reverse the process and extract information such as egomotion and depth from the visual optic flow alone [106]. It was concluded that the information is not conserved in the structure of optic flow and is lost in the process of projective transformation. Neither the relative depth nor the local surface orientation can be fully retrieved since they are unambiguously stored in the optic flow. Thus, if such information is to be extracted, separating optic flow into translation and rotational parts requires further assumptions, prior knowledge, cognitive experience, or complementary information.

Much of the previous research has been focused on the relative motion between the observer and the surrounding scene. However, optic flow may also emerge from independently moving objects, which Layton et al. investigated from a neurophysiological point of view [72]. By studying the optic flow, the closely related focus of expansion, and the local optic expansion of objects, so-called *looming*, they developed artificial neuro-models inspired by those found in natural biological processes. It was mainly designed to mimic the human perception of heading in the presence of independently moving objects in a parallel pathway fashion. They concluded that their models could reproduce the heading biases in the same direction and magnitude as in human experimental data.

Throughout the early research on optic flow, it has been chiefly regarded in two-dimensional image space or retinal space. However, with the advancements of complementary sensor frameworks such as *lidars* and *depth vision systems*, or even using *sensor fusion*, this could be extended to three-dimensional flow in three-dimensional space. The three-dimensional flow is referred to as *scene flow* [133], describing every visual point in the scene with an associated velocity vector. Although scene flow might appear slightly different from optic flow at first glance, its implications pose significant challenges for numerical estimation methods.



(a) Divergence. (b) Curl.

(c) Side-translation. (d) Combined.

Figure 2.1: Optic flow visualizations. Four examples of emergent optic flow fields of a perpendicular plane using two different visualization techniques: quiver and hue-saturation-value (HSV) color encoding. Quiver visualization is commonly used for *sparse* optic flow fields where the direction and the magnitude are represented with the length and direction of the arrows. It provides an intuitive interpretation utilizing a small set of discrete flow field vectors. However, the color encoding offers a continuous visualization of the *dense* optic flow field by encoding the direction to hue and the magnitude to saturation. Here, four different flow fields are illustrated: divergence caused by pure v_x, curl caused by pure ω_x, side-translation solely caused by v_y, and a full combination of the former. The location of the *focus of expansion* (FOE) is denoted with a star.

Much of the previous work for estimating optic flow is not directly applicable to corresponding scene flow, thus creating a substantial chasm between optic flow and scene flow research.

In summary, the early works of optic flow quickly identified its importance to local navigation as it provided additional tangible means of perceiving egomotion in theory. However, extracting translation and rotational movements using solely optic flow is non-trivial, as depth is lost due to visual projection. Due to the history and broad adoption of optic flow, the term has become ambiguous as optic flow more widely refers to visual motion flow fixed to the locomotor axis (or head). While retinal optic flow specifically refers to visual motion flow as perceived in the retina of the agent, it accounts for retinal stabilization of the visual image.

2.2 Retinal optic flow

As emphasized previously, retinal optic flow incorporates the gazing dynamics and head movements of the agent into the visual flow fields. A simplified illustration of the emergent retinal optic flow is shown in Fig. 2.2, demonstrating the visual dynamics between visual features and the projected motion to the human retina.

Much of the work related to optic flow may be transferred and extended to the field of retinal optic flow. However, some challenges are apparent when considering retinal optic flow. For example, how does one consider and estimate gazing and fixation dynamics to incorporate them into the existing optic flow framework? Early research has successfully demonstrated the reconstruction of sparse retinal optic flow fields by Calow and Lappe [22]. In their work, they used extensive data sets of egomotion, eye movement, and depth structure to generate what they call *true* retinal motion fields. Although this thesis does not go deep into the detailed research on a neuron-level analysis, Calow and Lappe provide and consider a hypothetical framework on how humans, on a neural level, may process retinal optic flow using an encoder for the flow. In their modeling of retinal optic flow, where polar transformation is applied, they could reproduce similar properties of real motion-sensitive neurons found in the middle temporal area of the human brain. While their method of reconstructing the flow produces accurate retinal optic flow fields, it significantly demands the sensor framework and sensor fusion techniques.

2.2.1 Retinal image stabilization and gazing behaviors

Gazing is described as the bio-mechanical coordination effort of directing the visual sensation system of the observer. Retinal image stabilization may be achieved by continuously directing and maintaining gaze fixation on a targeted visual feature. Various processes contribute to the retinal image stabilization, but active

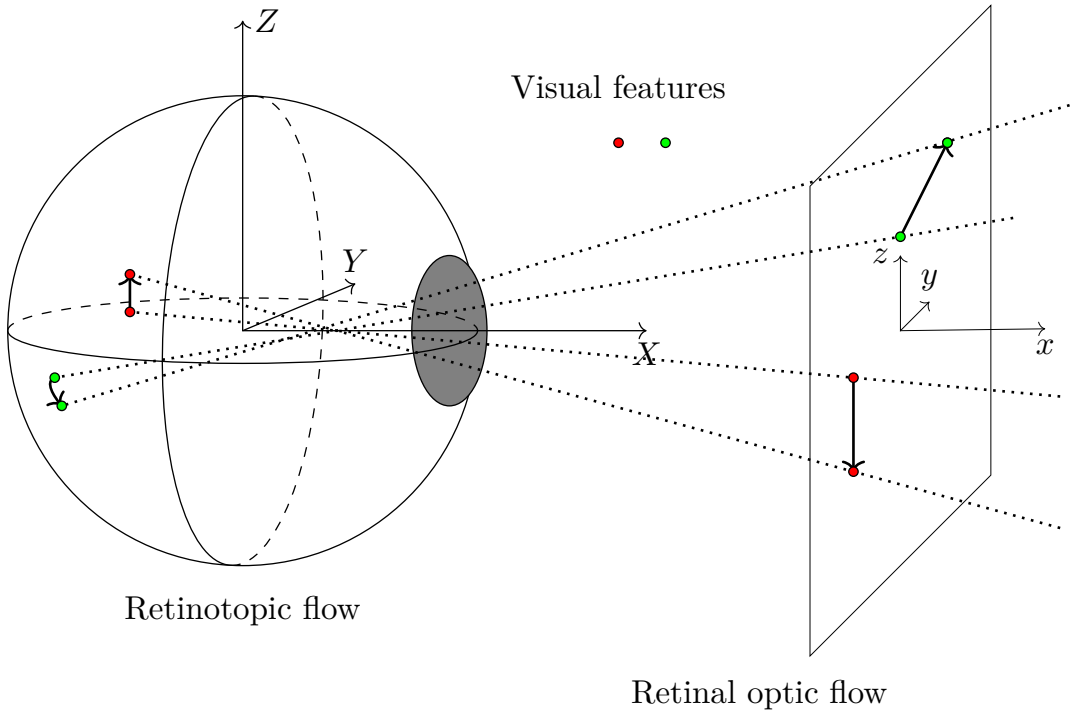


Figure 2.2: Retinal optic flow. An illustration of visual flow from motion is projected to the retina of an eye, producing the sensation of the emergent *retinal optic flow* and retinotopic flow. The main difference between retinal optic flow and retinotopic flow is the last stage of stereographic projection, i.e. projection to the sphere skewing coordinate system. By actively fixating the gaze, a stabilized retinal image at the fovea can be achieved via so-called *smooth pursuit*. The visual motion projected during such fixation differentiates the general head-fixed optic flow and retinal optic flow.

fixation, vestibulo-ocular reflexes, and cervico-ocular reflexes are mainly credited. Without retinal image stabilization, the vision would easily become blurry and misinterpreted by cognitive functions, and may result in dizziness or nausea due to cognitive dissonance, as previously discussed.

Vestibulo-ocular reflexes (VOR) correct and adjust the vision for mainly the head movements, both translational and rotational, via the vestibular sensation system located in the inner ears of humans. Anecdotally, many have probably experienced blurred vision caused by erratic eye movements known as *nystagmus* when sitting in a spinning office chair and abruptly halting the spinning; as the vestibulo-ocular reflexes are actively incorrectly correcting and compensating for the previously exposed angular velocity. The same sensation system is responsible for equilibrioception and spatial orientation to orient to gravity, identifying which direction is upwards. By sensing the fine and jerky head movements, the VOR can inhibit or exhibit signals to the eye muscles to readjust and realign the eyes for stabilization. *Cervico-ocular reflexes* (COR) work similarly but mainly use the sensation in the neck to compensate for head-body movements.

In terms of directed eye movements, there are typically five distinct types or classifications: saccades, passive fixations, active fixations, smooth pursuits, and post-saccadic oscillations. Some researchers discern active fixations from smooth pursuits as the former is characterized as low-velocity and the latter as high-velocity from some threshold. For the purpose of this thesis, these two will be considered the same and used interchangeably as it has proven challenging to distinguish between the two fully [100]. Albeit separated by a velocity threshold, they functionally achieve the same purpose of maintaining visual tracking.

A saccade is the rapid and jerky eye movement where the observer temporarily and cognitively becomes blind during the movement by cognitively blocking, or canceling, the visual sensation known as *saccadic suppression* to avoid further processing the retinal image [127]. This suppression may already occur at the retina as some research has found in animals [60]. Surprisingly, in a recent study, a complete suppression is not found during a saccade as humans are still capable of on-the-fly correcting or jump-starting their saccade amid the movement, implying there might exist continuous degrees of cognitive suppression [116].

A passive fixation is when the eyes are locked relative to the eye socket; therefore, rotating the head will also rotate the line of gaze. In contrast, active fixation and smooth pursuit are achieved when the gaze is stable and fixed on a visual target, for example, when tracking a moving object. Active fixation has been reported to reach angular velocities up to $0.5\,\mathrm{rad\,s^{-1}}$, while any greater velocities tend to induce saccadic behavior with velocities of $17\,\mathrm{rad\,s^{-1}}$ [100]. Lastly, post-saccadic oscillation is the brief, temporary transient state characterized by small oscillations in the eye movement occurring after a saccade.

These gazing behaviors can be related in Eq. (2.1) when modeling the formation of retinal optic flow. The biomechanical reflexes and smooth pursuit attempt to

null the projected motion flow in the fovea (origin) in the eye, i.e. $\vec{Q}(0,0) \approx \vec{0}$, by accurately and precisely counter-rotating the eyes by controlling the attached fast-twitching muscles. This process must be highly performant and simultaneously rapid to stabilize the retinal image properly, thus demanding a tight control loop of the reflexes and eye muscles. The resulting retinal optic flow field in the proximity of the *fovea*, where the *visual acuity* is the greatest, is nearly nulled, achieving a retinal image stabilization.

Significant advancements in eye-tracking technology have been made in recent years, which in turn have stimulated and reinvigorated retinal optic flow research, experimental psychology, and cognitive sciences. One such innovation is the compact and wearable design, allowing the research participant to move unhindered while wearing the sensors effortlessly. Active or passive vision sensors may be used depending on the chosen numerical methodology for detecting and estimating human gazing. Salvucci and Goldberg [114] made a taxonomy study of five general algorithms for identifying and estimating saccades and fixations: (1) velocity-based, (2) hidden Markov model-based, (3) dispersion-based, (4) minimum spanning tree-based, and (5) area-based algorithm. In their work, they investigated the main advantages and disadvantages of the five, highlighting their suitability for various contexts. The work mainstreamed the development of eye-tracking research, enabling new types of eye-tracking technologies and improving existing systems.

The new high-quality eye data, provided by modern eye-tracking sensors, enables significantly improved data processing. This has allowed researchers to study in detail what, how, and why humans look at things in various contexts, something that previously has been challenging or even impossible. For example, Matthis, Yates, and Hayhoe investigated how humans use gazing to plan and control their foot placement when walking in natural terrain [83]; and Nishizono et al. characterized and studied how professional formula drivers shift their blinking behavior patterns depending on specific driving contexts [98]. Pekkanen and Lappi developed a hidden Markov model-based method using *naive segmented linear regression* (NSLR) to simultaneously denoise and classify eye movements by assuming constant gazing velocity into the four categories mentioned previously [100]. NSLR has been further improved by Johari et al. [62] by using a clustering algorithm to find the linear segments instead.

As previously mentioned, humans tend to intelligently and skillfully guide their gaze through the scene using knowledge-driven gaze control [58, 138]. Simultaneously, a debate was initiated on whether human sensorimotor behavior could be understood as a model-free strong online control, an inherent internal model-based control creating the anticipation used for control, or a composite hybrid of the two [150]. Tuhkanen et al. investigated this issue and found that human predictive gazing behaviors emerged during curvilinear locomotion in a simulated environment [129]. They argue that research participants must have an internalized model to facilitate these predictive gazing behaviors. This contrasts with a model-free

online control, where the control is immediately triggered based on currently available information. The latter case implies that the person can solely guide their locomotion without involving additional complex or higher cognitive functions, such as prediction using readily available visual information. Their findings suggest that humans must have internal models to support predictive behavior. This notion may be partially supported by the study in Paper A, looking at the curve negotiation and maintenance. The research participants may have internalized models to skillfully sample relevant visual information through proper gaze control for aiding their mission-critical control, see Section 5.1.2, which highlights and discusses the gazing distributions.

To summarize, humans express eye movements as biomechanical reflexes such as smooth pursuit and vestibulo-ocular reflexes, which achieve image stabilization at the retina in the eyes. Stabilization is necessary to maximize the acute clarity of visual sensation and further process visual information. While it is widely accepted that gazing mainly achieves retinal image stabilization, it is unclear how exactly proper gaze control aids locomotor control beyond solely stabilization, as humans tend to look where they want to go.

2.2.2 Perception for curvilinear locomotor control

In the context of human local navigation, the scientific community agrees that the retinal optic flow field is used for human locomotor control. However, it is still unclear *how* the visual motion information is perceived, interpreted, and utilized. A notorious point of disagreement is whether the direction of egomotion, i.e. *heading* of the locomotor axis, is used, or even *necessary* in human locomotor control.

As previously mentioned in Section 2.1, the complete retrieval of information such as egomotion and depth perception can not be exactly determined from solely optic or retinal optic flow. Despite this, Warren and Fajen still argued that there is *sufficient* estimation of extra-retinal information that could be exploited for navigational tasks [139, 140]. Hence, they proposed a simple control law for on-foot using a linear combination of egocentric heading error and optic flow error

$$\dot{\phi}_{\text{ego}} = -k(\phi_{\text{ego}} - \psi_{\text{g}}) - wvk(\phi_{\text{of}} - \psi_{\text{g}}) \tag{2.2}$$

where ϕ_{ego} is the current heading from egomotion (egocentric), ψ_{g} the current target goal, k the constant coefficient turning rate, w the optic flow measure dependent on the visual structure of the environment, and v the egomotion speed. If optic flow is to be exploited in a texture-rich environment, the proposed control law is divided into two contributing parts. The optic flow control part governs during high speed, and the extra-retinal control governs during low speed. This control was also extended to accommodate obstacle avoidance and path routing by slightly modifying the model.

Contrary to locomotor strategies exploiting extra-retinal information, Kim and Turvey were the first to show that exploiting the formation of retinal optic flow patterns may be sufficient for locomotor control in curvilinear paths on a textured plane [64]. In their work, they discussed how the projected retinal flow patterns may be used to perceive under-steering and over-steering by integrating the optokinetic gazing of the agent, i.e. smooth pursuit. However, due to the technological limitations of their time, all of their work related to retinal optic flow is entirely conducted in simulation.

Simultaneously, in parallel to the work of Kim and Turvey, Lappe et al. published a literature review on the visual heading perception during egomotion, concluding that heading can be estimated rapidly and accurately from retinal optic flow [69]. However, Lappe et al. further argue that extra-retinal information is needed to resolve the perception problem in ambiguous heading situations. This is supported by Wilkie and Wann when they investigated how humans navigate in situations with degraded visual information, such as nightfall [144]. It is suggested that the human brain has learned to solve and discern the retinal optic flow into motion flow relative to the locomoting body, using optokinetic information as an example, thus perceiving the locomotor heading.

Wann and their colleagues further developed and expanded the idea of Kim and Turvey on retinal optic flow. They argued that extra-retinal information, such as heading and eye movement, is unnecessary for sufficient human locomotor control in curvilinear motion. Furthermore, the model presented by Fajen and Warren could only explain the special case of walking and could not be extended or generalized to curvilinear locomotor control, maintaining steering for road vehicles as explored in Paper A. Wann et al. argued that humans could pivot their locomotor axis and take a straight line trajectory to the target [137], thus indirectly questioning the generalization of the locomotor controls using extra-retinal cues, such as in Eq. 2.2 by Warren and Fajen.

Instead, the works of Kim et al. and Wann et al. suggested retinal optic flow as the primary information source guiding general locomotor control. They presented a mathematical framework using the perceived retinal optic flow of the *intended path* [64, 138]. A theoretical steering control could be realized by continuously minimizing the horizontal flow field component of the intended path, which Wann et al. call *nulling flow curvature with active fixation*. Here, it is to be noted that the intended path of the agent implies volition with short-term planning or anticipation of their local navigation. This strategy is based on the simple heuristic principle of *you should look where you are going* as inspired by the instruction commonly taught in professional driving schools and as a response to *where we look when we steer* by the work of Land and Lee [68]. Assuming that the agent locomotes in a curvilinear trajectory and fixates their gaze onto a point above the current curvilinear path and parallel to the ground surface plane, i.e. towards the horizon. Then, the retinal optic flow field of the intended path can be described

as a ratio of horizontal and vertical components

$$\frac{\dot{U}}{\dot{V}} = \frac{\sqrt{2}(r^2 - x_{\mathrm{p}}^2 - z_{\mathrm{p}}^2)}{(y_{\mathrm{e}} - y_{\mathrm{p}})((x_{\mathrm{p}} + r)\sqrt{1 + \cos\theta} + z_{\mathrm{p}}\sqrt{1 - \cos\theta})} \tag{2.3}$$

as further explained in detail in their work [138]. In this mathematical derivation using their reference coordinate system, one may conclude that all points on the intended path, i.e. $x_{\mathrm{p}}^2 + z_{\mathrm{p}}^2 = r^2$, result in the numerator being zero, implying no visual horizontal flow. Eq. (2.3), as presented in this work, is in the correct form and differs slightly from the erroneous original presented in the Supplementary Material by Wann and Swapp [138]. Despite this error, it does not affect the qualitative analysis made by the authors.

Further research shows that both retinal flow and extra-retinal information could support locomotor control through a flexible combination similar to the works of Warren and Fajen. Using the damped harmonic oscillator in physics as an inspiration, Wilkie et al. [145, 144] proposed

$$\ddot{\theta} = k_1(\beta_1 \dot{P}_{\mathrm{RF}} + \beta_2 \dot{P}_{\mathrm{ERD}} + \beta_3 \dot{P}_{VD}) + k_2(\beta_4 P_{\mathrm{ERD}} + \beta_5 P_{\mathrm{VD}}) - b\dot{\theta} \tag{2.4}$$

where θ is the heading, k_i control gain coefficients, b some damping factor, β_i weights balancing the perception variables, and $\dot{P}_{\{\mathrm{RF,VD,ERD}\}}$ perception quantity variables for retinal optic flow (RF), visual direction (VD) and extra-retinal target direction (ERD). This model attempts to incorporate all available perception signals into a single process to generate a corrective control signal.

In parallel with developing the nulling flow curvature strategy as presented above, Wann, Land, Wilkie, and their colleagues [137, 138, 146] propose a simplified heuristic of using the emergent retinal optic flow patterns to maintain proper locomotor control in curvilinear motion as follows:

R1: Fixate the gaze close to the intended path.

R2: The retinal optic flow directions of the intended path will then indicate under-steering or over-steering through the horizontal component of the motion field.

R3: If the retinal optic flow directions of the intended path are downward, then the future path and intended path coincide, and no further locomotor correction is needed.

However, if detailed and explicit enough, these proposals implicitly suggest the classical engineering approach, proportional–integral–derivative (PID) control, to continuously regulate a process variable to a set point. This type of controller does not reflect how humans typically exert movements via what we understand as reaching, let alone accounting for response times and the intermittency of human

behavior. Aspects of human intermittent control and its inherent delayed response will be further discussed in Section 3.3.

Zhao and Warren attempt to characterize different approaches to modeling human sensorimotor behavior either as strong online-based, strong model-based, or a hybrid of both [150]. However, with the current understanding from Paper A, and in line with recent research in human sensorimotor modeling, a whole distinction between strong online and strong model characterization can not easily be made. For example, anticipation and prediction use internalized models originating from past learned experiences to form a personal perceptual understanding (model-based control) [129, 57, 55, 148]. While in neurocognitive research and supported in Paper A, humans initiate and complete the muscle motor responses before any other higher auxiliary cognitive processes, such as body motor planning, are fully finished (online control) [103, 142, 52, 148, 42]. As the evidence points both ways, it may be concluded that it is most likely a hybrid of both approaches when understanding human sensorimotor behavior, as it can not solely be explained by a strong online or a strong model-based hypothesis.

Matthis et al. [83, 82, 84] revisited human locomotor control on foot and visual motion perception in naturalistic environments with modern eye-tracking technology. Their initial findings concluded that humans are highly adaptive in their locomotor and gazing behavior to cope with varying demanding environments while traversing. This observed on-the-fly adaptability in human locomotion is also supported by the works of Hanna, Fung, and Lamontagne [54], suggesting that humans cognitively adapt and re-weight relevant sources of information to enhance locomotor capabilities, in this case, extra-retinal and visual motion information. Further, visual motion fixed to the head is significantly chaotic due to the bipedal gait of humans [84]. Thus, the role of optic flow for human locomotion should be reconsidered in favor of retinal optic flow instead, further emphasizing the importance of retinal image stabilization and its interpretation as perceptual cues.

The work presented in Paper A attempts to incorporate the main theories presented by Matthis et al., Kim et al., and Wann et al. regarding the interpretation of retinal optic flow for locomotor control of curvilinear motion. However, there is an inherent challenge in quantifying what is referred to as the *intended path* and the *future path* if such quantities exist, given that only the visual information and egomotion are known. Given that the agent is supposed to maintain their current future path so that this path lies within their designated road lane, then the designated road lane should contain their internalized projected intended path. Thus, it is arguably that the designated road lane visually contains the intended path, as one of the main objectives of the control is to stay within some boundaries, such as lane markings. This is assumed and used in the naive numerical approach of estimating and interpreting the retinal optic flow presented in Section 4.5.

2.2.3 Smooth pursuit during curvilinear locomotion

To geometrically relate the case of an agent fixating their gaze on their circular path ahead, let us consider the following in the plane parallel to the ground. Using the *inscribed angle theorem*, as shown in Fig. 2.3a, it can be stated that

$$\varphi = 2\theta \tag{2.5}$$

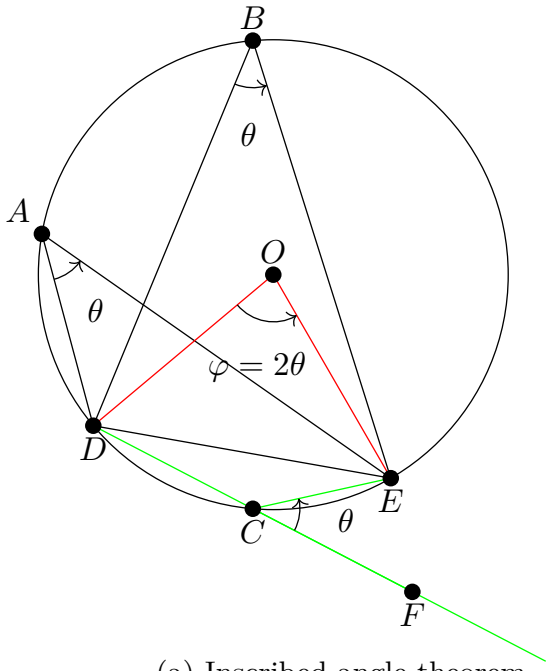
for any point A in the set of points M on the circular arc which the angle $\angle DAE = \theta$ can be constructed. The complement set M^c, which is the complement arc to complete the circle, and the supplementary angle $\angle DCE = \theta$ can be constructed as illustrated in Fig. 2.3a at point C.

Let the points A and C converge to the point D from each side such that it establishes the special case of a tangent AF where $\angle FAE = \angle FDE = \angle FCE = \theta$, as shown in Fig. 2.3b. This is the approximate case (disregarding the height dimension) when the agent fixates their gaze at point E on the circular path from their point A, which establishes the horizontal gaze angle θ while the heading of the vehicle is the tangent GF. The time derivative of Eq. (2.5) is thus simply

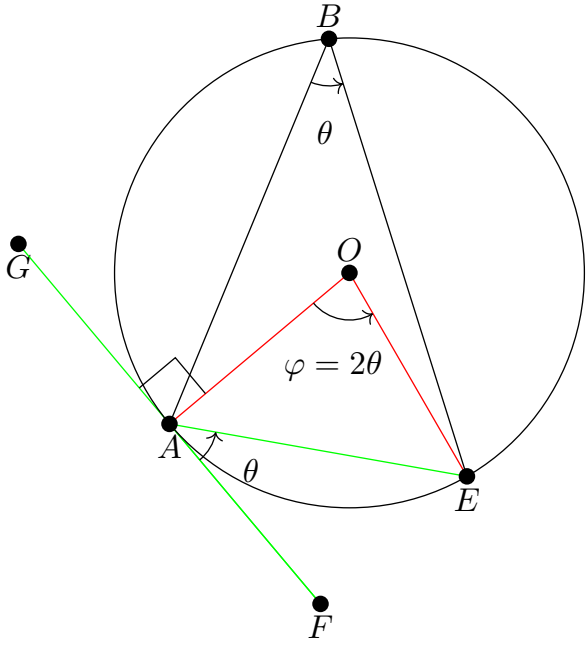
$$\dot{\varphi}(t) = 2\dot{\theta}(t) \tag{2.6}$$

where $\dot{\varphi}(t)$ is the vehicular yaw angular velocity and $\dot{\theta}(t)$ is the horizontal gaze velocity. This proves that the vehicular yaw angular velocity is always twice as large as the horizontal gaze velocity during an *exact* smooth pursuit i.e. the agent perfectly fixates their gaze on the curvilinear path.

To summarize, previous work studying visual perception for human locomotor control on foot and in vehicles looked at how the retinal optic flow (and optic flow) and the heading from egomotion can be exploited for local navigation. Unfortunately, the suggested control mechanisms do not consider human realistic movement featuring characteristics of reaching, latency, and intermittency stemming from a satisficing behavior. Furthermore, the human sensorimotor system can not be explained as solely simple online control (immediately triggering a response) or model-based (intermediate interpretation and modeling), but rather as a hybrid of both. Humans internalize past experiences to form anticipation and prediction to consistently and rapidly initiate informed movement responses. At the same time, they can quickly adapt their locomotor behavior to cope with changing and dynamic environmental circumstances. Many of the proposed control schemes are rooted in using active gaze fixation to form the retinal optic flow pattern, and a geometrical proof is presented relating vehicular yaw angular velocity and horizontal gaze velocity.



(a) Inscribed angle theorem.



(b) Special case of tangent.

Figure 2.3: Geometrical proof relating vehicular yaw angular velocity $\dot{\varphi}$ to horizontal gaze velocity $\dot{\theta}$. (a) shows the inscribed angle theorem, where (b) is the special case of tangent, which can be used to relate an agent at point A fixating their gaze on point E on their curvilinear path around O.

Chapter 3

Neuromuscular behavior

The thesis considers human movements, specifically reaching, and how these reaching movements are expressed as corrective responses associated with stimulus cues. Human movement control fundamentally differs from those typically found in classical control, such as the PID controller, as mentioned in Section 2.2.2. Despite this, human control is often modeled using classical control approaches, which may oversimplify human characteristics and yield an inaccurate representation of human behavior. Furthermore, much research regarding human behavior in perception and control and its applications has been carried out in isolation, further widening the gaps between the fields.

This chapter will discuss human behavior related to muscular control and movements, mental chronometry, and corrective responses to perception cues, harmonizing the current understanding of the human sensorimotor system across fields.

3.1 Motor control and muscle coordination

The human central nervous system compiles and interprets sensory information and propagates signals to lower motor neurons connected to muscle fibers, to accomplish a muscle contraction response. The signals are conveyed via axons in neural networks, where neurons fire and propagate an electric action potential through biochemical reactions in their cell membranes. When humans initiate a voluntary muscle contraction, the signals originate and are transmitted to the muscles, often via the spinal cord, from the so-called *motor cortex* of the brain. The motor cortex has three specialized areas: the *primary motor cortex*, the *premotor cortex*, and the *supplementary motor area*. The former two are mainly credited for

preparing, planning, guiding, and coordinating across body sides, while the latter is considered the primary contributor to generating neural signals to the muscles.

There are multiple competing theories on how the central nervous system copes with the grand scheme and the complexity of muscle coordination involving various muscles and joints when executing and accomplishing a movement task. *Muscle synergy* is one of the competing theories. It hypothesizes that learned patterns govern time-varying modulation, rate of recruitment, and strength of specific muscles to achieve the intended movement patterns [31, 16, 26, 32], for example when walking, hand waving, and finger pointing. The suggested muscle coordination is done by mapping initial states and task-specific goals to time-varying activation of muscle synergies on a higher level, contrary to individually micro-controlling each muscle primitive in a heavily centralized manner [31, 16]. The achieved limb movement is functionally illustrated in Fig. 3.1, where the primary motor cortex, after selecting a motor program, fires neural signals to activate selected muscle synergies, which in turn directly recruit individual muscles.

d'Avella et al. developed a mathematical framework describing muscle activation using the theoretical framework of muscle synergies. Consider K number of muscles activated by J time-invariant muscle synergies, then the muscle activation can be expressed as

$$\vec{m}(t) = \sum_{j=0}^{J-1} c_j(t)\vec{w}_j \tag{3.1}$$

where $m(t)$ is the K-dimensional vector describing muscle activations, $c_j(t)$ is the coefficient modulating strength and time activation of the synergy, and $\vec{w}_j$ is the K-dimensional muscle synergy vector involving the K muscles. If the synergy is time-variant, then the equation can be modified by transferring the time modulation of the coefficient c_j to the synergy $\vec{w}_j$, resulting in

$$\vec{m}(t) = \sum_{j=0}^{J-1} c_j\vec{w}_j(t - t_j) \tag{3.2}$$

where t_j is the time shift for activating the muscle synergy, and now c_j is time-invariant and strictly modulates the activation strength. It is further depicted as a cognitive actuation subsystem, a model proposed by Markkula et al. [81] shown in Fig. 3.3b, which will be further discussed in Section 3.3 in the context of intermittent control and human sensorimotor model.

Intriguingly delving deeper into the development of muscle synergies and their possible origin, Dominici et al. investigated how basic or rudimentary movements develop into sophisticated ones from being neonates (newborn human infants) to becoming adults [32]. The basic patterns of the motor neuron (motoneuron) signals responsible for activating the stepping movements were retained throughout life. They suggest that these basic stepping movements are augmented and evolved into

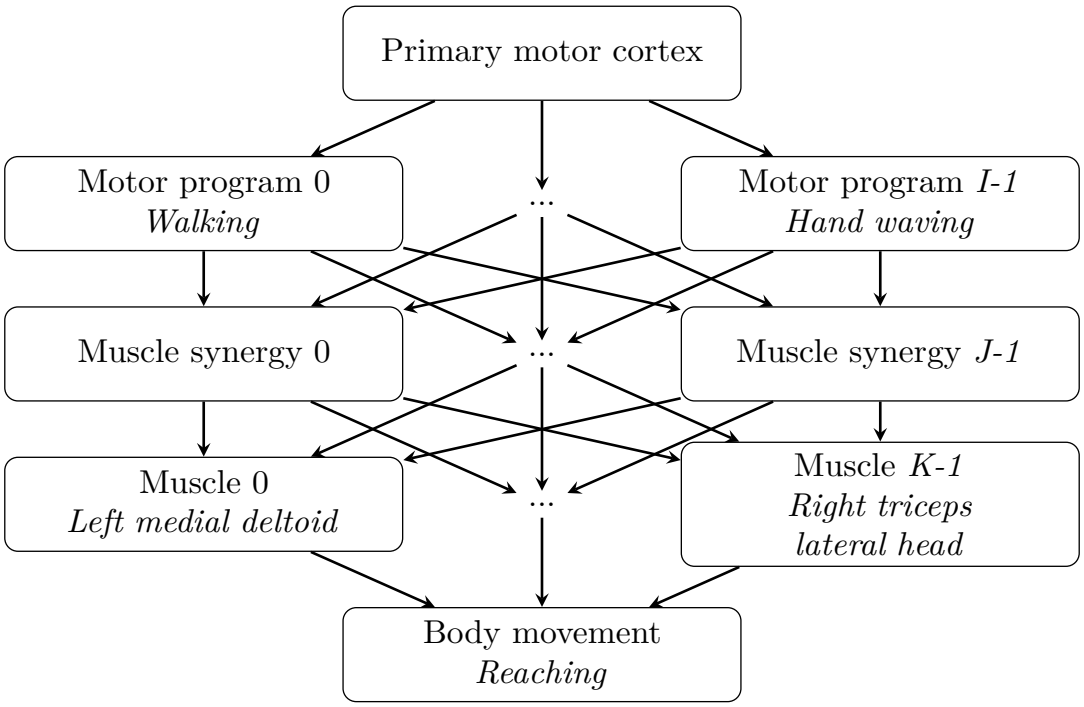


Figure 3.1: Execution and coordination of muscles via muscle synergies. The diagram illustrates an abstraction of how the human central nervous system executes movement, from the primary motor cortex in the brain to a particular muscle. Accomplishing a movement is seemingly effortless but a complex endeavor, often involving multiple muscles contracting asynchronously and coordinated at different joints. It is thought that our brain does not individually control each muscle through micromanagement but executes previously learned and known motor programs, such as walking or hand waving. This approach would alleviate the cognitive load from the brain, delegating muscle synergies to other parts of the central nervous system.

the sophisticated human bipedal locomotion. Furthermore, similar motoneuron signal patterns for basic stepping are found in other mammals in early infancy, suggesting that the first building blocks for movement share a common ancestral origin for mammals, and one might inherit these shared rudimentary movements through genetic material. But then, how do humans enhance or augment muscle synergies when learning or adapting to new movements? Cheung et al. further investigated how these muscle synergies may be developed when adapting new movement patterns in human runners [26]. They suggest the processes of *fractionation* (splitting) and *merging* (combining) of existing synergies to create new muscle synergies, which result in an altered limb movement. These suggestions imply that muscle synergies are highly adaptive and tailored to each body of the individual in question, and further their application.

To summarize, muscle synergy is proposed to simplify how the human central nervous system coordinates muscle contractions, producing effortless yet highly complex movement patterns such as walking, finger-pointing, or hand waving. Instead of micro-managing the individual muscle primitives, the central nervous system governs on a high cognitive level by internally mapping initial states and task-specific goals to muscle activation and recruitment via so-called muscle synergies.

3.2 Modeling human reaching movements

When humans intentionally actuate rapid limb movement from an initial point to a target point, the movement often has a logistic shape in position and, by extension, a bell-shaped velocity profile. This is the neurological biomechanical phenomenon known as *reaching* [53, 104]. Reaching was initially observed and studied in two dimensions due to experimental restrictions [88], but has later been observed in three dimensions as well [4, 53]. Assuming the velocity profile is symmetric around a mode (optima) for a single ballistic reaching correction, it may be expressed and approximated through a Gaussian function. Thus, each correction movement velocity may be described as

$$\dot{\delta}_{\mathrm{b}}(t, a_{\mathrm{I}}, \mu_{\mathrm{I}}, \sigma_{\mathrm{I}}) \triangleq a_{\mathrm{I}} \exp \frac{-(t-\mu_{\mathrm{I}})^2}{2\sigma_{\mathrm{I}}^2}, \tag{3.3}$$

where a_{I} is the correction strength, μ_{I} is the mode of correction determining its timeliness by time shifting, and σ_{I} is the correction rate parameter. To determine the positional sigmoid function, it is derived as the primitive function of Eq. (3.3)

$$\delta_{\mathrm{b}}(t, a_{\mathrm{I}}, \mu_{\mathrm{I}}, \sigma_{\mathrm{I}}) = \int a_{\mathrm{I}} \exp \frac{-(t-\mu_{\mathrm{I}})^2}{2\sigma_{\mathrm{I}}^2} \mathrm{d}t = a_{\mathrm{I}}\sigma_{\mathrm{I}}\sqrt{\pi/2}\ \mathrm{erf}(\frac{t-\mu_{\mathrm{I}}}{\sqrt{2}\sigma_{\mathrm{I}}}) + C \tag{3.4}$$

where $\mathrm{erf}(\dots)$ is the Gauss error function and cannot be expressed through elementary functions, and C is some constant determined by a known condition in

position (usually a boundary condition). Then, to get the full travel length of the ballistic movement, the correction is

$$\Delta\delta_{\mathrm{b}}(a_{\mathrm{I}},\sigma_{\mathrm{I}}) = \int_{-\infty}^{\infty} \dot{\delta}_{\mathrm{b}}(t,a_{\mathrm{I}},\mu_{\mathrm{I}},\sigma_{\mathrm{I}})\mathrm{d}t = a_{\mathrm{I}}\sigma_{\mathrm{I}}\sqrt{2\pi}. \tag{3.5}$$

The travel length is entirely determined by the correction strength and the speed parameter, which is to be expected. Furthermore, this is the corrected form using the definite integral of the Gaussian function compared to the one presented in the works of Benderius and Markkula [11].

A numerical approximation of the Gauss error function can be used for computation using Taylor series expansion

$$\mathrm{erf}(t) = \frac{2}{\sqrt{\pi}}\int_0^t e^{-x^2}\mathrm{d}x = \frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{(-1)^n t^{2n+1}}{n!(2n+1)}. \tag{3.6}$$

Naturally, in a similar fashion to deriving the position profile, the acceleration profile is the time derivative of Eq. (3.3) as follows

$$\ddot{\delta}_{\mathrm{b}}(t,a_{\mathrm{I}},\mu_{\mathrm{I}},\sigma_{\mathrm{I}}) = \frac{\mathrm{d}}{\mathrm{d}t}a_{\mathrm{I}}\exp\frac{-(t-\mu_{\mathrm{I}})^2}{2\sigma_{\mathrm{I}}^2} = \frac{a_{\mathrm{I}}(\mu_{\mathrm{I}}-t)}{\sigma_{\mathrm{I}}^2}\exp\frac{-(t-\mu_{\mathrm{I}})^2}{2\sigma_{\mathrm{I}}^2}. \tag{3.7}$$

Using the properties of the Gaussian functions and inspiration from statistics, one can define onset and offset timings for each reaching correction using the standard deviation parameter which is equivalent to σ_{I}. The onset and offset timings are defined as

$$t_{\mathrm{I},\{\mathrm{on,off}\}} \triangleq \mu_{\mathrm{I}} \mp 2\sigma_{\mathrm{I}} \tag{3.8}$$

which then encompasses approximately 95.47 % of the correction movement of a duration of $t_{\mathrm{I,off}} - t_{\mathrm{I,on}} = 4\sigma_{\mathrm{I}}$. In the reaching movement trajectory, the onset is equivalently defined at approximately 2.28 % and offset at approximately 97.72 %.

To construct a complex reaching movement, each ballistic reaching movement correction is superposed as follows

$$\dot{\delta}(t) = \sum_{(a_{\mathrm{I}},\mu_{\mathrm{I}},\sigma_{\mathrm{I}})\in\Omega_{\dot{\delta}}} \dot{\delta}_{\mathrm{b}}(t,a_{\mathrm{I}},\mu_{\mathrm{I}},\sigma_{\mathrm{I}}), \tag{3.9}$$

where $\Omega_{\dot{\delta}}$ is the set containing all correction parameters for each ballistic correction. Theoretical examples of human reaching using the Gaussian function are shown in Fig. 3.2 with a simple single movement and two superposed ballistic corrections creating the resulting complex movement.

Plamondon further extended the model by considering the asymmetry as observed in the experimental data of human reaching [104], modifying Eq. (3.3) to support a right-sided heavy tail

$$\dot{\tilde{\delta}}_{\mathrm{b}}(P(t),a_{\mathrm{I}},\mu_{\mathrm{I}},\sigma_{\mathrm{I}}) \triangleq \frac{a_{\mathrm{I}}}{P(t)}\exp\frac{-(\ln(P(t))-\mu_{\mathrm{I}})^2}{2\sigma_{\mathrm{I}}^2}, \tag{3.10}$$

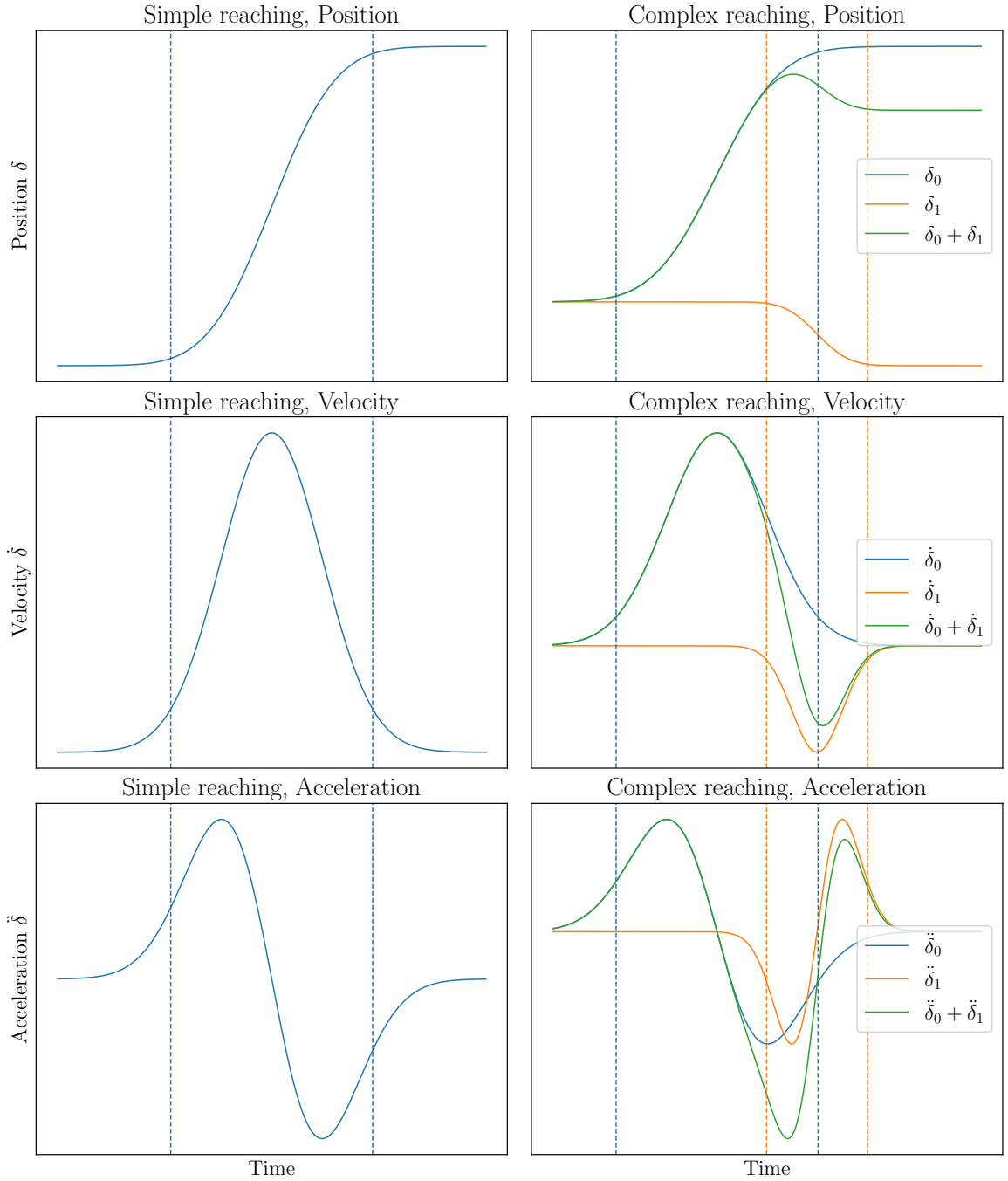


Figure 3.2: Human reaching movement patterns. Theoretical reaching movements described using the Gaussian function to approximate the velocity profile of reaching are demonstrated in a simple and a complex case. The complex movements are constructed through the superposition of two simple ballistic movements. The vertical lines indicate the onset and offset timings at approximately 2.2 % and 97.8 % respectively of each correction movement trajectory.

where

$$P(t) = \frac{t - t_{\text{on}}}{t_{\text{off}} - t_{\text{on}}}, \quad \text{for} \quad t_{\text{on}} \leq t \leq t_{\text{off}}, \tag{3.11}$$

$t_{\{\text{on,off}\}}$ are the onset and offset timings for the reaching correction. Despite better approximation, this modification convolutes the interpretation of the model, as both the strength and the mode of the correction are no longer easily interpretable directly in Eq. (3.10) compared to Eq. (3.3). Thus, for this thesis, the bell-shaped velocity will be considered symmetric, i.e. absence of a heavy tail, and approximated with Gaussian functions of each ballistic correction.

To summarize, humans commonly exert limb movements as reaching movements. A ballistic reaching movement is characterized by having a sigmoid curve of the distance and, by extension, a bell-shaped velocity profile. The velocity profile may be approximated and expressed by a Gaussian function with three parameters determining its strength, rate of movement, and timeliness. The general reaching movement may be overlapped by the superimposing simpler ballistic corrections, creating a train of superposed ballistic corrections resulting in a complex movement pattern.

3.3 Sufficient response through intermittent control

The biomechanical actuation and control of limbs with intermittency, or *intermittent control* phenomenon, emerges from the fact that muscles can only generate pulling forces through contraction, often around a joint. Thus, biological systems are evolutionarily optimized for energy efficiency, co-contraction, muscle synergy, etc. Since energy is a valuable resource, things ought to change *only* when there is a sufficient need for it, which may partially explain the existence of the intermittent control phenomenon. Despite intermittent control being a well-established topic in experimental psychology [134, 29], it is only recently gaining attention outside of its field of inception when studying human control behavior [11, 81].

Perfiliev extensively investigated the organization of reaching movements, in particular how the programming and updating are performed in humans (and cats) [102, 103]. It was found that the sensory signals relevant to the initiation of reaching are transformed into neural motor programs (commands). Additional on-the-fly corrections may be applied if the sensory feedback of the response does not align with the internal prediction of reaching or if the target changed from the time of the initiation of the response, e.g. when pursuing a moving target. Furthermore, the research conducted by Yeom, Kim, and Cheung further indicates that all neural cog nitive processes related to motor control are completed during motor planning [148]. This implies that motor planning processes are more informative as an internal predictor to evaluate and assess the initiated response. These findings

provide a possible explanation for the quick and overlapping reaching corrections, or *trains of ballistic corrections*, observed in human reaching movements.

Interestingly and perhaps yet unsurprisingly, it has been found that human-operated control interfaces with constraint movement, such as the steering wheel, share the same fundamental properties as human reaching and intermittency [11, 120]. Extending the works of Benderius and Markkula[11, 81], Paper A presents and demonstrates a novel computational method using a meta-heuristic optimization method, particle swarm optimization, to identify trains of ballistic corrections in experimental data, which is further detailed in Section 4.6.

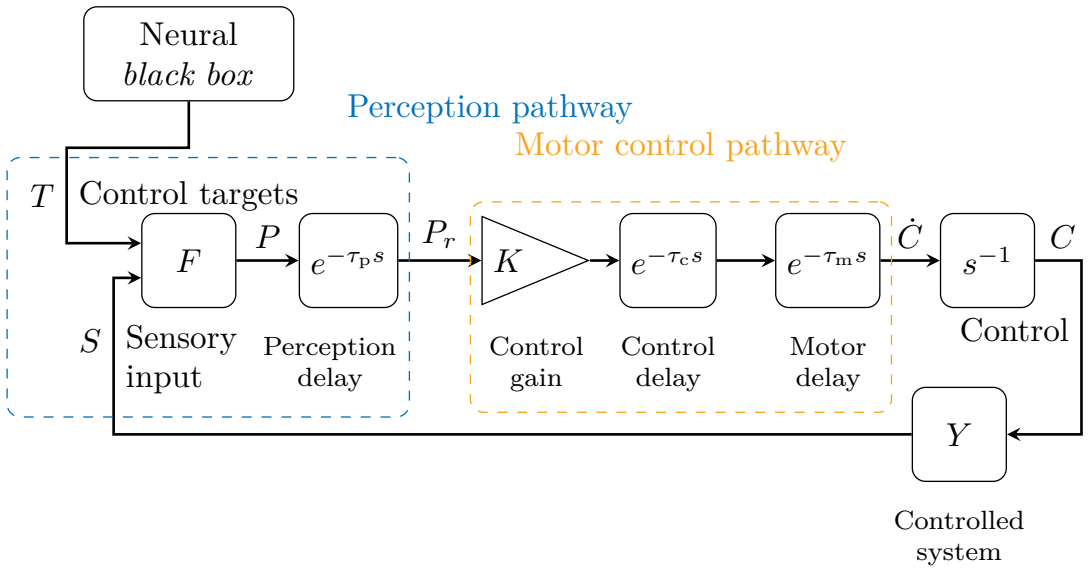
Markkula and colleagues [81] have proposed a computational framework model reflecting human cognitive sensing and control behavior properties, a *sustained sensorimotor control* as illustrated in Fig. 3.3. It considers and models the distinct cognitive subsystems: perception, control, and muscular pathways, using the driving steering behavior as an applied study case of intermittent control. This proposed sensorimotor model features interesting concepts such as *activation* motor primitives and muscle synergies (previously discussed in Section 3.1 and in Fig. 3.1), internalization of *prediction and anticipation* of its actuation, and *evidence accumulation* to trigger an output response in one common computational framework achieving intermittent control, see Fig. 3.3b.

These cognitive functions collectively contribute to the resulting trains of ballistic corrections creating complex limb reaching movements to achieve the task at hand [29, 103, 102, 81]. This mimics the biological behavior of actively initiating a ballistic response only when there is a *sufficient* need for it, i.e. *satisficing* behavior [9, 122, 117]. This fundamentally differs from a continuously optimized control behavior like the commonly used PID controller [128]. Furthermore, as implied by this sensorimotor model (neural black box and prediction), by Paper A, and in Section 2.2.2, the human control behavior can not be solely explained as strong online or pure model-based control but as a hybridization of both approaches.

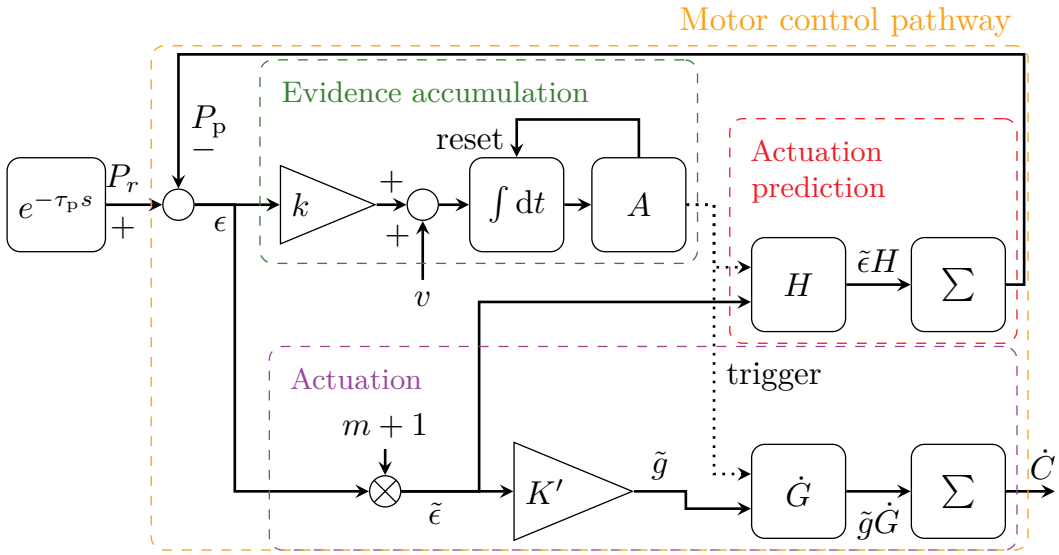
Using the generic sensorimotor model by Markkula et al. [81], see Fig. 3.3a, to describe the total delayed response time

$$\tau_{\mathrm{d}} \triangleq \tau_{\mathrm{p}} + \tau_{\mathrm{c}} + \tau_{\mathrm{m}} \approx t_{\mathrm{I,on}} - t_{\mathrm{s,on}}, \tag{3.12}$$

where $\tau_{\{\mathrm{p,\ c,\ m}\}}$ are the constituent delays for perception, control, and muscular activation pathways, respectively, achieving a timely response. This should approximately match the time difference from a stimulus onset $t_{\mathrm{s,on}}$ to respective movement onset $t_{\mathrm{I,on}}$. This was extensively investigated in Paper A for retinal optic flow, optic flow, heading, and lateral position as an information source to trigger a ballistic reaching movement in the form of a steering correction. Suggested by using an online control model, effectively and directly relating the stimulus signal S to reaching corrections $\dot{C}$ and C, as shown in Fig. 3.3b. Out of the four examined information sources, retinal optic flow had the strongest correlation



(a) Generic sensorimotor control system



(b) Motor control pathway featuring intermittent control

Figure 3.3: Sustained sensorimotor model by Markkula et al. [81]. (a) illustrates a generic linear control system, imagining the human perception and motor control pathways in state-space Laplacian frequency domain representation. The different response delays $\tau_{\{p,\,c,\,m\}}$ are also presented for perception, control, and muscular activation pathways. $\dot{C}$ and C are velocity and positional control variables, respectively, for the resulting responses, e.g. limb movements. The motor control model proposed by Markkula et al. is shown in (b), featuring *evidence accumulation* that transforms the generic control in (a) into an intermittent control. The *actuation* of the muscle synergies subsystem is functionally depicted in Fig. 3.1.

(using Spearman's correlation) to the strength of correction strength a_{I} as defined in Eq. (3.4) at response time $\tau_{\mathrm{d}} \approx 0.14\,\mathrm{s}$.

To summarize, humans divert their attention to sufficiently solve tasks at hand, producing intermittent control behavior as we understand it. Intermittent control fundamentally differs from classical control approaches such as PID control, as corrections are applied intermittently in contrast to continuous evaluation and correction. It has been observed that the human sensorimotor system is better explained as an intermittent controller, which then explains why humans apply corrections on the fly. This results in trains of ballistic correction as sufficient responses to stimuli.

Chapter 4

Robotics for bionic navigation

This chapter will discuss robotic systems in terms of hardware and software, and further implementations and applications for local navigation using vision sensing and visual motion flow perception. In particular, it concerns how to fairly evaluate the performance of optic flow estimation methods for real-time scheduling and application in embedded systems (Paper C). This work presents and demonstrates a software architecture to support and facilitate robotics research and development agnostic to hardware platforms (Paper D). Finally, localization applications of exploiting optic flow are investigated and evaluated for robotics in maritime settings (Paper B) in addition to using optic flow methods on unconventional topographic data to determine egomotion (Paper E). The software implementations and applications presented here permeate the underlying work supporting this thesis and the research findings presented in the appended papers.

4.1 Vision sensors

At the fundamental level, human color vision at the retina can distinguish the three primary base colors of red, green, and blue. To reconstruct the full range of the visual colors, the human vision perception cognitively recreates the other colors by blending the three. The additive superposition of the primary colors can, therefore, create the perceived color in humans despite their inability to sense that particular color directly. The underlying technique exploits how human color perception functions and encodes vision information in these three base colors to

recreate the representation of the scenery functionally. Implementing the same idea in cameras is to use a well-established Bayer filter on top of an image sensor when capturing the three primary colors, or three color channels, often referred to as *RGB*. If the camera sensor only captures the imagery as a single-channel image, it is referred to as a monochrome camera.

In traditional vision sensors, the fundamental measurement procedure is to measure the light illumination over the image sensor for a fixed time duration referred to as the *exposure time*. This can be done either for the entirety of the frame using a *global shutter*, or on a row-by-row basis using a *rolling shutter*. The former ensures that the whole sampled image originates from the same time instant, while the latter allows for more efficient use of the hardware and related computational resources. However, the rolling shutter may introduce motion artifacts due to non-synchronous row-by-row basis sampling, resulting in distorted depicted objects.

RGB vision is fairly well-used in robotics applications, as additional information may be embedded or extracted from colors. Computer vision algorithms and robotics control implementation often favor shorter latency, simplification of color blending, and smaller data footprints, making the monochrome approach as common as RGB vision. Optic flow estimation methods using monochrome vision are far more common for these purposes. More recent optic flow estimators often fully exploit the three color channels, further enhancing their accuracy and precision performance at the cost of additional data throughput and computational resources.

In contrast to the conventional way of interpreting vision using three or single-channel images generated from fixed-length exposure time, there is an alternative way to do it using so-called *event-based camera* vision or *dynamic vision sensor*. These bio-inspired event cameras fundamentally differ from traditional vision sensors. Instead of measuring collected illumination in a fixed time, they measure the per-pixel time to fill an illumination buffer, after which they send an event, i.e. in effect measuring the illumination dynamics in the time domain. Thus, the challenges of rolling shutter, motion blur, and light exposure are absent in event-based vision technology.

This event-based vision sensor technology is likely to be more suited to visual motion processing, which has successfully been demonstrated in the related works of Scarramuzza and Gehrig et al. [44, 23, 87]. An example of this is the works of Gehrig et al. where an event-based vision was used with a deep learning approach estimating the optic flow [45]. Their result showed an improvement of 12 % in the error metrics compared to other state-of-the-art methods. Continuing their work, they presented and demonstrated dense optic flow estimation continuous in time using data from event camera sensors, mitigating the problems of discrete time sampling of traditional camera sensors [46]. Another example is where intensity estimation (image reconstruction) and optic flow estimation were simultaneously computed during challenges of rapid motion and high dynamic range scenes [6].

Furthermore, this novel technology enables an asynchronous data pipeline

compared to the problematic synchronous sampling data flow. In asynchronous data pipelines, less data has to be transferred and filtered through the data processing. In essence, only important or relevant data are sampled and detected as events. In contrast to the event-based sensor, a traditional camera sensor, for example, *Full-HD* 1080p ($1920 \times 1080 \times 3$) at 144 Hz video stream produces a data bandwidth of approximately $0.89\,\mathrm{GB\,s^{-1}}$ assuming an 8-bit pixel channel representation. In reality and practically, the bandwidth is much less, thanks to various video and image compression techniques at the cost of additional computational resources and potentially loss of data integrity. Further, a higher resolution of *Ultra-HD* 2160p ($3840 \times 2160 \times 3$) at the same frame rate would be roughly $3.58\,\mathrm{GB\,s^{-1}}$, a fourfold increase compared to Full-HD. These magnitudes of data bandwidth pose significant challenges for data processing, both on software and hardware levels, which often negatively affect overall run-time performance in robotics applications.

To summarize, past traditional vision sensors are heavily inspired by and adapted for human vision using red, green, and blue data representation, commonly referred to as RGB vision cameras. Historically, this has been sufficient for capturing and storing images to recreate static imagery for the human eye. However, the increasing precision, accuracy, and robotic applications of RGB image sensors present new challenges in managing and processing the sensor information. Event-based camera technology is a novel competing vision technology that has successfully demonstrated its suitability for robotics applications compared to the RGB counterpart, particularly optic flow estimation.

4.2 Robotic software

There are many challenges in designing software architecture when considering robotics and its applications. Some of these challenges are highlighted and addressed in this thesis. In many developed research robotic platforms, e.g. road vehicles and marine vehicles, pipes-and-filters-based data processing approaches are often utilized. This effectively makes them *cyber-physical* systems, defined as "integrations of computation and physical processes," meaning their interaction with the physical world is driven by software algorithms and vice versa. While the pipes-and-filters-based approach is a straightforward and simpler method to implement, there is an argument to be made about the code base becoming complicated, vulnerable, monolithic, and challenging to maintain. One way to mitigate these shortfalls, as investigated and demonstrated in Paper D, is to adopt *containerization* and *microservice* design paradigm, which was done with special consideration towards road vehicles. This was demonstrated in various vehicles of different scales operating in settings emphasizing the high demand for reliability, traceability, and deployment.

4.2.1 Microservice design, real-time computing, and embedded system deadlines

According to Lewis and Fowler [76], a microservice design paradigm may be described as *a suite of small services, each running in its own process and communicating with lightweight mechanisms.* To support such a paradigm, middleware software handling communication and established message specifications protocol has been used as demonstrated in the Paper D in the example of a fleet of various vehicles. Thanks to the very design of the communication protocol, a microservice can be regarded as a highly specialized program block with well-defined input and output signals. The methodology in the Paper C fully implements this, where the sequential images and optic flow estimation are clearly defined as input and output, respectively. Understanding this process created a structured and easy-to-use methodology where run-time executable binaries were easily swapped in place when producing the results.

Furthermore, since the microservice binaries can be packed as *self-sufficient* containers, they can be deployed on various host machines agnostic to CPU architecture. This creates run-time measurements normalized to a single host machine, thus producing valuable, comparable, and fair results. See Table 4.1 and Paper C for further detailed results and discussions. It is noted that the methodology used in Paper C does not favor initialization and clean-up of runtime execution. This places GPU-accelerated algorithms at a significant disadvantage. For example, for *FlowNet2*, the pre-trained weights of its neural network need to be parsed and uploaded to the internal GPU memory.

Relating to the execution and run-time, when machines are programmed to perform a task, they are expected to provide a response before a time constraint, or a *deadline*. Depending on the importance of the task or the severity of the task outcome, deadlines are often classified as *hard*, *firm*, or *soft*, where the severity here ranges from the harshest to the least. This is easily imagined in a critical situation where a subsystem is expected to perform a task, such as anti-lock braking systems in road vehicles, where fatality is a possible outcome. For such reason, embedded real-time systems are designed and constructed to execute programs or instructions that comply with predetermined deadlines. Moreover, it is challenging to guarantee predictability or real-time computing as software systems grow larger and ever more advanced. In general, robotic locomotion should be considered a safety-critical task, so if robotic locomotion depends on optic flow estimation, hard deadlines ought to be defined with required accuracy. However, from a pedantic view, the results in Paper C showed that the majority of the optic flow estimators at their current implemented state do not fully satisfy the requirement for real-time execution and, by extension, are not suitable for a live real-time robotic system in the narrow sense.

Table 4.1: Optic flow estimator performance evaluation results from Paper C. The run-time of various optic flow estimators was evaluated on the complete Sintel dataset, and their average end-point error (AEE) and average angular error (AAE) were evaluated on the final dataset. The evaluation counterpart of Sintel is shown as a reference where applicable. The blue-colored cells are the best performance per category.

Algorithm	Run-time / (s)				AEE / (pixel)		AAE / (rad)
	mean	std. dev.	99th	longest		ref.	
Deepflow [141]	1.027	0.031	1.129	1.236	6.497	7.212	0.187
DIS[66]	0.109	0.006	0.124	0.203	5.857	10.127	0.214
DualTVL1 [149, 115]	6.057	6.674	21.693	23.577	9.113	9.861	0.289
Farneback [38]	0.220	0.017	0.245	0.278	11.492		0.584
FlowNet2 [61]	3.473	0.070	3.604	5.847	2.220	5.739	0.109
Lucas-Kanade [20, 80]	0.199	0.018	0.232	0.246	8.727		0.236
PCA-Flow [147]	0.287	0.059	0.436	0.465	7.843	8.652	0.257
SimpleFlow [125]	0.869	0.042	0.953	0.990	9.150	13.364	0.319

Notably, commonly available personal computers with general-purpose operating systems are not real-time computing machines. These machines are primarily designed to adopt broader hardware support, shared computing resources, and a better user experience. For example, GPUs are specifically designed to parallel-process data for graphics pipelines, which is often time-sensitive and is of significant importance for the user experience. For this reason, to maximize the utilization, the operating system kernel allows for variable and tighter computational timing for smoother graphics rendering, in contrast to well-defined fixed deadlines. However, it is possible to make personal computers comply with deadline computing through software techniques such as a customized and specialized deadline or task scheduler, or efficient locking protocols.

Like many other computer vision algorithms, optic flow estimators typically require high-end features found in general-purpose computation nodes, such as *general-purpose computing on graphics processing units* (GPGPU). The maturity and suitability of using graphics processing units in real-time computing settings are improving and may be considered for soft deadline applications [34, 35]. However, specialized features like GPGPU often require closed-source user-space drivers to operate. Therefore, new and more specialized estimator implementations are needed for future applications where optic flow should be further exploited for locomotion in robotics, especially for real-time embedded systems with hard and firm deadlines.

To summarize, modern robots are increasingly deployed in challenging and varying environments and are increasingly tasked with critical applications. This has resulted in wide and complex configurations of cyber-physical systems used for robotics. With complex cyber-physical systems, the accompanying software is often very sophisticated, complex, and tailored for a specific system and particular application. Adopting adequate software design and deployment techniques, such as microservice architecture, hardware abstraction, and containerization, may address and ease software complexity management in robotics compared to the monolithic development approach. Simultaneously, adopting such designs and techniques also allows for more fair benchmarking and evaluation comparisons in robotics research.

4.3 Optic flow estimation methods

The neurocognition and behavioral aspects of optic flow have been considered in the previous Section 2. However, methods to numerically estimate the optic flow in a temporal sequence of images have not yet been discussed. In this section, a brief overview of optic flow estimation will be presented.

Raudies, Andreev, Wildes et al. created a descriptive taxonomy overview of various optic flow constraints used in numerical estimation in the literature [108, 3, 143]. Many of the optic flow estimation computation routines can be generalized

and described with the help of the well-known continuity equation in physics. The continuity equation for optic flow can be described as

$$\partial_t \rho + \nabla_{xy} \cdot (\vec{v} \cdot \rho) = \sigma \tag{4.1}$$

where ρ is the density of some quantity, ∇_{xy} the divergence operator with respect to x and y, $\vec{v}$ the optic flow, and σ the generation or destruction of the quantity ρ. The equation is kept to a generalized quantity ρ since different algorithms use a different quantity in their applications to estimate the optic flow $\vec{v}$. It should be noted that most optical flow estimations assume that the quantity ρ is conserved, thus simplifying the expression to

$$\partial_t \rho + \nabla_{xy} \cdot (\vec{v} \cdot \rho) = 0. \tag{4.2}$$

This form serves as a common starting point for many traditional naive optic flow estimation methods. Imposing additional constraints further transforms the equation to a more approachable and computationally feasible form.

Optic flow constraint equation

Several additional assumptions are made before detailing the form of the optic flow constraint equation. Let the quantity be a scalar-valued intensity quantity $\rho = g(x, y, t)$ and assume that $\nabla_{xy}\vec{v} = \vec{0}$ where x, y are spatial dimension variables, for example, in image space. The assumption is motivated by the argument that in a contained local pixel domain in an image, the brightness can not be accumulated or lost due to the brightness flux. If a *brightness constancy* assumption is made, then the following approximation can be constructed

$$g(x + \Delta x, y + \Delta y, t + \Delta t) \approx g(x, y, t). \tag{4.3}$$

Further, if an approximation of the first order is performed on $g(x, y, t)$ in the local proximity, it may also be written as

$$g(x + \Delta x, y + \Delta y, t + \Delta t) \approx g(x, y, t) + (\nabla_{xyt} g(x, y, t)) \cdot (\Delta x, \Delta y, \Delta t). \tag{4.4}$$

Combining these two equations establishes that the first-order correction must be

$$(\nabla_{xyt} g(x, y, t)) \cdot (\dot{x}, \dot{y}, 1) = 0, \tag{4.5}$$

which is a simpler form of the continuity equation, see Eq. (4.2). This is, for example, the backbone of the famous Lucas–Kanade optic flow algorithm [80, 20].

Hessian constraint

Similarly to the previous optic flow constraint equation, but using a different quantity $q = \nabla_{xy} \cdot f$ under the same assumption that $\nabla_{xy}\vec{v} = \vec{0}$ directly yields the Hessian constraint equation

$$\nabla\partial_t g(x, y, t) + H_{\mathrm{g}}\vec{v} = 0 \tag{4.6}$$

where H_{g} is the Hessian matrix of the scalar-valued function g. The Hessian constraint equation was developed in the work of Uras et al. [131].

Phase constancy constraint

Instead of using the brightness attribute in an image, Fleet and Jepson suggested using the image phase attribute [41]. This changes the quantity $q = \phi$ and using the previous constraint $\nabla_{xy}\vec{v} = \vec{0}$ results in

$$(\dot{x}, \dot{y}, 1) \cdot \nabla_{xyt}\phi(x, y, t) = 0. \tag{4.7}$$

This approach is practical and feasible because Gabor filters may be used to compute the image phase ϕ, thus creating an alternative estimation of optic flow.

While more than 200 listed approaches advance the estimation of optic flow and improve existing methods, most traditional estimators can be traced back to the continuity equation, except for end-to-end estimators such as deep neural network approaches. This creates challenges when attempting to gain an overview of the landscape of proposed optic flow estimators. However, optic flow estimation benchmarks and leaderboards exist and may provide a general overview of state-of-the-art methods (see Section 4.3.2 for dataset providers and leaderboard maintainers).

4.3.1 Machine learning approach

Since the introduction of *convolutional neural network* (CNN) in the 1980s, various fields have accelerated in computer vision research, and not least in optical flow estimation research. CNN has been proven to be very suitable for image interpretation, such as shape detection, object recognition, and motion perception. Before the inception of CNN in computer vision, using multilayer perceptrons with hidden layers was computationally ineffective and underperforming due to the complexity in tuning weights between the layers and neurons. With CNNs, clustered pixel neighborhoods can be processed and considered a simultaneous input to the network layer. Furthermore, CNN is shift-invariant, meaning that apparent translational or rotational displacement of the kernel can help the filter to generalize learned patterns.

In recent years, the machine learning approach has dominated optic flow estimation research and outperformed in accuracy, precision, and run-time. There are several factors contributing to this, especially synergistic development and research in deep neural networks. Like most deep neural network technologies, CNN greatly benefits from specialized hardware such as GPUs, which excel in parallel data processing and computation. The rapid development of modern GPUs specializing in neural network computing and tensor computing has accelerated the performance of CNNs and further widened the performance gap. This can, for example, be observed in the benchmark leaderboards of popular datasets as discussed in Paper C, where the top performing algorithms are mainly machine learning approaches [47, 85, 86, 5, 21].

4.3.2 Benchmarking performance and evaluation

It is common in the computer vision community to compare their proposed algorithms and estimators. In the case of optic flow, the common performance metrics are *average end-point error* (AEE) and *average angular error* (AAE). These error metrics were first introduced and used by Fleet, Jepson, and Barron et al. [41, 7] and have become the *de facto* standard when evaluating optic flow estimators, see Fig. 4.1. The AEE is computed using Euclidean distance between the estimated flow field $\vec{Q}_{\text{est}}(\vec{p})$ and the ground truth $\vec{Q}_{\text{gt}}(\vec{p})$ as follows

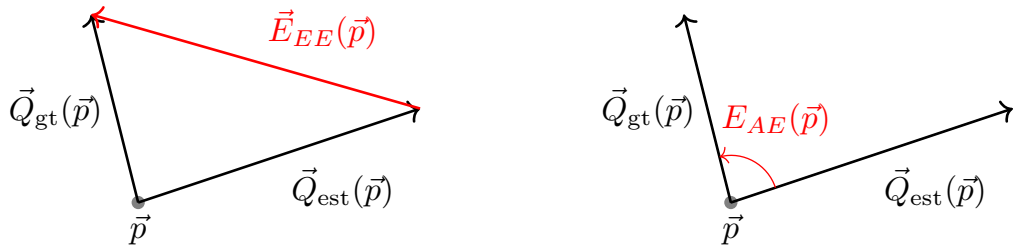
$$E_{\text{AEE}} = \frac{1}{N}\sum_{\vec{p}} ||\vec{E}_{\text{EE}}(\vec{p})||_2 = \frac{1}{N}\sum_{\vec{p}} ||\vec{Q}_{\text{est}}(\vec{p}) - \vec{Q}_{\text{gt}}(\vec{p})||_2 \tag{4.8}$$

where N is the number of sample points $\vec{p}$, typically the image area for dense optic flow. Similarly, for the AEE under the assumption that neither the estimated nor the ground truth vectors are null, the mean of the cosine angles is computed as

$$E_{\text{AAE}} = \frac{1}{N}\sum_{\vec{p}} E_{\text{AE}}(\vec{p}) = \frac{1}{N}\sum_{\vec{p}} \arccos \frac{\vec{Q}_{\text{est}}(\vec{p}) \cdot \vec{Q}_{\text{gt}}(\vec{p})}{||\vec{Q}_{\text{est}}(\vec{p})||_2 ||\vec{Q}_{\text{gt}}(\vec{p})||_2}. \tag{4.9}$$

These optic flow metrics are widely used to benchmark available published data sets of various types (synthetic and experimental) for computer vision research groups and developer communities. Some of these established and maintained data sets with public leaderboards are *KITTI* [47, 85, 86], *Middlebury* [5], *Sintel* [21], and *Reeds* [10, 118, 18]. These data sets are used in Paper C to investigate and address the problem of fair performance evaluation of algorithms and real-time computing, see Section 4.2.1 and Table 4.1.

As investigated in Paper C, performing a *fair* run-time evaluation is challenging in computer vision research. This is due to the complexities of executing binaries and their interaction with the hardware during run-time. Thus, it is an acknowledged problem in the scientific community without an established or standardized



(a) End-point error. (b) Angular error.

Figure 4.1: Opic flow error metrics. The panels illustrate two standard methods of evaluating optic flow performance: end-point error $E_{EE}(\vec{p})$ shown in (a) and angular error $E_{AE}(\vec{p})$ shown in (b) evaluated at the point $\vec{p}$.

methodology to perform a fair and comparable run-time evaluation. Researchers and software developers developing novel software often accumulate *technical debt*, which frequently results in degradation in software quality and neglect of software deployment [1]. Traditionally, computational complexity analysis has been used to classify and evaluate different software methods. However, doing so for software for run-time execution or real-time computing, in general, would be challenging and arguably meaningless.

Nevertheless, observing the general trends in the optic flow estimation leaderboards reveals that methods continuously improve accuracy and runtime, sometimes without compromising the two. Simultaneously, machine learning approaches dominate the top-performing methods in the optic flow category on the leaderboards. This phenomenon is not only limited to optic flow but is the general case for most categories, if not all.

To summarize, most of the traditional naive optic flow estimation methods can be traced back to the 'continuity equation' for optic flow, similar to that in physics. The different estimation methods make some assumptions that transform the continuity equation into a straightforward computational form. In contrast, end-to-end black box approaches like convolutional neural networks outperform their traditional counterpart. This can be seen in well-established benchmarking data sets with public leaderboards such as KITTI and Sintel. End-point and angular error averages are well-adopted performance benchmarking metrics used in optic flow estimation research.

4.4 Optic flow-based odometry

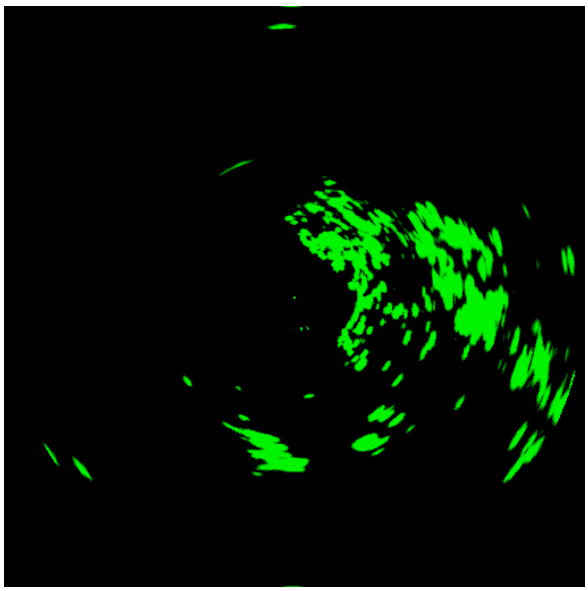
Inspired by humans using visual motion flow perception to estimate their egomotion, a similar approach may be exploited to solve the localization problems in robotics. By tracking the features of the environment, it is possible to deduce the relative motion of the sensor and its environment. This can be done in various ways, but the most common approach is to use *epipolar geometry* to determine egomotion. This can be done using either temporal (mono) or spatial (stereo) pairs of images and a camera model to determine the intersection of features. Optic flow in this context is primarily used to track the features and, by extension, determines the time-varying epipolar lines of these features. The approximate intersections of the epipolar lines determine the trajectory of the camera. In the case of mono-vision systems, the scaling in estimating the egomotion is lost due to the process of projective transformation. The scaling can be preserved if additional depth information is available, e.g. stereo vision or complementary information from lidars, radars, or GNSS. The epipolar geometry odometry using optic flow and mono-vision camera system is investigated in Paper B, particularly how well the method performs for marine vehicles in littoral settings.

In contrast to using epipolar geometry to determine egomotion, Paper E demonstrates a novel method, *topographic flow*, which applies optic flow on unconventional topographic data instead of visual data. The method strictly depends on the topographic data, thus making it agnostic to specific sensor systems. The topographic data in Paper E is constructed from a fully rotating active sensing radar system. The emitted radar beams from the sensor are partially reflected off the objects and registered. The non-reflected part penetrates the object and reflects upon the next occluded object, allowing for multiple object detections from a single emitted ray beam, see Fig. 4.2. However, the topographic data may be derived from any sensor data capable of reliably capturing the surrounding environment.

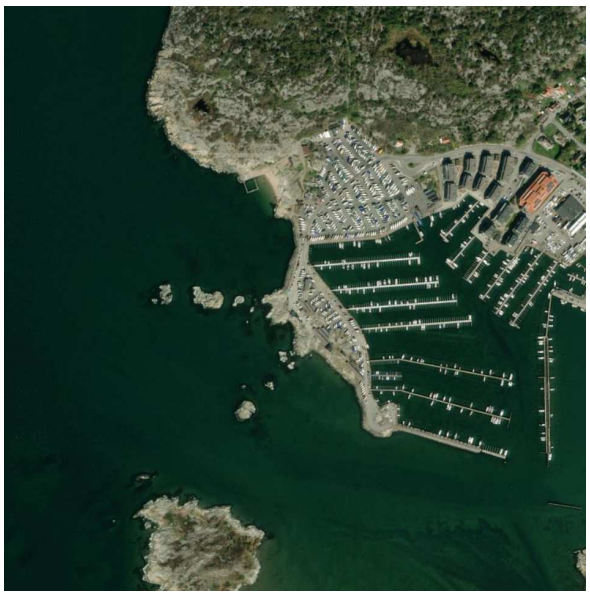
The topographic flow is based on the motion analysis of the emergent optic flow or simply analyzing the translational and rotational transformations of the topographic map. The emergent optic flow described in Eq. 2.1 for a perpendicular plane (to the viewpoint) may be reduced to three degrees of freedom: longitudinal velocity v_x, lateral velocity v_y, and angular yaw rate ω_z of the vehicle. This results in the following equation

$$\vec{Q}(x,y) = \begin{pmatrix} \dot{x} \\ \dot{y} \end{pmatrix} = -C \begin{pmatrix} v_x \\ v_y \end{pmatrix} - \omega_z \begin{pmatrix} -y \\ x \end{pmatrix} \tag{4.10}$$

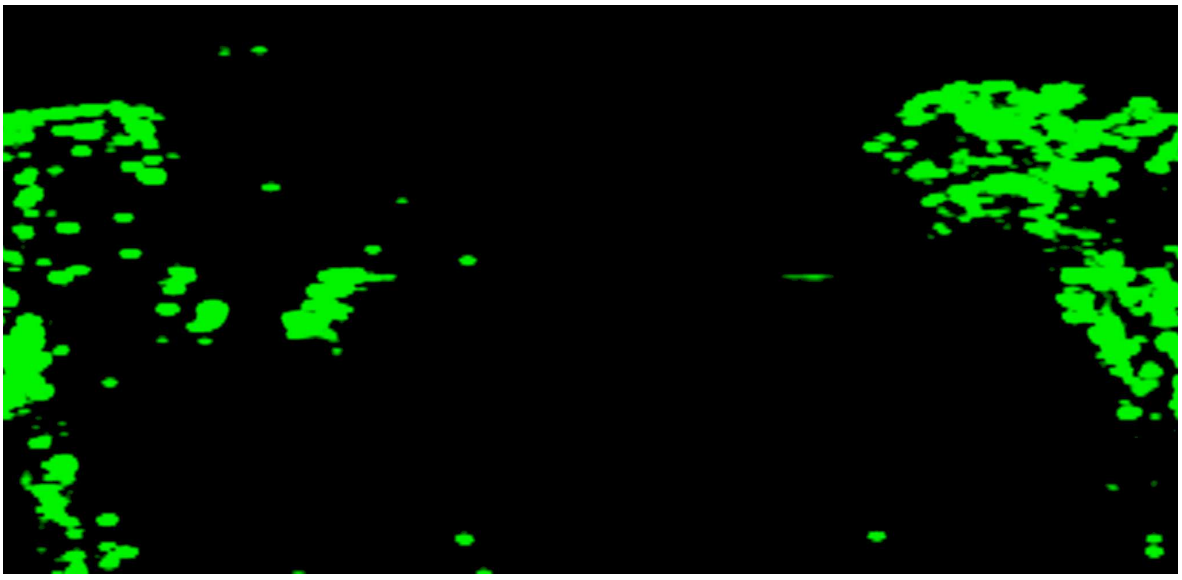
where C is some scaling constant relating the real world to the image space, (x, y) are coordinates in the image space aligned with the coordinate system of the vehicle, $(\dot{x}, \dot{y})$ is the optic flow field at (x, y). Eq. (4.10) is applied for a set of landmarks on the topographic map, computing their flow field for each temporal update of the map, yielding an egomotion estimate of the vehicle.



(a) Cartesian topography data.



(b) Satellite image.



(c) Cylindrical topography data.

Figure 4.2: Topography data constructed from radar. The figure shows the topography data used for optic flow-based odometry. The topographic map constructed from radar data is depicted in (a) a Cartesian coordinate system and (c) a cylindrical coordinate system. A satellite image of the corresponding island region outside of Göteborg is shown in (b). The satellite image is used under the OpenStreetMap ODbl license, found at `www.openstreetmap.org/copyright`.

To summarize, optic flow techniques can be leveraged to address the localization problem commonly encountered in robotic navigation. Odometry estimation can be carried out in various ways, with optic flow primarily used for feature tracking. This is then used to determine egomotion through epipolar geometry or topographic transformation, for example. Due to the nature of projective transformation in visual cameras, the depth information is often lost, resulting in arbitrary scaling of unit lengths. However, lengthless unit estimates such as angular velocities and orientations do not suffer from this problem.

4.5 Retinal optic flow estimation

Further extending optic flow to retinal optic flow may be carried out by accounting for the human gazing kinematics, see Section 2.2. This has been made possible thanks to the rapid development and maturity of modern wearable eye-tracking systems. One naive numerical estimate of the retinal optic flow is to combine the head-fixed optic flow estimation with the gazing measurement of the agent. Then, the retinal optic flow using head-fixed optic flow can be formulated as

$$\vec{Q}_{\text{rof}}(\vec{Q}_{\text{of}}, \vec{p}, \vec{p}_{\text{g}}) = \vec{Q}_{\text{of}}(\vec{p}) - \vec{Q}_{\text{of}}(\vec{p}_{\text{g}}) \tag{4.11}$$

where $\vec{Q}_{\text{rof,of}}$ are the vector fields for retinal optic flow and optic flow respectively, $\vec{p}$ is some point in the vector field, and $\vec{p}_{\text{g}}$ is the gaze point. An example illustration and visualization of the resulting retinal optic flow field is shown in Fig. 4.3. This formulation assumes that the agent achieves *exact* smooth pursuit during active fixations, as it significantly simplifies computations. As a result of the assumption and the formulation of retinal optic flow, the flow field around the gaze point is

$$\vec{Q}_{\text{rof}}(\vec{Q}_{\text{of}}, \vec{p}_{\text{g}}, \vec{p}_{\text{g}}) = \vec{0}. \tag{4.12}$$

This stems from the fact that the retinal optic flow at the point of gaze during perfect smooth pursuit is null. Given the vector field and a point of interest, the flow angle can be computed as

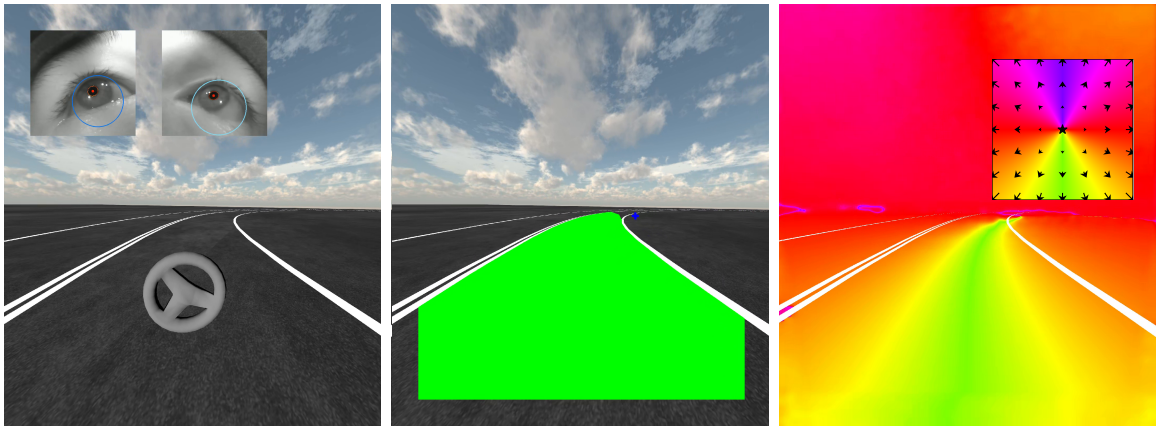
$$\theta(\vec{Q}, \vec{p}) = \arctan \frac{\vec{Q}(\vec{p}) \cdot \vec{e}_y}{\vec{Q}(\vec{p}) \cdot \vec{e}_z} \tag{4.13}$$

where $\vec{Q}$ is the vector field, e.g. retinal optic flow, $\vec{p}$ is the point of measurement, and $\vec{e}_y$ and $\vec{e}_z$ are the orthonormal basis vectors in the image space. This quantity can be used to form a judgment or estimation of the locomotor performance of the agent, as demonstrated in Paper A, inspired by the retinal flow patterns suggested by Kim et al. [64] and Wann et al. [138]. A naive approach to this is to compute

the 'spatial circular mean angle of the retinal optic flow over the intended path' as

$$\bar{\theta}_{\text{rof}} = \arg\Big(\sum_{\theta_{\text{rof}} \in \Omega_{\text{lane}}} \exp(i\theta_{\text{rof}}) \Big), \tag{4.14}$$

where θ_{rof} is the directional angle of the dense retinal optic flow field (exemplified in Fig. 4.3c), Ω_{lane} is the set containing the dense visual flow of the visual lane (corresponds to the green area in Fig. 4.3b), and i is the imaginary unit. For simplicity, 'spatial circular mean angle of the retinal optic flow over the intended path' will now be referred to as *retinal optic flow angle* for the remainder of this work.



(a) Point of view. (b) Lane and gaze point. (c) Retinal optic flow.

Figure 4.3: Visual view, computer vision and image processing, data visualization and representation. An example of retinal optic flow computed using experimental data in virtual reality. (a) shows the point of view in virtual reality. (b) visualizes the post-processed data rendered on top of the raw image. The gaze point is illustrated as a blue crosshair, and the lane segment as the green area. (c) demonstrates the reconstructed retinal optic flow by combining motion analysis of consecutive temporal images and sampled gaze data. The reference color coding of pure divergence is shown in the top right corner as a reference (the visualization makes left and right directions symmetric). The green color is equivalent to the retinal optic flow pointing in the downward direction.

To summarize, a naive numerical computational viable approach to quantify and reconstruct the retinal optic flow field is presented using measured gazes, head-fixed vision, and optic flow estimation. This approach assumes that the agent achieves exact pursuit during active gazing, significantly simplifying the computation. In addition, the derived quantity retinal optic flow angle is introduced, which may be used to perceive the locomotor performance of the agent suggested in the literature.

4.6 Identifying ballistic reaching corrections

The natural movements of the limbs exerted by humans are often characterized by reaching, as discussed in Section 3.2, where the velocity profile of the limb is bell-shaped and the positioning is sigmoidal. In reality, the reaching movements are seldom very clean, i.e. consisting of one single ballistic reaching correction, as humans intermittently assess and correct performance on-the-fly according to their task at hand, see Section 2.2.2 and 3.3. This results in the naturalistic limb movements often being convoluted and complex, consisting of multiple superposed ballistic reaching corrections.

To identify the constituent ballistic correction using Eq. (3.3) and (3.9) in Paper A, the following objective function to be minimized may be used

$$f(z(t), \Omega_{\dot{\delta}}) = \frac{\text{cols}(\Omega_{\dot{\delta}})^{\alpha}}{\Delta t} \int_{t_i}^{t_i+\Delta t} |z(t) - \sum_{(a_i,\mu_i,\sigma_i)\in\Omega_{\dot{\delta}}}^{N_\delta-1} \dot{\delta}_{\text{b}}(t, a_i, \mu_i, \sigma_i)| \text{d}t, \tag{4.15}$$

where

$$\Omega_{\dot{\delta}} = \begin{pmatrix} a_0 & a_1 & \dots & a_{N_\delta-1} \\ \mu_0 & \mu_1 & \dots & \mu_{N_\delta-1} \\ \sigma_0 & \sigma_1 & \dots & \sigma_{N_\delta-1} \end{pmatrix} \tag{4.16}$$

is the matrix of N_δ triplet parameters describing each constituent ballistic correction in a reaching movement, $z(t)$ is the measured movement signal to be decomposed for a defined time interval $t \in [t_i, t_i + \Delta t]$, and α is a balancing parameter for the number of corrections to avoid overfitting. Note that $\text{cols}(\Omega_{\dot{\delta}}) = N_\delta$, but here the left-hand side is the preferred form as the dimensionality of the matrix is also under consideration for optimization. The resulting object score is intentionally normalized for the time duration in order to make a comparable goodness fit for different movement patterns of various durations.

In the context of human driving behavior and under assumptions of normal steering handling with sufficiently small steering wheel angles, the speed parameter σ_I as defined in Eq. (3.3) was empirically found to be bounded $\sigma_{\text{I}} \in [0.0707\,\text{s}, 0.1732\,\text{s}]$ which was used in Paper A. This was initially implicitly suggested and used in the works of Benderius and Markkula when studying human driving behavior [11, 81]. By these empirical constraints, one single ballistic steering correction movement duration is then within 0.28 s to 0.69 s.

The corrections applied to the steering wheel are preliminarily selected on the criteria introduced in the works of Benderius and Markkula [11]. Here, the example of the steering wheel movement is used, but the generality of the application still holds in one dimension. A simple moving average with a constant window determines the time intervals $t \in [t_i, t_i + \Delta t]$ where the significant movement is

detected. It is described as follows

$$\bar{\dot{\delta}}(t) = \frac{1}{2\epsilon} \int_{t-\epsilon}^{t+\epsilon} \dot{\delta}(\zeta) \mathrm{d}\zeta \tag{4.17}$$

and for the discrete case

$$\bar{\dot{\delta}}(t_n) = \frac{1}{N} \sum_{k=\frac{1-N}{2}}^{\frac{N-1}{2}} \dot{\delta}(t_{n+k}), \tag{4.18}$$

where $\dot{\delta}(t)$ is the steering wheel rate and 2ϵ is the kernel width (or N in the discrete case, chosen odd). By checking the contiguous fulfillment of the condition in time

$$\bar{\dot{\delta}}(t) \geq \dot{\delta}_{\text{still}} \tag{4.19}$$

where δ_{still} is some defined constant threshold for tuning detection sensitivity, yields the starting time t_i and the duration Δt of the movement. The deflection angle, which is defined as

$$\Delta\delta = \max_{t \in [t_i, t_i + \Delta t]} (\delta(t)) - \min_{t \in [t_i, t_i + \Delta t]} (\delta(t)), \tag{4.20}$$

is used to determine if the overall steering movement is significant, removing negligible and small movements, i.e.

$$\Delta\delta \geq \delta_{\text{min}} \tag{4.21}$$

where δ_{min} is a parameter. Furthermore, an additional condition is imposed in the case of the steering wheel to remove complex hand-to-hand coordination by using

$$\max_{t \in [t_i, t_i + \Delta t]} (|\delta(t)|) < \delta_{\text{max}} \tag{4.22}$$

where δ_{max} is a parameter. Numerical parameters used in this work and Paper A to detect steering movements of interest are presented in Table 4.2.

Table 4.2: Parameters used for detecting and selecting steering movements in raw steering angle data.

Parameter	Value	Note
$\dot{\delta}_{\text{still}}$	0.0349 rad	Detect movement
N	3	Discrete kernel width
δ_{min}	0.0349 rad	Remove smaller movements
δ_{max}	1.7453 rad	Remove larger movements

4.6.1 Particle swarm optimization

Finding the optimal solutions, if they exist, among all feasible solutions is known as an optimization problem. The optimization problem can be written in the following form

$$\arg\min_{\vec{x}} \quad f(\vec{x}) \tag{4.23}$$

$$\text{subject to} \quad g_i(\vec{x}) \leq 0, \quad i = 0, \ldots, m-1 \tag{4.24}$$

$$h_j(\vec{x}) = 0, \quad j = 0, \ldots, p-1 \tag{4.25}$$

where $\vec{x}^*$ is the solution, $f(\vec{x})$ is the objective function to be minimized, and $g_i(\vec{x})$ and $h_j(\vec{x})$ are some known constraints. For trivial reasons, such as easier implementation, it might be more suitable to invert the problem to maximize the objective function instead of minimizing it. Then the optimization problem becomes

$$\arg\max_{\vec{x}} \frac{1}{f(\vec{x})} = \arg\min_{\vec{x}} f(\vec{x}), \quad f(\vec{x}) \neq 0. \tag{4.26}$$

This thesis will consider the optimization problem by maximizing the objective function.

Approaching and solving the optimization problem of decomposing the sum of Gaussian functions is a complex and challenging task as defined in Eq. (4.15), i.e.

$$\arg\max_{\Omega_{\dot{\delta}}} \frac{1}{f(z(t), \Omega_{\dot{\delta}})}. \tag{4.27}$$

One contributing factor is that N_δ is not known, which determines the dimension of the search space and is itself part of the optimization problem. In addition, local optima may exist in different search spaces, and the objective function is not guaranteed to be continuous or differentiable. These challenges make common classical optimization approaches unfeasible or even impossible.

Particle swarm optimization (PSO) is a meta-heuristic bio-inspired optimization method inspired by and exploiting the emergent self-organized phenomenon of social collective flocking behavior observed in the real world when groups of living animals emerge e.g. murder of crows or school of fish. This flocking behavior is illustrated in Fig. 4.4 using flocking particles or artificial life programs called *Boids* (bird-oid object) first introduced by Craig Reynolds in 1986 [110, 63]. This stochastic optimization approach emulates Boids interacting locally with each other and further encodes their positions in the multidimensional search space. This approach has shortfalls as the yielded proposed solution $\vec{x}^*$ is not guaranteed to be a global optimum and is computationally inefficient compared to the other classical optimization methods. In return, the PSO does not impose many strict formal requirements on the objective function or the search space.

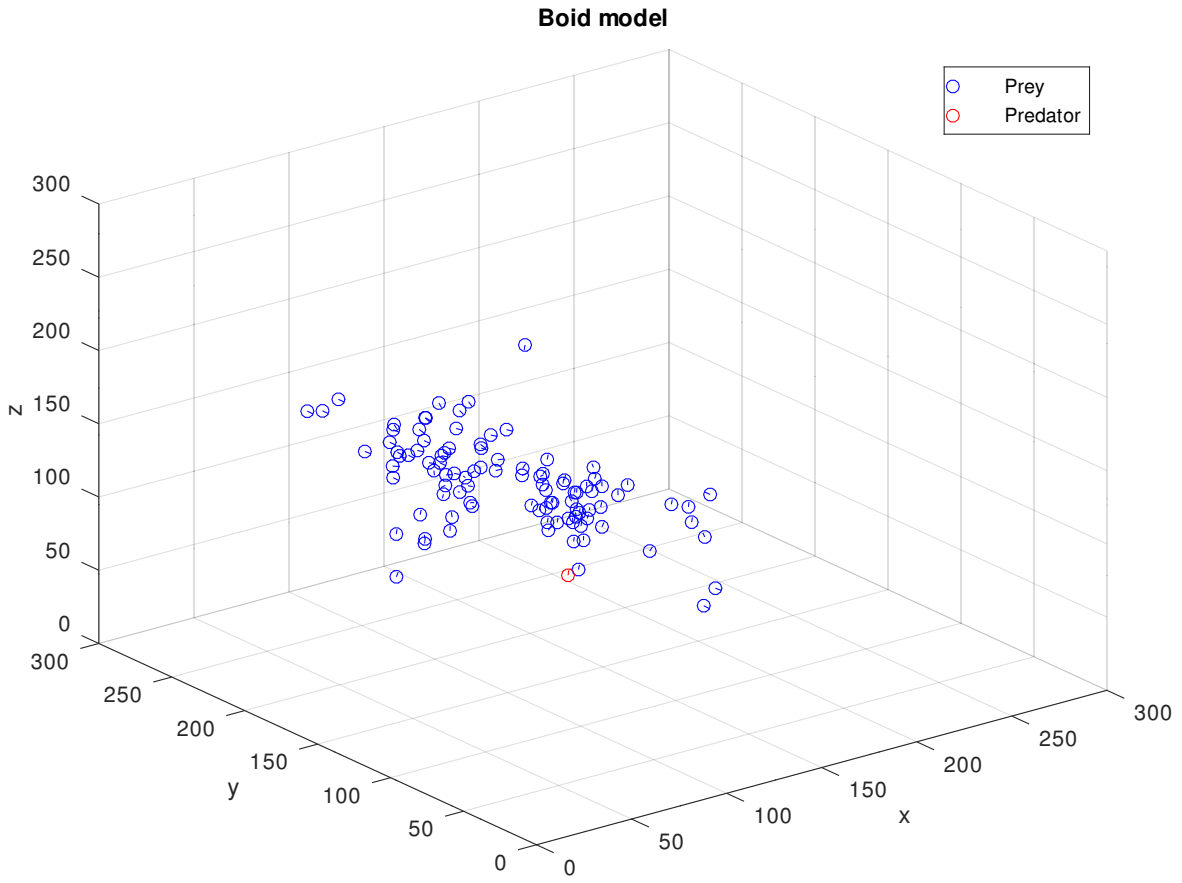


Figure 4.4: Emergent flocking behavior. The Boid model demonstrates the emergent flocking behavior using three simple rules: *separation*, *alignment*, and *cohesion*. This model further features a predator that drives the flock around.

Essentially, the swarms of particles exhaustively explore the multidimensional search space with emulated mass inertia. The emergent swarming behavior can be created by imposing three simple rules on each particle: i. *Separation*: avoid crowding, ii. *Alignment*: aligning the ego velocity to the local group, iii. *Cohesion*: creating a local swarm attraction. Additional tweaks to the implementation encourage the swarms to *explore* and to then *exploit* the search space in finding the optimum for the objective function. To favor exploration in the initial phase, the particles are made individualistic and asocial to other particles (ignoring) and then gradually made more social. This creates the emerging social swarming behavior to hone in on the convergence of a candidate solution by flocking around the optima.

To promote faster convergence in PSO, symmetry in the defined search space is exploited when solving the problem of decomposing the sum of Gaussian functions. This is due to the principle of superposition having *commutative* property, thus multiple different solutions are trivially effectively the same. For example, consider the following

$$\Omega_{\delta,0} = \begin{pmatrix} a_0 & a_1 \\ \mu_0 & \mu_1 \\ \sigma_0 & \sigma_1 \end{pmatrix} \quad \text{and} \quad \Omega_{\delta,1} = \begin{pmatrix} a_1 & a_0 \\ \mu_1 & \mu_0 \\ \sigma_1 & \sigma_0 \end{pmatrix} \tag{4.28}$$

which both solves

$$f(z(t), \Omega_{\delta,0}) = f(z(t), \Omega_{\delta,1}). \tag{4.29}$$

Using this fact, sorting the columns in Ω by μ_i, see Eq. (4.16), in ascending (alternatively descending) such that

$$\mu_0 \leq \mu_1 \leq \cdots \leq \mu_{N_{\delta-1}} \tag{4.30}$$

reduces the search space complexity. This effectively decreases the total number of local optima by merging trivially duplicated symmetric solutions.

Paper A successfully demonstrates and implements[1] [94] the standard PSO implementation [135] with additional tweaks to identify ballistic corrections by solving Eq. (4.15), see Fig. 4.5 on synthetic data and Fig. 4.6 on experimental human-operated steering wheel data. The PSO algorithm is further detailed as pseudo-code in Algorithm 1, and the parameter values shown in Table 4.3.

To summarize, complex and superimposed reaching movements can be decomposed into constituent ballistic corrections using the assumption of the superposition principle. The velocity profile of each ballistic correction can be approximated using the Gaussian membership function of three parameters. The approximation makes it possible to identify each correction in the experimental data of complex reaching movements. The particle swarm optimization, a bio-inspired meta-heuristic stochastic optimization method, was used to solve the decomposition of overlapping reaching movements. This novel approach yields numerical values for the three parameters for each constituent ballistic correction.

[1] Open source C++ implementation: `https://github.com/bjornborg/tinyso`

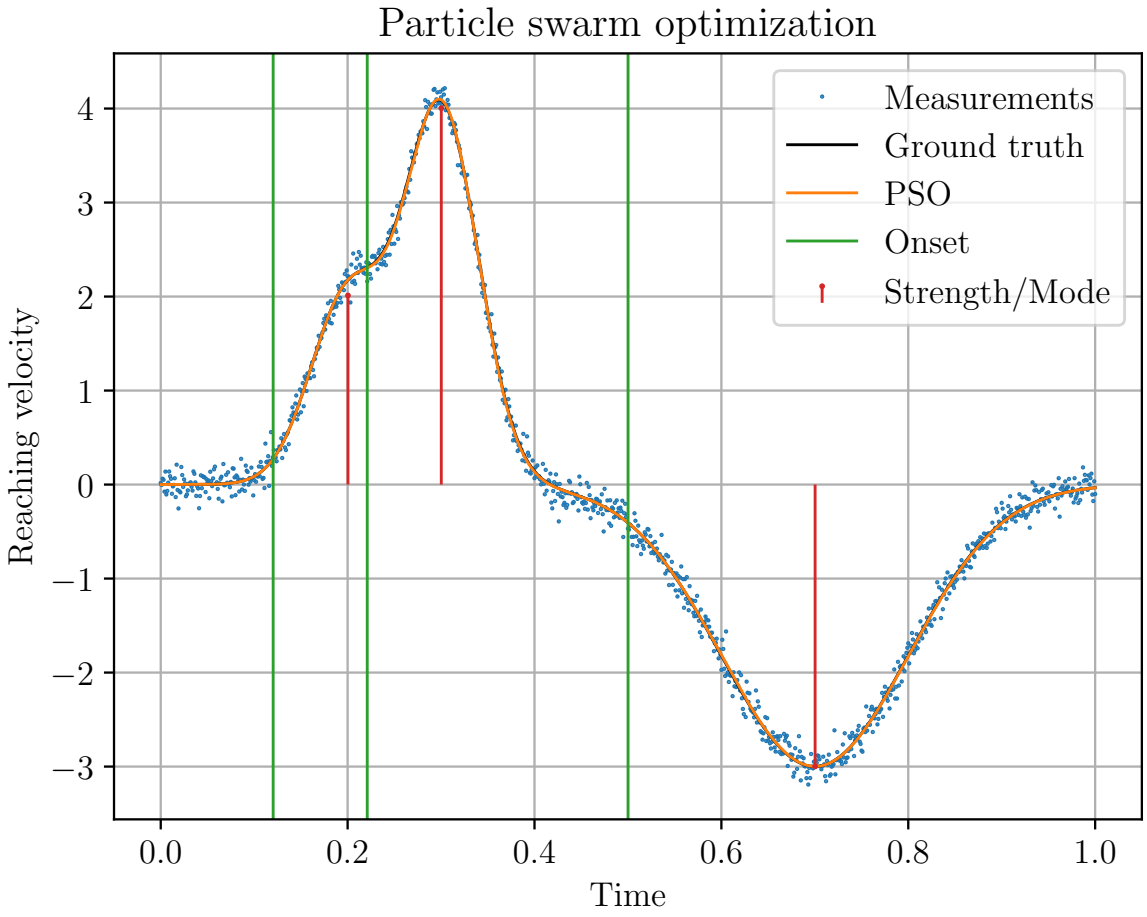


Figure 4.5: Identifying ballistic reaching correction on synthetic data. The figure demonstrates particle swarm optimization successfully identifying ballistic reaching correction in synthetic measurement data padded with normally distributed noise $n(t) \sim \mathcal{N}(0, 0.1)$.

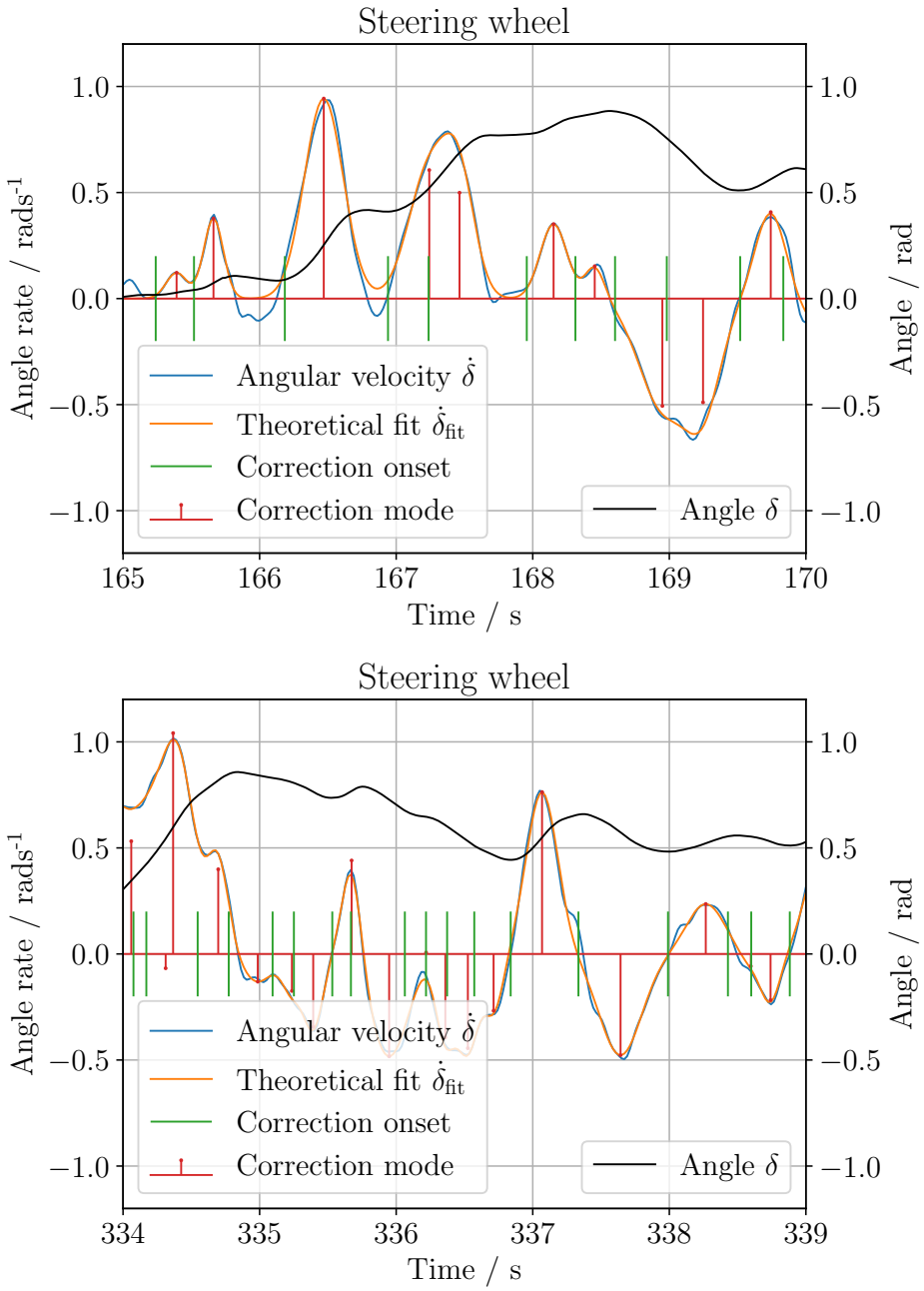


Figure 4.6: Identifying ballistic reaching correction on experimental data from humans. Particle swarm optimization is used to identify ballistic reaching correction in experimental data on a human-operated steering wheel.

Algorithm 1 Particle swarm optimization algorithm with.

Require: $S_{\min} > 0, N_{\delta,\max} > 0, \vec{x}_{\min} < \vec{x}_{\max} \in \mathbb{R}^3, g(\vec{x}) : \mathbb{R}^{3N_\delta} \to \mathbb{R}$
Ensure: $\mathbb{R}^{3N_\delta} : \text{repeat}(\vec{x}_{\min}, N_\delta) < \vec{x}^* < \text{repeat}(\vec{x}_{\max}, N_\delta), g(\vec{x}^*) > 0$

procedure ParticleSwarmOptimization($z(\vec{t}), \vec{x}_{\min}, \vec{x}_{\max}$)
  Converged ← False
  $N_\delta \leftarrow 1$
  while not Converged **do**
    $\mathbb{S} \leftarrow$ InitializeSwarm($N, \vec{x}_{\min}, \vec{x}_{\max}$) ▷ Implementation in Alg. 2
    for $n \leftarrow 0, \ldots, K-1$ **do**
      $\mathbb{S} \leftarrow$ EvaluateSwarm($\mathbb{S}, z(\vec{t})$) ▷ Implementation in Alg. 3
      $\mathbb{S} \leftarrow$ UpdateVelocityPosition($\mathbb{S}$) ▷ Implementation in Alg. 4
      $w \leftarrow \text{Clamp}(\beta w, w_{\min}, w_{\max})$ ▷ Decay particle inertia
    end for
    Criteria ← $(p_{\text{candidate}}.s_{\text{score}} > \mathbb{S}.p_{\text{best}}.s_{\text{best}})$ ▷ Swarm best
    Criteria ← Criteria **and** $p_{\text{candidate}}.s_{\text{score}} > S_{\min}$ ▷ Threshold score
    Criteria ← Criteria **or** $N \geq N_{\delta,\max}$ ▷ Maximum iterations
    if Criteria is met **then**
      Converged ← True
    else
      $p_{\text{candidate}} \leftarrow \mathbb{S}.p_{\text{best}}$ ▷ Store the candidate solution
      $N_\delta \leftarrow N_\delta + 1$ ▷ Increase the dimensionality of search space
    end if
  end while
  return $\vec{x}^* \leftarrow p_{\text{candidate}}.\vec{x}$
end procedure

Algorithm 2 Particle swarm optimization: Initialization.

Require: $N_\delta > 0, \vec{x}_{\min} < \vec{x}_{\max}$

function InitializeSwarm($N_\delta, \vec{x}_{\min}, \vec{x}_{\max}$)
  $\vec{x}_{\min} \leftarrow \text{repeat}(\vec{x}_{\min}, N_\delta)$ ▷ Repeat the array, length $3N_\delta$
  $\vec{x}_{\max} \leftarrow \text{repeat}(\vec{x}_{\max}, N_\delta)$
  for $p_i \in \mathbb{S}$ **do** ▷ $\vec{r} = (r_0, \ldots, r_{3N_\delta - 1}), r_i \in [0, 1]$
    $\vec{x}_i \leftarrow \vec{x}_{\min} + \vec{r} \odot (\vec{x}_{\max} - \vec{x}_{\min})$ ▷ Hadamard product $\odot$
    $\vec{v}_i \leftarrow \gamma \mathrm{d}t^{-1}(\frac{\vec{x}_{\min} - \vec{x}_{\max}}{2} + \vec{r} \odot (\vec{x}_{\max} - \vec{x}_{\min}))$
    $p_i \leftarrow (\vec{x}_i, \vec{v}_i)$
  end for
  return $\mathbb{S}$
end function

Table 4.3: Parameters for the particle swarm optimization and the objective function. The table presents the numerical parameter values used in the particle swarm optimization and the objective function for identifying the ballistic steering wheel corrections.

Parameter	Value	Note
N_P	10000	Swarm population size
K	400	PSO max steps
α	1.0	Curve fitting parameter
γ	1.0	Time precision factor
$\mathrm{d}t$	1.0	Discrete time step size
v_max	1.1	Particle speed limit
w_max	1.4	Max particle inertia for exploration
w_min	0.3	Min particle inertia for exploitation
β	0.991	Inertia decay factor
c_1	3.0	Cognitive factor
c_2	$4.0 - c_1 = 1.0$	Social factor
K	400	Maximum number of steps in each PSO iteration
$N_{\delta,\mathrm{max}}$	6	Maximum Gaussian components for N_δ
S_min	200	Minimum fitness

Algorithm 3 Particle swarm optimization: Evaluation

Require: $\mathbb{S}, f(\vec{x})$ ▷ Swarm, objective function

function EVALUATESWARM($\mathbb{S}, z(t)$)
  for $p_i \in \mathbb{S}$ **do**
    $\vec{x} \leftarrow p_i.\vec{x}$
    $s \leftarrow g(z(t), \vec{x})$ ▷ Objective score see Eq. (4.15)
    if $s > p_i.s_\mathrm{best}$ **then** ▷ Store the personal best score
      $p_i.s_\mathrm{best} \leftarrow s$
      $p_i.\vec{x}_\mathrm{best} \leftarrow \vec{x}$
    end if
    if $s > \mathbb{S}.p_\mathrm{best}.s_\mathrm{best}$ **then** ▷ Store the swarm best score
      $\mathbb{S}_i.p_\mathrm{best} \leftarrow p_i$
    end if
  end for
  return $(S.p_\mathrm{best}.s_\mathrm{best})$
end function

Algorithm 4 Particle swarm optimization: Update step.

Require: $\mathbb{S}, c_1, c_2, v_{\max}, w$

function UPDATEVELOCITYPOSITION($\mathbb{S}$)
  $\vec{x}_{\text{sb}} = \mathbb{S}.p_{\text{best}}.\vec{x}$ ▷ Swarm best
  for $p_i \in \mathbb{S}$ **do**
    $\vec{x} \leftarrow p_i.\vec{x}$
    $\vec{x}_{\text{pb}} \leftarrow p_i.\vec{x}_{\text{best}}$ ▷ Personal best
    $\vec{v} \leftarrow p_i.\vec{v}$
    $\vec{v} \leftarrow w\vec{v} + c_1 r_1 \frac{\vec{x}_{pb} - \vec{x}}{\mathrm{d}t} + c_2 r_2 \frac{\vec{x}_{sb} - \vec{x}}{\mathrm{d}t}$ ▷ Uniform random $r_i \in [0, 1]$
    if $||\vec{v}||_2 > v_{\max}$ **then** ▷ Restrict velocity below a threshold
      $\vec{v} \leftarrow v_{\max} ||\vec{v}||_2^{-1} \vec{v}$
    end if
    $\vec{x} \leftarrow \vec{x} + \vec{v}\mathrm{d}t$ ▷ Euler forward method
    $p_i.\vec{x} \leftarrow \vec{x}$ ▷ Update particle
    $p_i.\vec{v} \leftarrow \vec{v}$
  end for
  return $\mathbb{S}$
end function

Chapter 5

Discussion

This chapter presents and discusses the main results of this thesis. The first section presents and discusses the reconstruction and estimation of retinal optical flow using experimental data collected from human research participants (**RQ1**). In the second section, the retinal optic flow may be related to steering maintenance when locomoting in curvilinear paths (**RQ1** and **RQ2**). The third section presents further results of characterizing heavy-tailed velocity profiles. In the last section, estimations of egomotion and self-localization in marine settings using optic flow are investigated and presented in the context of robotics and embedded systems (**RQ3** and **RQ4**).

5.1 Retinal optic flow for locomotor control

Since the retinal optic flow is the main topic of this thesis, there is an emphasis on reconstructing and interpreting the retinal optic flow. Much of the previous work on retinal optic flow for locomotion has primarily been discussed in theory [64, 146, 138, 140, 106, 139]. Most of these studies, although not limited to them, have heavily simplified the retinal optic flow. Matthis et al. [82, 84] are the first to reconstruct the retinal optic flow using naturalistic video data numerically and to study how the retinal optic flow may influence human locomotor abilities on foot. However, their study is carried out on *sparse* retinal optic flow fields, which limits their interpretation and analysis.

This work further extends the fields to be represented as dense vector fields, allowing for complete and more accurate interpretation and post-processing at a substantial increase in computational resource cost (see Paper C). This dense reconstruction of retinal optic flow is implemented and demonstrated in Paper A,

where research participants operated a road vehicle to reach a destination. Demonstration videos are available in open data [93] with the author *acting* as a research participant.

The retinal optic flow angle, as a naive interpretation of the retinal optic flow pattern (see Eq. (4.14)), can be used as a quantity for maintaining steering (**RQ1**). As shown in Paper A, the retinal optic flow angle seems to be more strongly associated with the steering correction in the low-latency response at 0.14 s compared to the heading case of 0.44 s. The findings raise interesting questions:

1. Are the triggering events fully independent of each other? If so, are the two cues fully operating in their own regimes, retinal optic flow-based cues in the low-latency and the heading-based cues in the high-latency?

2. Alternatively, are the retinal optic flow cues an outcome from corrections based on the heading cue or vice versa?

Further investigation using better variable control in experiments is needed to settle these discussions, to isolate heading-based cues and retinal optic flow-based cues from each other, and observe and determine the effects. The alternative is to show the existence of statistical significance through ANOVA on large data sets.

One of the direct artifacts and outcomes of visual flow is the FOE. Lappe et al. pointed out that many researchers often erroneously synonymously use FOE for egomotion heading by making the example of retinal optic flow [69]. The confusion is directly apparent when retinal image stabilization is achieved, as the FOE is located as the gaze point of the scene (at the fovea on the retina). Matthis et al. further investigated the role of the FOE in head-centered optic flow and retinal optic flow fields [84]. Their work studies the sparse patterns of integrated paths in the visual flow field, providing further insight into how humans perceive heading. They conclude and emphasize that the role of FOE and head-fixed optic flow should be re-evaluated in the context of human locomotion, which is in line with and supported by the findings in Paper A, research of Lappe et al. [69], and Land and Lee [68]. Continuing the work of Matthis et al., further assessment of the details in dense retinal optic flow field around the FOE may reveal how the visual patterns may be used for high-speed curvilinear locomotor control.

Although retinal optic flow was studied explicitly and mainly for steering behavior in Paper A, some interesting observations were made about the maintenance and evaluation of longitudinal control in humans. It is assumed, and arguably so, that retinal optic flow could be used to estimate the egomotion speed. However, in the absence of a speedometer and inability to experience forces in the virtual simulation, almost all the research participants kept their vehicle speed high at an average of $v_{\mathrm{l}} \approx 13\,\mathrm{m\,s^{-1}}$ in curve bends, which is equivalent to 0.57g lateral force. The high speed significantly contributed to the loss of vehicle control when research participants applied too large steering corrections, resulting in failed trials.

The initial observation challenges the notion that retinal optic flow in the visual periphery during high-speed context could accurately and properly be used for speed assessment or maintenance.

5.1.1 On the assumption of exact smooth pursuits

In estimating the retinal optic flow in this work, it is assumed that the research participants achieve exact smooth pursuit when they fixate their gaze on the scene. The assumption can be motivated by looking at the spatial gaze distribution of human experimental data from Paper A as illustrated in Fig. 5.1. It may be observed that the research participants tend to consistently and skillfully place their gaze point on their lane. This assumption of smooth pursuit significantly simplifies the computation described with Eq. (4.11) by using the fact that smooth pursuits achieve exact retinal image stabilization in the fovea, since this feature is captured and exploited in Eq. (4.12). As such, this idealization may heavily impact the validity of the result downstream in the analysis. Thus, it is required to further discuss the idealization of the smooth pursuit for retinal optic flow estimation.

The NSLR algorithm developed by Tuhkanen et al. [100] has been implemented[1] and applied to quantify and identify periods of fixed gazing velocities in raw gazing time series data, an example is shown in Fig. 5.2 for a small select time series for one research participant. The direct results yielded from NSLR may be used to extract and characterize the low-velocity smooth pursuit from saccades.

It was theoretically shown in Section 2.2.3 that during curvilinear motion, when the agent achieves an exact smooth pursuit, the yaw angular velocity of the vehicle is twice the horizontal gaze velocity. Using the smooth pursuit characteristics from NSLR and comparing them to the vehicular kinematics yields the optokinetic results presented in Fig. 5.3. All the data sampled are collected specifically during the curve bends, hence the dead band $[-0.2, 0.2]$ in vehicle yaw velocity. Due to the sensitivity of linear regression to outliers, *random sample consensus* (RANSAC) has been applied to reject outliers in the data [40]. The linear regression should be compared to the theoretically expected value of ± 2.0 (the sign depends on the definitions of coordinate systems). The slope of the linear regression of the experimental data is found to be lower, possibly caused by retinal slippage. The discrepancy implies that the gaze is slightly slipping or not entirely in sync with the yaw rotation velocity of the vehicle. However, the discrepancy is not large enough to invalidate the assumption of exact smooth pursuit for computing retinal optic flow.

[1] The C++ implementation is available at `https://gitlab.com/bjornborg/nslr`.

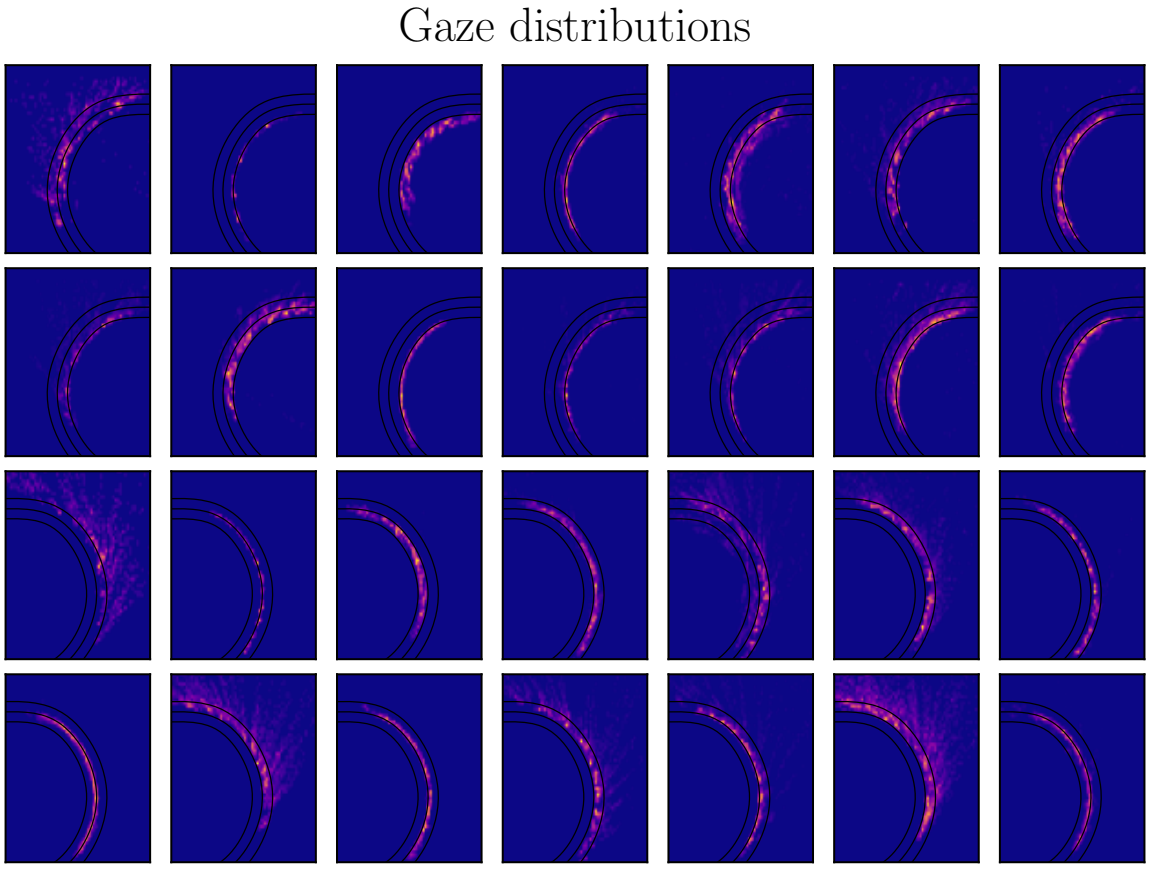


Figure 5.1: Human gaze distributions on road curve bends. The heat map of the gaze placement of human drivers maintaining mission-critical locomotion during left (bottom rows) and right curves (top rows). Each panel is the accumulated and aggregated data for each research participant and curve bend. Many research participants skillfully deploy the gazing strategy *"looking where you are going"*. This may be observed in the placement of the gaze points, which are often around the apex of the curve bends. Other research participants tend to fixate their gaze straight toward the horizon or where the vehicle is heading, creating straight lines in the heatmap or no registrations on the ground.

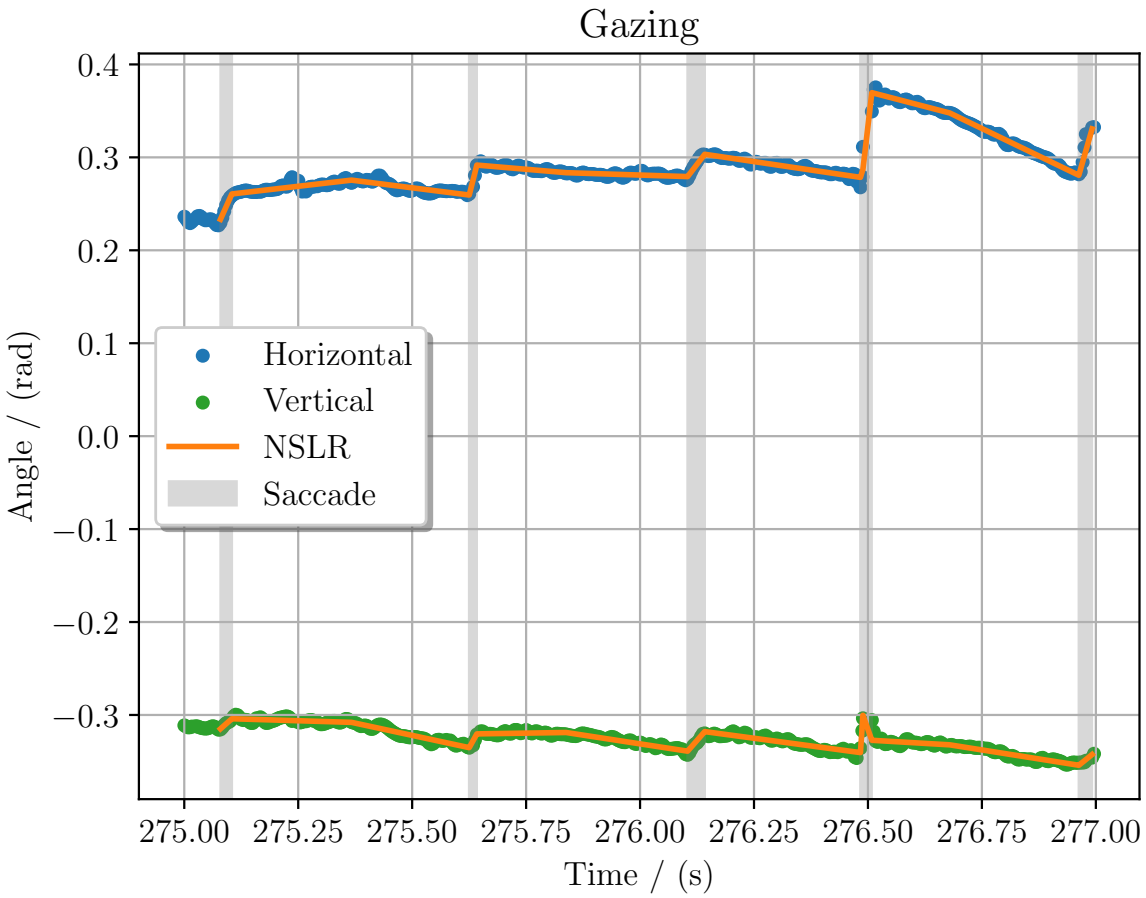


Figure 5.2: Gazing, smooth pursuits, and naive segmented linear regression (NSLR). The experimental data on human gazing are shown. The NSLR algorithm has been applied [100], yielding the defined periods of constant gazing speed, where periods of saccades have been highlighted. The algorithm is used to identify low-velocity smooth pursuits.

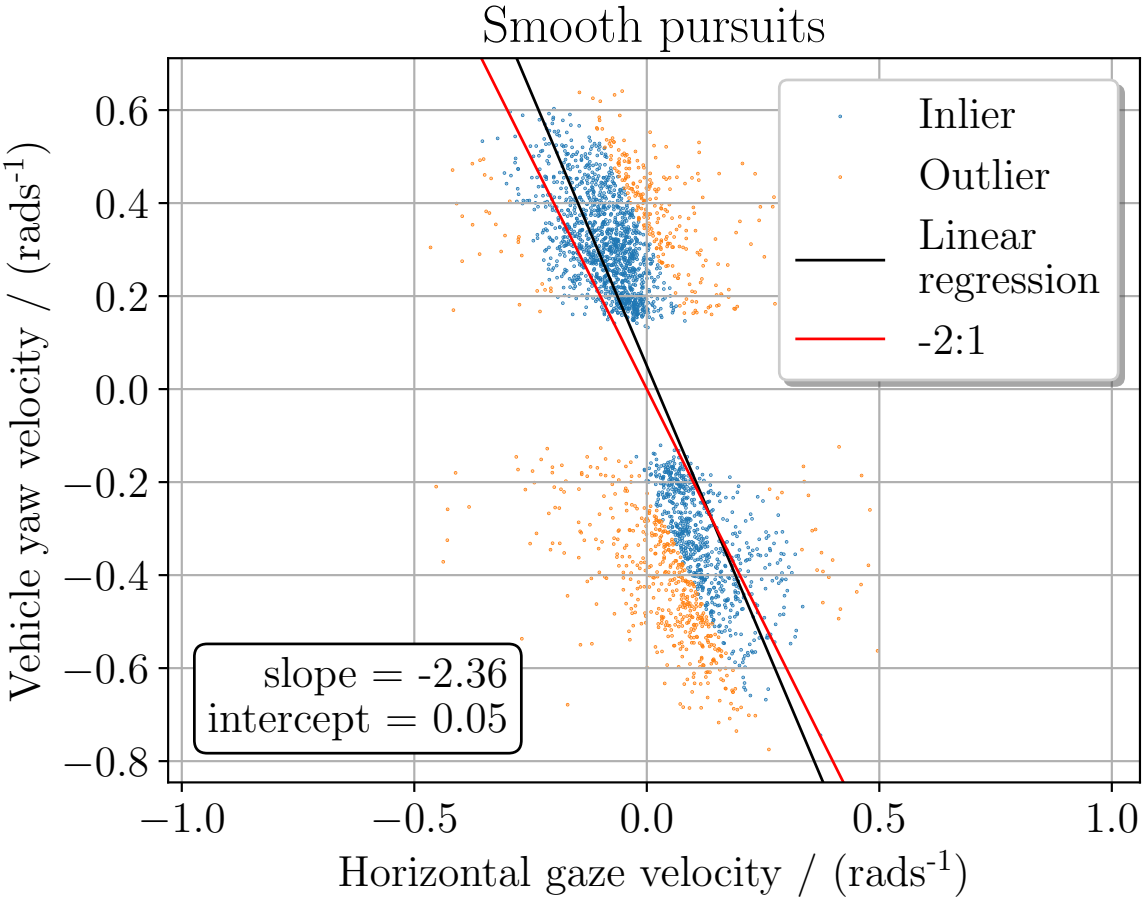


Figure 5.3: Optokinetic samples during curvilinear motion. The figure shows the optokinetic smooth pursuit data when humans maintain their motion on a curvilinear path. When humans produce exact smooth pursuit, i.e., perfectly tracking their gaze on a visual target on a curvilinear path, the vehicular yaw angular velocity is twice the horizontal gaze velocity. The experimental data show that the smooth pursuits here tend to be slightly lower than the predicted value of -2.0 by theory (negative due to the definition of the screen coordinate system).

5.1.2 Gaze placement on ground

Kim et al. [64] and Wann et al. [138] suggested humans skillfully deploy the strategy *"looking where you are going"* when maintaining proper steering in curve bends, inferred in Fig. 5.1. Most gaze points are distributed at the apex of the curve, for both left and right bends. This may imply the importance of correct gaze placement if the retinal optic flow-cue is to be used for correction. Furthermore, initial analysis of the data [91] suggests that this behavior is learned and deployed by skilled or experienced drivers, as argued by Wann et al. [138]. In contrast, inexperienced drivers tend to fix their gaze toward the horizon or straight ahead along the vehicle heading, as shown in Fig. 5.1 for some research participants.

When the gazing point gets too close, it is shifted towards a point farther away in the scene. From Paper A, it was observed that the gaze often shifted along the apex line in the curve bend in curvilinear motion. Using the experimental data [91], the placement characteristic is shown in Fig. 5.4. Here, the gaze time headway is defined as

$$T_{\mathrm{h}} = s|\vec{v}|^{-1}, \tag{5.1}$$

where s is the distance from the head to the gaze point fixation on the ground, and $\vec{v}$ is the vehicular velocity. Note that there is no strict and agreed-upon definition of time headway in three-dimensional space, as it is usually considered one-dimensional.

The launch of the saccade occurs close to $T_{\mathrm{h}} = 0.4\,\mathrm{s}$ (2.5th percentile), similar to findings in the works of Tuhkanen et al. [129] where it was estimated to be $T_{\mathrm{h}} \approx 1.0\,\mathrm{s}$ (lower estimate). The discrepancy may be explained by how the time headway is computed and the head position of the research participant above the ground. Another explanation may be that the research participants are conditioned to fixate their gaze on the visually placed waypoints, unlike our work, where there are no discrete waypoints to follow but only continuous lane markings.

To summarize, studying human research participants revealed a response time of $0.14\,\mathrm{s}$ *defined from retinal optic flow angle onset to ballistic steering correction onset using cross-correlation time delay analysis (and* $0.44\,\mathrm{s}$ *for heading-based). During the curve bend, it is shown that the vehicular yaw velocity is* -2.36 *times faster than the horizontal gaze velocity (c.f.* -2.0 *by theory). Despite the small discrepancy likely caused by retinal slippage, the assumption of exact smooth pursuit arguably still holds when reconstructing retinal optic flow. In addition, experimental data confirms that skilled drivers tend to intelligently fix their gaze in specific places on the ground, simply "looking where you are going". This gazing strategy arguably enables and suggests the use of retinal optic flow angle for vehicular locomotor control.*

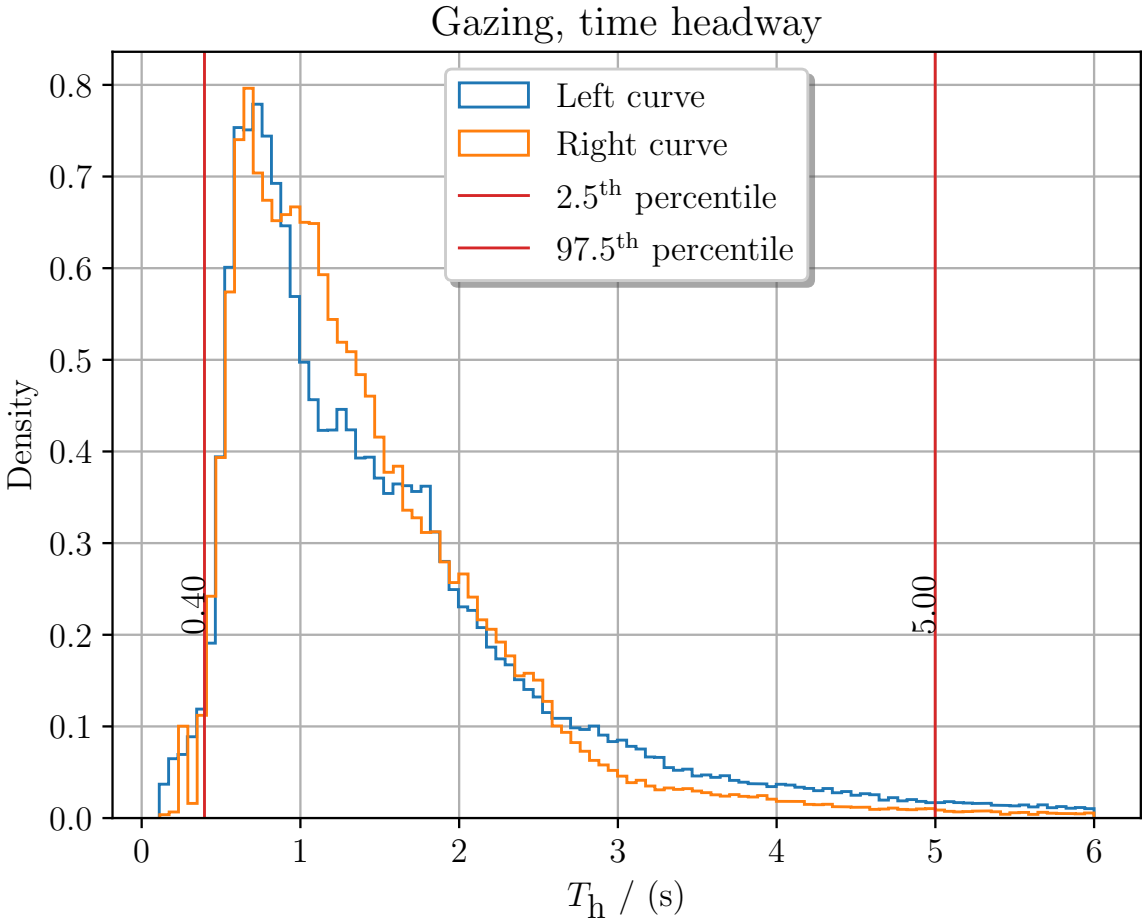


Figure 5.4: Human gazing time headway distributions. The figure shows the gazing time headway for left and right curve bends during curvilinear locomotion. It provides insight into how far and for how long humans tend to fixate on the curve bends. When the gazing point gets too close around $T_h = 0.4\,s$, the gaze is shifted a point further away.

5.2 Maintaining steering in curvilinear paths

The early steering models presented in Section 2.2.2 fail to capture the satisficing behavior observed in humans; moreover, in particular, the *reaching* movement characteristics and the *intermittency* property. These are some of the fundamental properties of realistic human movement and control. It is previously known that human control behaviors could be more accurately described and interpreted using intermittent control, in contrast to the more widely used and classical PID control. However, the latter has been selected and preferred over the former due to being easier to implement and well-established in engineering.

In Paper A and mentioned above, a connection was established (**RQ2**) using a time delay analysis of the steering adjustment travel angle $|\Delta\delta_b|$ to the visual cues of the retinal optic flow angle $\bar{\theta}_{\text{rof}}$ and the locomotor heading θ_{h}. This connection would have been impossible to establish without using the theories of intermittent control and reaching [11]. The paper explicitly investigates how these stimuli may immediately trigger (online control) and demand a response implied by the sustained sensorimotor control model suggested by Markkula et al. [81], see the illustrated example timeline in Fig. 5.5.

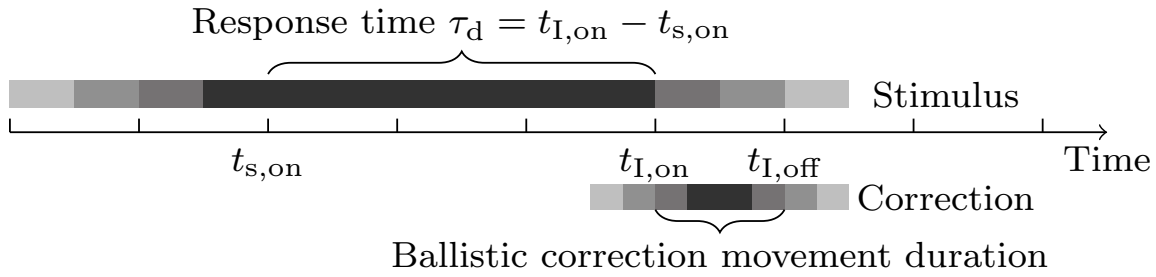


Figure 5.5: Timeline of human satisficing intermittent control. The figure illustrates a simplified human satisficing behavior with intermittent corrections using *evidence accumulation*. Once this perceived need for action becomes large enough at $t_{\text{s,on}}$, it triggers a corrective response at $t_{\text{I,on}}$, alleviating the need for action given that the expressed response is correct and proper. The difference between the two forms the human response time τ_{d}.

An investigation was carried out on how the degradation of visual information availability affects the steering control behavior. No significant differences were found in the characteristics of the correction travel angle $|\Delta\delta_{\text{b}}|$. The characteristic generally seems to follow the log-normal distribution $\log\mathcal{N}(\mu = 0.08, \sigma^2 = 0.75)$ as shown in Fig. 5.6 regardless. From the study, the visual information degradation mainly impacted the *timeliness* of initiating the correction, adding a latency of approximately 0.17 s for both retinal optic flow angle and locomotor heading.

It remains unclear what exact mechanisms are at play that contribute to the

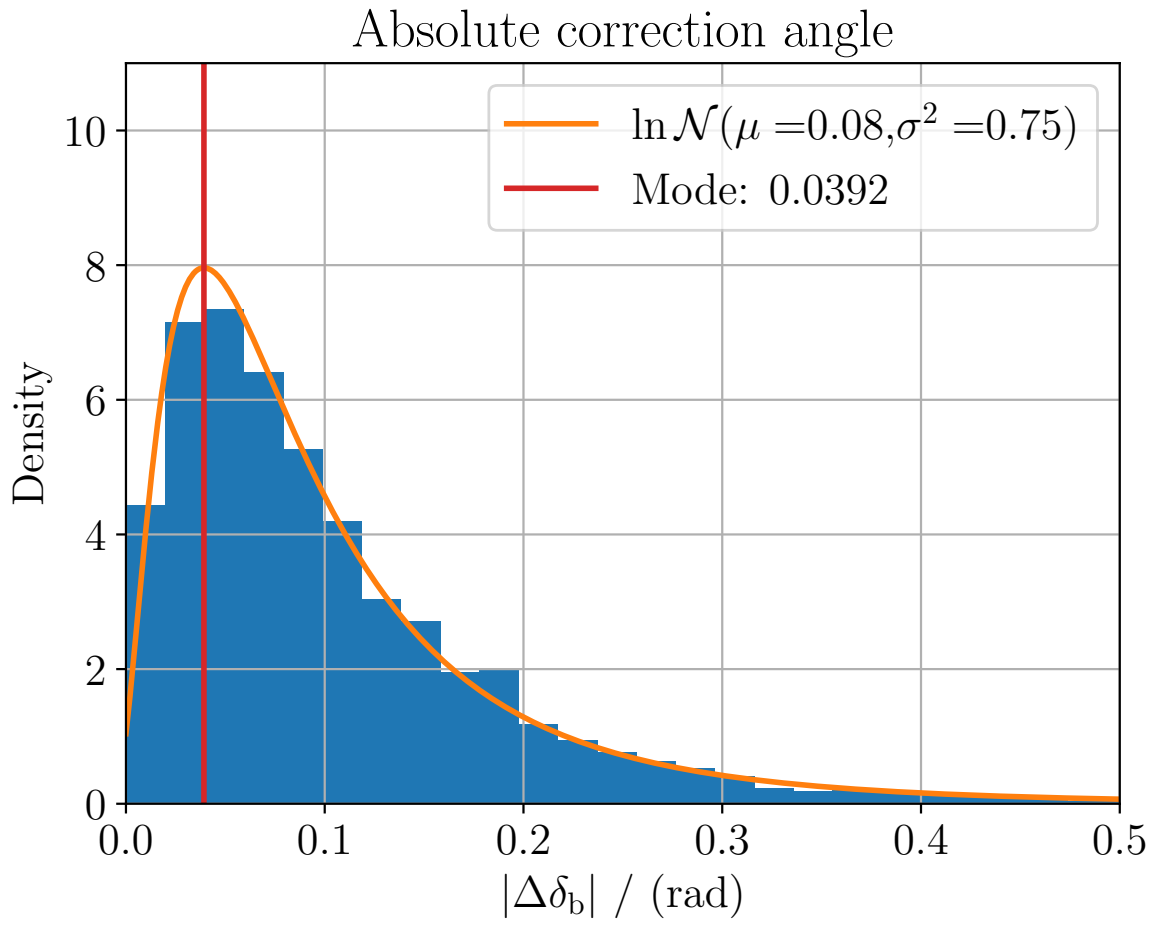


Figure 5.6: Absolute steering ballistic correction angle distributions. The figure shows the distribution of absolute steering ballistic correction travel angles. The distributions generally follow a theoretical log-normal distribution with parameters $\log \mathcal{N}(\mu = 0.08, \sigma^2 = 0.75)$.

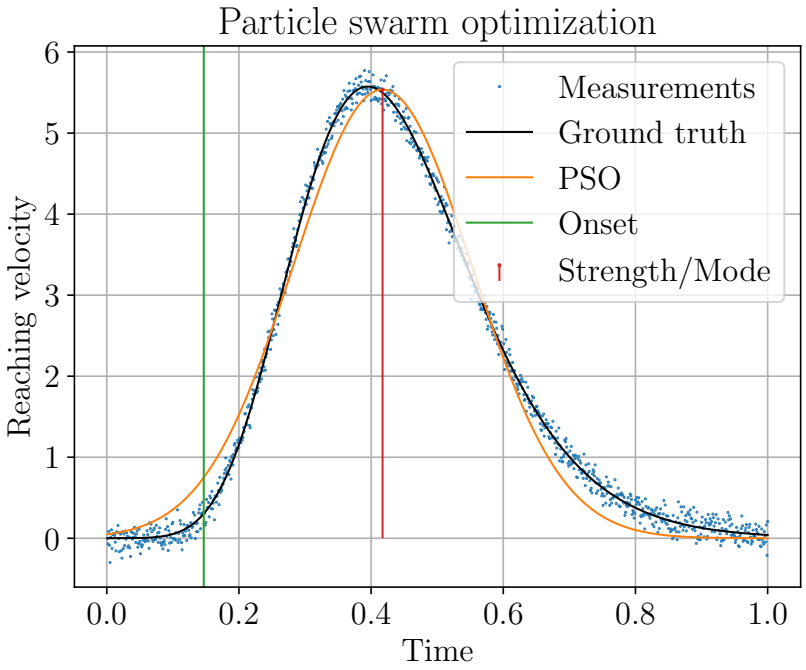
added latency to the correction. One possible explanation is that the degradation of visual information availability is compensated for by greater reliance on internalized models to produce prediction and anticipation. It could partially explain why the research participants anecdotally reported an increased difficulty in maintaining their steering when the steering wheel was invisible in virtual reality in Paper A.

While Paper A suggests an online corrective behavior in the curve bends, it does not provide convincing evidence for the same mechanism when humans enter and exit the curve bend using retinal optic flow or locomotor heading. The lack of convincing evidence implies that anticipatory and predictive behavior may play a role in human steering control in these contexts. Lehtonen et al. showed that gaze does not always directly result in steering correction [75]. Furthermore, they investigated the role of guiding fixations (located in the 1 s to 2 s in gazing time headway, see Fig. 5.4) and look-ahead fixations in the context of human driving in curve bends. They suggested that look-ahead fixations may play a role in planning egomotion for future correction [74, 37]. This internalized model could thus be in play when humans negotiate curve bends.

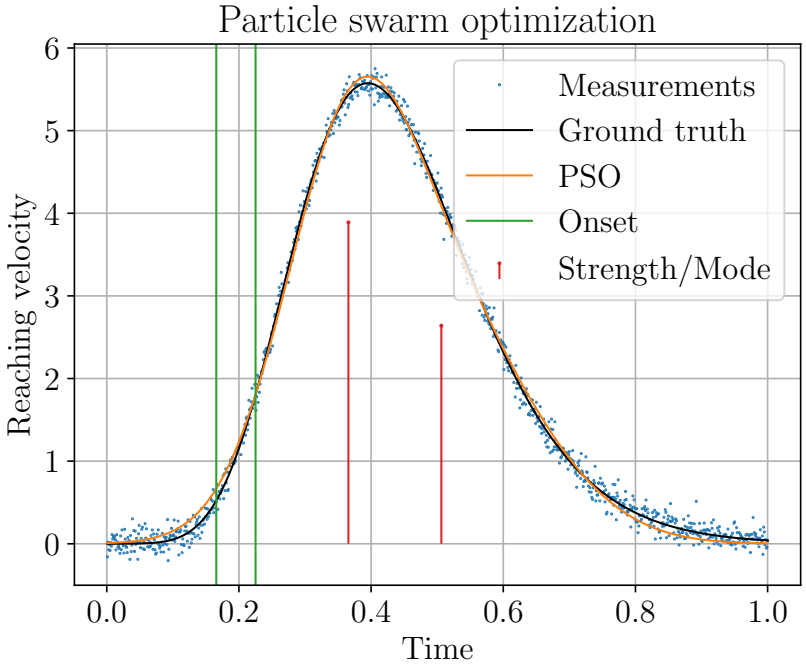
It should thus be concluded that a clear distinction of human sensorimotor models can not easily be made as either solely strong online or using internal models, using the definition of Zhao et al. [150]. The evidence points both ways, arguing that it could be modeled as online [102, 148] and internal-model [103, 55, 57, 129]. Hence, it may not be that useful to confine our understanding of the human locomotor control as such, but rather a complex hybrid sensorimotor system. The latter has been suggested and detailed by Markkula et al. [81], which simultaneously functions as online, producing highly *reactive* and *rapid* responses, with internal models creating the *anticipation* and *prediction* capabilities.

5.2.1 Heavy tail property in reaching velocity profiles

Early research has observed a heavy-tailed property in experimental data, creating the asymmetry in the human reaching velocity profile. In the works of Plamondon [104], it is argued that the heavy-tail property originates from the unsynchronicity of the many muscular contractions in the human neuromuscular system. The pre-study results in simulation from Paper A may provide a convincing alternative explanation to the heavy-tail property. It could be the emergence of multiple out-of-sync superposed ballistic corrections, illustrated in Fig. 5.7. The two different muscle synergies are coordinated to achieve the intended movement, but are slightly applied out of sync (see the onset timings in Fig. 5.7b). This explanation fits our current understanding of human movement coordination and the framework of muscle synergies better.



(a) One Gaussian function



(b) Two Gaussian functions

Figure 5.7: Approximating reaching movement velocity profile with a heavy tail. The figure demonstrates how multiple Gaussian membership functions may capture and approximate the heavy tail. (a) demonstrates how well the heavy tail is captured with a single Gaussian membership function and further in (b) with two Gaussian membership functions of different parameters. The synthetic data containing the heavy tail is padded with a normal distribution measurement noise $n(t) \sim \mathcal{N}(0, 0.1)$.

To summarize, applying our current understanding of intermittent control, motivated by the satisficing behavior observed in humans, to the experimental data argues that human sensorimotor control is a hybrid of reactive and predictive. Specifically, the locomotor control uses both online stimuli (reactive) and internal models (predictive) to trigger ballistic corrections in forms of reaching movements through evidence accumulation. This approach is aligned with the most recent research in limb movement coordination and muscle synergies. Further, our work finds an alternative explanation for the heavy-tail property seen in the experimental data of rapid limb movements. That is, multiple unsynchronized overlapping ballistic corrections with symmetric bell-shaped velocity profiles could also produce the heavy-tail property.

5.3 Bionic navigation in robotics

The combined findings of Paper C and Paper D suggest favorable outcomes to facilitate and enable modern software development in robotic vehicles when adopting the containerization and microservice design paradigm (**RQ4**). This has allowed for state-of-the-art research, development, and technology to be rolled out on the road as continuous testing and integration, compared to experimental methods made in simulations with limited transferable results for real applications. In Paper C, it was concluded that modern software development using containerized development in road vehicles is possible despite having various types and platforms of hardware architecture. This was demonstrated in robotics platforms ranging from miniature vehicles, racing cars, heavy-duty vehicles, passenger vehicles, and marine surface vehicles. The various computational resources were abstracted and generalized, with an attempt to increase the separation between software and hardware.

During the analysis in Paper A, it became apparent that the accuracy of reconstructing retinal optic flow would heavily depend on the performance of the optic flow estimator. Depending on the application, the optic flow estimator has to be carefully selected and properly evaluated. The complex environment of independently moving objects, challenging high dynamic range in the scene, and rapid head motion significantly impact the accuracy performance. As discussed in Paper C and intuitively, there is a clear relation between computational resources and performance accuracy, a trade-off of computational resources and accuracy; spending more computational resources may yield better accuracy results. However, this poses a challenge when perception is part of a time-sensitive critical system such as motion control. One possible solution to mitigate this is to use dynamic vision sensors to directly estimate the optic flow fields and eliminate the challenges of motion blur or scenes with high dynamic range lighting.

In Paper C, it was demonstrated that near real-time computing with optic flow estimators can be achieved using a software scheduler emphasizing run-time

consistency and fast computation, with some tolerance to a few isolated run-time computation outliers. Using the mental chronometry results from Paper A and neuro-cognitive science, it is implied that a computation deadline could be set 0.15 s as a whole from information sensation onset to control response onset, as inspired by human locomotor control. More specifically, the individual deadline could be divided as approximately 0.1 s for (visual) perception [81, 27, 107], and approximately 0.04 s for initiating the control adjustment [67, 89, 81].

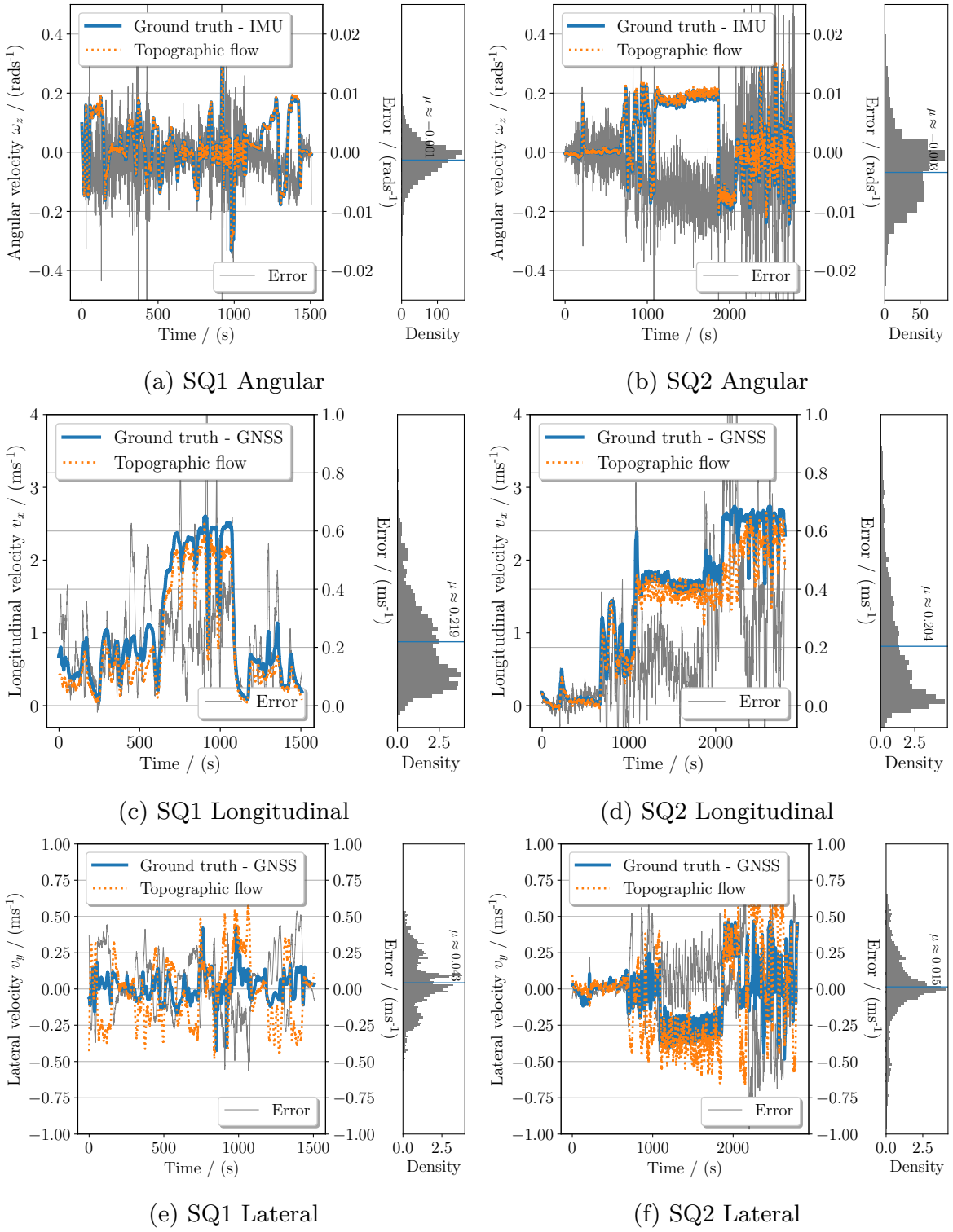
It was also demonstrated that the optic flow estimators could be implemented using a containerized approach, which showcases the modularity and flexibility of such a deployment strategy. Such development is preferable for continuous experimentation during a development phase and, more importantly, retrospective analysis for a fair comparison and traceability in deployed systems. So far, no dense optic flow estimator has been formally proven and demonstrated to exhibit real-time capabilities with sufficient accuracy.

Self-localization and odometry results and limitations

Inspired by the navigation of biological creatures, optical flow has been investigated and demonstrated for the applications of self-localization and odometry for a marine vehicle in coastal and littoral settings (**RQ3**). It is shown that local navigation based on optic flow is *possible* with emphasis on dimensionless estimates, as it does not suffer from the scaling problems. This is due to using a mono-camera (a perspective projection), which ambiguously truncates depth perception into image representation.

The topographic flow method, described in Paper E and introduced in Section 4.4, yielded the results shown in Fig. 5.8. The method deviates from the traditional epipolar geometry computation, showing potential applied to alternative information representations like abstracted topographic maps. However, in this particular instance, the work in Paper E suffers significantly from slow and asynchronous sampling, i.e. rolling shutter effect, which affects how the optic flow is computed since it estimates the pixel translation between two images. This problem originates from how radar technology functions, which can be avoided entirely if the topographic map is derived from an alternative sensor method based on synchronous sampling (global shutter), for example, a steady-state radar with 360° field of view.

The performance and feasibility of a traditional optic flow-based odometry method, direct sparse odometry (DSO) developed by Engel et al. [36], was investigated in littoral settings, see the odometry results in Fig. 5.9. The DSO method suffers challenges similar to many other mono-camera-based odometry approaches, i.e. scaling and relative reference alignment, which can only be mitigated through additional sensors, additional computing methods, or exact and elaborate calibrations. It is found that directly usable results are the dimensionless estimates, such



(a) SQ1 Angular

(b) SQ2 Angular

(c) SQ1 Longitudinal

(d) SQ2 Longitudinal

(e) SQ1 Lateral

(f) SQ2 Lateral

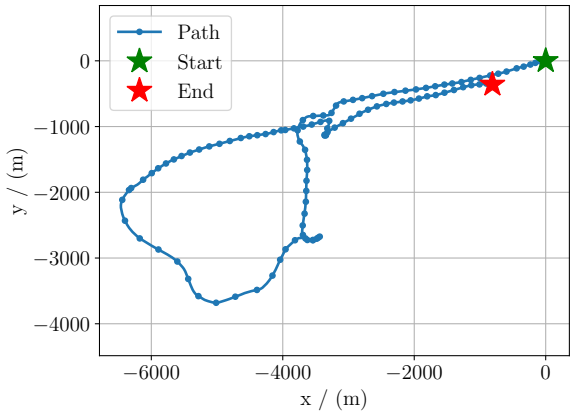
Figure 5.8: Topographic flow results. The angular velocity results of topographic flow are shown in (a) and (b), longitudinal velocity in (c) and (d), and lateral velocity in (e) and (f). Topographic flow is applied to two data sequences, SQ1 and SQ2, varying the sensor radar range configuration and vessel maneuvers. The error distributions are shown on the right of the time series panels.

as the angular velocity. It can thus serve as a complement or alternative to a gyroscope.

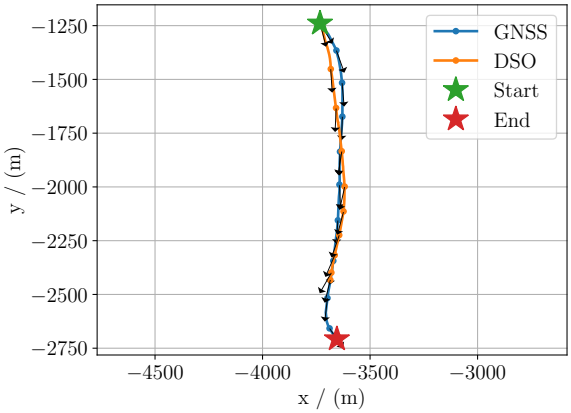
The applications discussed above mainly concern how to navigate locally in the environment, and thus do not provide a fixed solution for global positioning. However, making feasible attempts to estimate global positioning with additional information, such as a topographic map reference database, is possible. One such approach would be to implement, for example, a Monte Carlo localization i.e. particle filter localization, which uses a topographic map to establish localization of the agent. This is done by distributing filter particles representing a possible state in the map and updating the particles accordingly to the model kinematics. Impossible or unlikely states are discarded, resulting in the particle population converging to an optimal state (true state).

Thus, it has been shown that optic flow can be used for perception in navigation with caveats. Like flying insects such as *hymenopterans*, exact path integration is not necessary for successful navigation [33]. Similarly to the findings in this work, they argue that navigational cues such as optic flow are mainly relied upon for flight and that there are too many ambiguities to determine exact positioning. Instead, navigational capabilities are iteratively refined as hymenopterans familiarize themselves with the scene. Furthermore, to solve navigational ambiguities, honeybees help each other navigate by communicating a directional goal vector through *waggle dancing* where they perform a figure-eight dance. For robotics, perhaps researchers and developers should reconsider their navigation approach. In biology, the navigation task is only *sufficiently* solved, as opposed to exactly solved, such as optimal control.

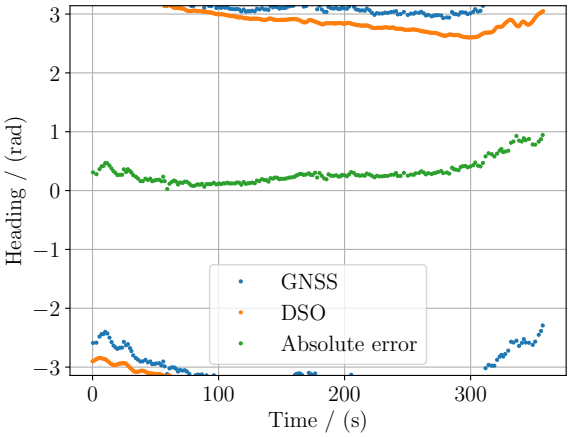
In summary, this thesis argues and demonstrates that real-world applications using bionic components for robotics can be carried out using containerization technology and a microservice design paradigm, particularly in research and development with rapid prototyping. A self-localization and an odometry based on optic flow are demonstrated using real-world and naturalistic data in marine settings with some success. However, the strict hard deadline requirements may need to be relaxed as complex computational bionic implementation may not consistently meet them. Inspired by human locomotor control behavior, it is observed that sufficient control could be an alternative to optimal control. It remains to be seen if the computational and locomotor performance of a bionic system can yield acceptable control behavior compared to its classical control counterpart for robotics.



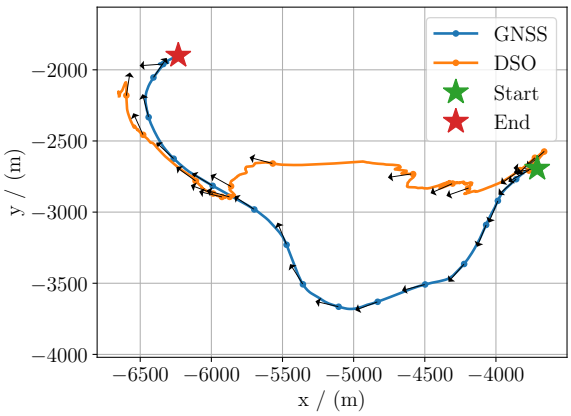
(a) Reference GNSS



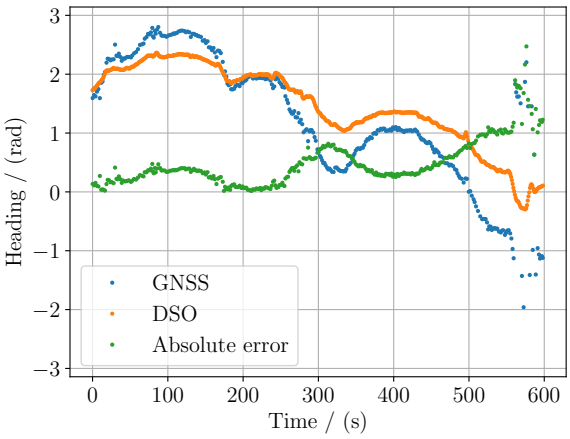
(b) Position



(c) Heading



(d) Position



(e) Heading

Figure 5.9: Direct sparse odometry results in littoral settings. The results from *direct sparse odometry* based on optic flow are shown. The complete GNSS positioning trace as the reference is shown in (a). Particular sections are highlighted, (a) showing vehicular position and (b) showing vehicular heading during the straight maneuver. An equivalent for the turning maneuver section is shown in (d) for position and (e) for heading.

Chapter 6

Concluding remarks and future work

The main conclusions of this thesis and the possible paths for future work are presented in this chapter.

6.1 Conclusions

The main conclusions from this thesis are as follows:

(i) Retinal optic flow is a major perception cue for human locomotor control. This thesis introduced and implemented a computational approach to quantify dense retinal optic flow field using experimental gazing and head-fixed visual video data (Paper A). The quantity *retinal optic flow angle* was derived and successfully utilized (**RQ1**) for analyzing human locomotor control in maintaining a curvilinear path, as inspired by the retinal optic flow path and flow lines as central concepts in *nulling flow curvature*-strategy proposed by researchers [138, 64].

(ii) Intermittent control in combination with theories of reaching and muscle synergies was introduced and discussed for human limb control and, by extension, human locomotor control (**RQ2**). A novel method for experimentally identifying individual ballistic corrections in simulated and real human data, utilizing particle swarm optimization, was successfully demonstrated (Paper A).

(iii) Further advancements in characterizing ballistic corrections as reaching movements were carried out and presented by studying the human-operated steering wheel. Using intermittent control for modeling human movement behavior has yielded a more detailed and accurate picture of human driver modeling (**RQ1**), compared to the contemporary models in narrow research fields.

(iv) Unification and harmonization of cross-disciplinary research contributions from human locomotor perceptions, control, and behavior were demonstrated using high-fidelity data [91, 93] and computing techniques. The study conducted on human research participants in Paper A, investigating human visual motion processing and intermittent control adjustment in high-speed curvilinear motion. The underlying work loosely connects the interaction between human perception, control, and locomotion,

(v) The human response time $\tau_{\mathrm{d}} \approx 0.14\,\mathrm{s}$ for retinal optic flow angle triggering a ballistic steering correction was found (**RQ2**) using cross-correlation analysis on data from human research participants; see Paper A. Using the characteristic of reaching, a well-defined definition of movement onset could be made, which excludes the movement travel time. Analogously, the response time for locomotor heading cue was found at $\tau_{\mathrm{d}} \approx 0.44\,\mathrm{s}$.

(vi) Optic flow-based odometry is a computational estimation technique of a general nature and may be applied beyond the traditional visual image domain. This idea is explored and demonstrated in Paper D on topographic map data derived from radar data. The radar sensor was mounted on a marine vehicle, and the odometry estimation was performed in coastal settings. Thanks to the detailed topographic data, well-defined topographic features were exploited for motion tracking using sparse optic flow, similar to visual feature identification techniques. Using IMU and GNSS as ground truth comparison unveils that the topographic flow performs considerably well in specific favorable conditions (**RQ3**) as static and feature-rich environments in littoral waters.

(vii) A traditional optic flow-based odometry, direct sparse odometry, was also investigated in a similar environment with comparable results (**RQ3**). However, it suffers from scaling issues in position estimation, which need to be dealt with in future work.

(viii) Microservice software design enables rapid prototyping and testing by abstracting the hardware layer, and harmonizing the data structure sent between microservices. The design allows researchers and developers to mainly focus on the core implementation of their prototype without worrying about the underlying robotic platform and computer architecture. Software modularity

is particularly favorable for robotics development (**RQ4**), as it mitigates the code complexity, increases run-time robustness, and enables real-time computing of complex tasks at the cost of increased deployment complexity. It simultaneously addresses other problems like prevalent irreproducible results and partially disclosed methods in the scientific community. The adoption enables continuous experimentation, continuous deployment, and continuous integration, which would greatly benefit the scientific community by increasing their impact, transparency, and comparability.

6.2 Future work

The thesis demonstrates the proof of concept of constructing and estimating the retinal optic flow field using human gazing data and head-fixed video feed. However, there is still much to be desired for optic flow regarding data volumes, representation, and sensors. Early work on estimating optical flow using event-driven sensors demonstrates promising accuracy, reliability, and faster data acquisition and processing. These aspects make the event-based camera far superior when estimating optic flow. Due to the limited data to process and fundamentally different data processing pipelines, it could enable high-quality retinal optic flow field implementation in a live robotic system and real-time applications.

Originally for Paper A, electromyography was planned for studying the muscle activation of the human research participants at four measurement points from the shoulder to the elbow. Unfortunately, it was quickly realized that more measurement points with better precision are needed. This is especially true when studying how the muscle synergies are expressed when operating a steering wheel, resulting in a steering correction. The movements expressing the steering corrections further involve muscle groups in the back (e.g. latissimus dorsi) and front chest (e.g. pectoralis minor) that must be considered. Furthermore, how these muscle groups form different configurations of muscle synergies creates the antagonistic dynamics of "pulling" and "pushing", rotating the steering wheel.

The interaction between local navigation, motion planning, and motion anticipation should be investigated to further the understanding of how humans solve the navigation task as a whole. This is indicated when humans negotiate into and out of curve bends, as the proposed heuristic strategy "look where you are going" fails to explain the mechanism of human locomotor control. From the initial glance, the mechanism likely stems from anticipation and internal models to produce steering corrections, which could explain the look-ahead fixation behavior reported by Lehtonen et al. [37].

Intermittent control and ballistic corrections are introduced and described in this work for studying humans, but not fully implemented in live robotics for actuating movements. The implementation would significantly progress the development of

a bionic robotic system with features such as evidence accumulation to trigger the ballistic correction, via, for example, a computational sustained sensorimotor model [81].

More research on superposed ballistic control is needed to fully answer whether it is multiple out-of-sync overlapping reaching movements or a single reaching movement. If it is overlapping ballistic reaching movements, it would better fit our current understanding of how the human brain coordinates complex movements through muscle synergies, but more evidence is needed. For example, this could be studied more carefully through electromyography in combination with motion capture of multiple joints.

Much work remains to interpret the retinal optic flow field for navigation. In this work, short-term steering maintenance is mainly considered. However, it is known that visual flows, e.g. looming, may inform emergency decisions like initiating braking or avoidance maneuvers. Another topic, as suggested by the works of Matthis et al. [83, 82], is to study the interpretation of the local retinal optic flow field around the gaze point. Intricate information may be embedded in the local field, aiding humans in their bipedal or curvilinear locomotor control.

Another future topic will be investigating how one can artificially generate retinal optic flow for robotics. That work would imply that gazing behavior needs to be artificially created, mimicking that in humans. Therefore, more research is required to understand *why* and *where* we are looking when we are going.

Finally, retinal optic flow remains to be implemented in a live robotic navigation system combined with intermittent control. Validation and comparison of such a system could bring a valuable analysis of the existing automated driving system concerning technology acceptance, user comfort, and system interpretability. More work is needed to improve computational performance and predictability related to retinal optic flow estimation and implementing artificial gazing behavior to support this type of bionic subsystem, emphasizing sufficiently hard deadlines for high-stakes decisions.

Chapter 7

Summary of included papers

This thesis consists of five papers investigating visual perception and locomotor control in humans (Paper A), how software design patterns and techniques for bionic research and robotics (Paper C and Paper D), and bio-inspired methods may be applied within robotics for navigational purposes (Paper B and Paper E).

7.1 Paper A

The primary purpose of Paper A is to investigate and understand human locomotor behavior, specifically how humans intermittently maintain proper steering during curve bends with the help of visual cues such as retinal optic flow and heading. An observational study was performed on fourteen research participants, eight of whom qualified for further analysis, driving in virtual reality on a two-lane S-shaped drive track equipped with an eye-tracking system. The simulated environment was designed to be simplistic and texture-rich to study the perceptual cues impacting human locomotor performance. A perceptual quantity, retinal optic flow angle, was derived using the immediately readily available visual cues (as may be seen from the point of view of the research participants), such as lane detection and segmentation. In addition, a novel method of identifying and approximating ballistic corrections in experimental positional reaching data using particle swarm optimization was developed and presented. In total, 3204 individual ballistic intermittent corrections were identified in the collected data, of which 2084 were used for the analysis. From this, a time-delay analysis using cross-correlation was performed, given the stimulus onset to the identified ballistic correction onset. It was revealed that human response time is approximately 0.14 s for retinal optic flow-based cues and 0.44 s for heading-based cues. The work showcased and demonstrated the intermittency

property in human neuromuscular control of muscle synergies, through the principle of *satisficing* behavior.

7.2 Paper B

Paper B applied and evaluated an existing monocular optic flow-based visual odometry system on a vessel in coastal settings. The motivation was that such systems were comparably unexplored for marine environments, as visual odometry systems are typically developed and target indoor use or on land. Using a simplistic single-camera system on a marine vehicle showed that the odometry system performed considerably well compared to GNSS regarding the orientation estimation under suitable and favorable conditions, such as visibility and feature-rich visual landmarks. It was clear from this work that additional work and further considerations were required to make it fully performant, as the marine environment imposed greater challenges for this particular odometry system. One challenge to properly estimate absolute positioning is that the method suffers from scaling problems. Another challenge of the method is the decomposition of the odometry into the rotational and translational contributions. However, the visual odometry system can complement existing and traditional navigational systems such as GNSS or IMU using sensor fusion approaches to increase localization reliability, robustness, and accuracy.

7.3 Paper C

Paper C addresses an evident challenge of reproducing research findings in scientific papers, especially in claimed run-time in performance benchmarks, without normalizing with respect to the used computational resources or retrospective performance analysis. Adopting containerization as demonstrated and proposed in Paper D showed that a fair and reproducible comparison can be made using optic flow estimators as the example. This may further aid the process of reproducing the research results and increase the transparency and credibility of the scientific community beyond paper results. The evaluations in the paper were based on computation using a real-time capable system. Using our methodology, the run-time results are normalized to a single system, which may provide more valuable comparison insight. Moreover, the performance uplift concerning error metrics and run-time metrics could be observed as contributed by technological advancements in the underlying dependency libraries, and computational resources are continuously improving (see Section 4.3.2). This allows researchers and developers to discuss the trade-offs of computational performance and accuracy and make a fair comparison across contributions and time.

7.4 Paper D

Paper D presents and details the accumulated experiences, best practices, and pitfalls of utilizing containerized software with the microservice design paradigm for self-driving road vehicles. This methodology of creating a microservice and packaging it into a self-contained software bundle, with all necessary dependency libraries, allows for clear binary traceability, software modularity, and easy deployment. Another motivating purpose was quickly introducing new researchers to the hardware and software, allowing them to focus on their core goals and problems. This way, decoupling computational resources from the algorithms made a dynamic restructuring of data processing pipelines possible. This allows for more flexible and interchangeable software blocks, resulting in convenient ways to introduce continuous integration, continuous deployment, and continuous experimentation. The work presented *OpenDLV* as an implementation of these design choices, including a standard message set, an *a priori* established communication protocol among the application microservices.

7.5 Paper E

Paper E explored and demonstrated a non-conventional application of optic flow to estimate odometry, which is presented as *topographic flow*. It is applied to topographic data instead of the conventional visual image data (compare this to visual odometry). By analyzing the motion of the pixels in the topographic map, it is possible to estimate the rotational and translational velocities of the agent caused by relative motion. The performance results when compared to GNSS as a ground truth showed mean errors of less than $0.003\,\mathrm{rad\,s^{-1}}$ for angular velocity, $0.04\,\mathrm{m\,s^{-1}}$ for lateral velocity, and $0.22\,\mathrm{m\,s^{-1}}$ for longitudinal velocity. This method was presented as a stand-alone alternative to the *de facto* standard GNSS used in navigation. The topographic data was solely derived from, but not limited to, experimental data from radar mounted on a marine vehicle. Due to the nature of optic flow, the challenge in odometry estimation lies in separating and identifying the translational and rotational components in the emergent optic flow analysis, similar to the challenges present in Paper B.

Bibliography

[1] M. A. Al Mamun et al. "Evolution of technical debt: An exploratory study". *2019 Joint Conference of the International Workshop on Software Measurement and the International Conference on Software Process and Product Measurement, IWSM-Mensura 2019, Haarlem, The Netherlands, October 7-9, 2019*. Vol. 2476. Ceur-ws. 2019, pp. 87–102.

[2] F. O. and. The Zeitraffer phenomenon, akinetopsia, and the visual perception of speed of motion: A case report. *Neurocase* **20**.3 (2014). Pmid: 23557277, 269–272. DOI: 10.1080/13554794.2013.770877. eprint: https://doi.org/10.1080/13554794.2013.770877. URL: https://doi.org/10.1080/13554794.2013.770877.

[3] R. Andreev, O. Scherzer, and W. Zulehner. Simultaneous optical flow and source estimation: Space–time discretization and preconditioning. *Applied Numerical Mathematics* **96** (2015), 72–81. ISSN: 0168-9274. DOI: 10.1016/j.apnum.2015.04.007. URL: https://www.sciencedirect.com/science/article/pii/S0168927415000781.

[4] C. Atkeson and J. Hollerbach. Kinematic features of unrestrained vertical arm movements. *Journal of Neuroscience* **5**.9 (1985), 2318–2330. ISSN: 0270-6474. DOI: 10.1523/jneurosci.05-09-02318.1985. eprint: https://www.jneurosci.org/content/5/9/2318.full.pdf. URL: https://www.jneurosci.org/content/5/9/2318.

[5] S. Baker et al. A database and evaluation methodology for optical flow. *International journal of computer vision* **92**.1 (2011), 1–31. DOI: 10.1007/s11263-010-0390-2.

[6] P. Bardow, A. J. Davison, and S. Leutenegger. "Simultaneous optical flow and intensity estimation from an event camera". *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016, pp. 884–892. URL: https://openaccess.thecvf.com/content%5C%5Fcvpr%5C%5F201

6/papers/Bardow%5C%5FSimultaneous%5C%5FOptical%5C%5FFlow%5C%5FCVPR%5C%5F2016%5C%5Fpaper.pdf.

[7] J. L. Barron, D. J. Fleet, and S. S. Beauchemin. Performance of optical flow techniques. *International journal of computer vision* **12**.1 (1994), 43–77. DOI: 10.1007/bf01420984.

[8] W. Beggs, R. Sakstein, and C. Howarth. The Generality of a Theory of the Intermittent Control of Accurate Movements. *Ergonomics* **17**.6 (1974), 757–768. DOI: 10.1080/00140137408931422. eprint: https://doi.org/10.1080/00140137408931422. URL: https://doi.org/10.1080/00140137408931422.

[9] O. Benderius. "Modelling Driver Steering and Neuromuscular Behaviour". English. PhD thesis. Chalmers University of Technology, 2014, p. 65. ISBN: 9781392736432. URL: https://publications.lib.chalmers.se/records/fulltext/205699/205699.pdf.

[10] O. Benderius, C. Berger, and K. Blanch. *Are we ready for beyond-application high-volume data? The Reeds robot perception benchmark dataset*. 2021. DOI: 10.48550/arXiv.2109.08250. arXiv: 2109.08250 [cs.CV]. URL: https://arxiv.org/abs/2109.08250.

[11] O. Benderius and G. Markkula. Evidence for a fundamental property of steering. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* **58**.1 (2014), 884–888. DOI: 10.1177/1541931214581186. eprint: https://doi.org/10.1177/1541931214581186. URL: https://doi.org/10.1177/1541931214581186.

[13] J. D. A. Bermudez. "Real-time Visual Flow Algorithms for Robotic Applications". PhD thesis. The Australian National University (Australia), 2017.

[14] "Bell-shaped Speed Profile". *Encyclopedia of Neuroscience*. Ed. by M. D. Binder, N. Hirokawa, and U. Windhorst. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 375–375. ISBN: 978-3-540-29678-2. DOI: 10.1007/978-3-540-29678-2_599. URL: https://doi.org/10.1007/978-3-540-29678-2%5C%5F599.

[15] "Point-to-Point Movements". *Encyclopedia of Neuroscience*. Ed. by M. D. Binder, N. Hirokawa, and U. Windhorst. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 3173–3173. ISBN: 978-3-540-29678-2. DOI: 10.1007/978-3-540-29678-2_4627. URL: https://doi.org/10.1007/978-3-540-29678-2%5C%5F4627.

[16] E. Bizzi and V. C. Cheung. The neural origin of muscle synergies. *Frontiers in Computational Neuroscience* **7** (2013). ISSN: 1662-5188. DOI: 10.3389/fncom.2013.00051. URL: https://www.frontiersin.org/journals/computational-neuroscience/articles/10.3389/fncom.2013.00051.

[17] C. Blakemore and G. F. Cooper. Development of the brain depends on the visual environment. *Nature* **228**.5270 (1970), 477–478. DOI: 10.1038/228477a0.

[18] K. Blanch. "Beyond Application Datasets and Automated Fair Benchmarking". English. PhD thesis. Chalmers University of Technology, 2023, p. 65. ISBN: 9781392736432. URL: https://research.chalmers.se/publication/537236/file/537236%5C%5FFulltext.pdf.

[20] J.-Y. Bouguet et al. Pyramidal implementation of the affine lucas kanade feature tracker description of the algorithm. *Intel corporation* **5**.1-10 (2001), 4. eprint: http://robots.stanford.edu/cs223b04/algo_affine_tracking.pdf.

[21] D. J. Butler et al. "A naturalistic open source movie for optical flow evaluation". *European Conf. on Computer Vision (ECCV)*. Ed. by A. F. et al. (Eds.) Part IV, LNCS 7577. Springer-Verlag, Oct. 2012, pp. 611–625. DOI: 10.1007/978-3-642-33783-3_44.

[22] D. Calow and M. Lappe. Efficient encoding of natural optic flow. *Network: Computation in Neural Systems* **19**.3 (2008), 183–212. DOI: 10.1080/09548980802368764.

[23] M. Cannici and D. Scaramuzza. "Mitigating Motion Blur in Neural Radiance Fields with Events and Frames". *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2024, pp. 9286–9296. eprint: https://openaccess.thecvf.com/content/CVPR2024/papers/Cannici_Mitigating_Motion_Blur_in_Neural_Radiance_Fields_with_Events_and_CVPR_2024_paper.pdf.

[24] J. C. K. Cheng and L. Li. Effects of reference objects and extra-retinal information about pursuit eye movements on curvilinear path perception from retinal flow. *Journal of Vision* **12**.3 (Mar. 2012), 12–12. ISSN: 1534-7362. DOI: 10.1167/12.3.12. eprint: https://arvojournals.org/arvo/content_public/journal/jov/932803/jov-12-3-12.pdf. URL: https://doi.org/10.1167/12.3.12.

[25] E. C. Cherry. Some Experiments on the Recognition of Speech, with One and with Two Ears. *The Journal of the Acoustical Society of America* **25**.5 (Sept. 1953), 975–979. ISSN: 0001-4966. DOI: 10.1121/1.1907229. eprint: https://pubs.aip.org/asa/jasa/article-pdf/25/5/975/18731769/975_1_online.pdf. URL: https://doi.org/10.1121/1.1907229.

[26] V. C. Cheung et al. Plasticity of muscle synergies through fractionation and merging during development and training of human runners. *Nature communications* **11**.1 (2020), 4356. DOI: 10.1038/s41467-020-18210-4.

[27] E. P. Cook and J. H. Maunsell. Dynamics of neuronal responses in macaque MT and VIP during motion detection. *Nature neuroscience* **5**.10 (2002), 985–994. DOI: 10.1038/nn924.

[28] S. A. Cooper et al. Akinetopsia: acute presentation and evidence for persisting defects in motion vision. *Journal of Neurology, Neurosurgery & Psychiatry* **83**.2 (2012), 229–230. DOI: 10.1136/jnnp.2010.223727.

[29] K. J. Craik. Theory of the human operator in control systems 1: I. The operator as an engineering system. *British Journal of Psychology. General Section* **38**.2 (1947), 56–61. DOI: 10.1111/j.2044-8295.1947.tb01141.x. URL: https://doi.org/10.1111/j.2044-8295.1947.tb01141.x.

[30] A. d'Avella. "Reaching Movements". *Encyclopedia of Neuroscience*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 3363–3367. ISBN: 978-3-540-29678-2. DOI: 10.1007/978-3-540-29678-2_4936. URL: https://doi.org/10.1007/978-3-540-29678-2%5C%5F4936.

[31] A. D'avella and F. Lacquaniti. Control of reaching movements by muscle synergy combinations. *Frontiers in Computational Neuroscience* **7** (2013). ISSN: 1662-5188. DOI: 10.3389/fncom.2013.00042. URL: https://www.frontiersin.org/journals/computational-neuroscience/articles/10.3389/fncom.2013.00042.

[32] N. Dominici et al. Locomotor Primitives in Newborn Babies and Their Development. *Science* **334**.6058 (2011), 997–999. DOI: 10.1126/science.1210617. eprint: https://www.science.org/doi/pdf/10.1126/science.1210617. URL: https://www.science.org/doi/abs/10.1126/science.1210617.

[33] M. Egelhaaf and J. P. Lindemann. Path integration and optic flow in flying insects: a review of current evidence. *Journal of Comparative Physiology A* (2025), 1–27. DOI: 10.1007/s00359-025-01734-9. URL: https://doi.org/10.1007/s00359-025-01734-9.

[34] G. A. Elliott and J. H. Anderson. Globally scheduled real-time multiprocessor systems with GPUs. *Real-Time Systems* **48**.1 (2012), 34–74. DOI: 10.1007/s11241-011-9140-y.

[35] G. A. Elliott and J. H. Anderson. "Robust Real-Time Multiprocessor Interrupt Handling Motivated by GPUs". *2012 24th Euromicro Conference on Real-Time Systems*. 2012, pp. 267–276. DOI: 10.1109/ECRTS.2012.20.

[36] J. Engel, V. Koltun, and D. Cremers. Direct Sparse Odometry. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **40**.3 (Mar. 2018), 611–625. ISSN: 1939-3539. DOI: 10.1109/tpami.2017.2658577.

[37] H. K. Esko Lehtonen Otto Lappi and H. Summala. Look-ahead fixations in curve driving. *Ergonomics* **56**.1 (2013). Pmid: 23140361, 34–44. DOI: 10.1080/00140139.2012.739205. eprint: https://doi.org/10.1080/00140139.2012.739205. URL: https://doi.org/10.1080/00140139.2012.739205.

[38] G. Farnebäck. "Two-Frame Motion Estimation Based on Polynomial Expansion". *Image Analysis*. Ed. by J. Bigun and T. Gustavsson. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003, pp. 363–370. ISBN: 978-3-540-45103-7. DOI: 10.1007/3-540-45103-x_50.

[39] T. E. Feinberg and J. Mallatt. The evolutionary and genetic origins of consciousness in the Cambrian Period over 500 million years ago. *Frontiers in psychology* **4** (2013), 667. DOI: 10.3389/fpsyg.2013.00667.

[40] M. A. Fischler and R. C. Bolles. "Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography". *Readings in Computer Vision*. Ed. by M. A. Fischler and O. Firschein. San Francisco (CA): Morgan Kaufmann, 1987, pp. 726–740. ISBN: 978-0-08-051581-6. DOI: 10.1016/B978-0-08-051581-6.50070-2. URL: https://www.sciencedirect.com/science/article/pii/B9780080515816500702.

[41] D. J. Fleet and A. D. Jepson. Computation of component image velocity from local phase information. *International journal of computer vision* **5**.1 (1990), 77–104. DOI: 10.1007/bf00056772.

[42] J. P. Gallivan et al. Decision-making in sensorimotor control. *Nature Reviews Neuroscience* **19**.9 (2018), 519–534. DOI: 10.1038/s41583-018-0045-9.

[43] P. Gawthrop et al. Intermittent control: a computational theory of human control. *Biological cybernetics* **104** (2011), 31–51. DOI: 10.1007/s00422-010-0416-4.

[44] D. Gehrig and D. Scaramuzza. Low-latency automotive vision with event cameras. *Nature* **629**.8014 (2024), 1034–1040. DOI: 10.1038/s41586-024-07409-w.

[45] D. Gehrig et al. "End-to-end learning of representations for asynchronous event-based data". *Proceedings of the IEEE International Conference on Computer Vision*. 2019, pp. 5633–5643. eprint: https://openaccess.thecvf.com/content_ICCV_2019/papers/Gehrig_End-to-End_Learning_of_Representations_for_Asynchronous_Event-Based_Data_ICCV_2019_paper.pdf.

[46] M. Gehrig, M. Muglikar, and D. Scaramuzza. Dense Continuous-Time Optical Flow From Event Cameras. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **46**.7 (2024), 4736–4746. DOI: 10.1109/tpami.2024.3361671.

[47] A. Geiger, P. Lenz, and R. Urtasun. "Are we ready for autonomous driving? The KITTI vision benchmark suite". *2012 IEEE Conference on Computer Vision and Pattern Recognition*. 2012, pp. 3354–3361. DOI: 10.1109/cvpr.2012.6248074.

[48] A. P. Georgopoulos, J. F. Kalaska, and J. T. Massey. Spatial trajectories and reaction times of aimed movements: effects of practice, uncertainty, and change in target location. *Journal of Neurophysiology* **46**.4 (1981). Pmid: 7288461, 725–743. DOI: 10.1152/jn.1981.46.4.725. eprint: https://doi.org/10.1152/jn.1981.46.4.725. URL: https://doi.org/10.1152/jn.1981.46.4.725.

[49] A. P. Georgopoulos. On reaching. *Annual review of neuroscience* (1986). DOI: 10.1146/annurev.ne.09.030186.001051.

[50] J. J. Gibson. The perception of the visual world. *The American Journal of Psychology* (1950).

[51] S. C. A. M. Gielen. "Motor Control Models". *Encyclopedia of Neuroscience*. Ed. by M. D. Binder, N. Hirokawa, and U. Windhorst. Berlin, Heidelberg: Springer Berlin Heidelberg, 2009, pp. 2428–2431. ISBN: 978-3-540-29678-2. DOI: 10.1007/978-3-540-29678-2_3584. URL: https://doi.org/10.1007/978-3-540-29678-2%5C%5F3584.

[52] N. Gosselin-Kessiby, J. Messier, and J. F. Kalaska. Evidence for Automatic On-Line Adjustments of Hand Orientation During Natural Reaching Movements to Stationary Targets. *Journal of Neurophysiology* **99**.4 (2008). Pmid: 18256170, 1653–1671. DOI: 10.1152/jn.00980.2007. eprint: https://doi.org/10.1152/jn.00980.2007. URL: https://doi.org/10.1152/jn.00980.2007.

[53] B. Grimme, J. Lipinski, and G. Schöner. Naturalistic arm movements during obstacle avoidance in 3D and the identification of movement primitives. *Experimental brain research* **222** (2012), 185–200. DOI: 10.1007/s00221-012-3205-6.

[54] M. Hanna, J. Fung, and A. Lamontagne. "Is extra-retinal information needed to control steering of locomotion in presence of a rotational optic flow?" *2008 Virtual Rehabilitation*. 2008, pp. 54–59. DOI: 10.1109/icvr.2008.4625122.

[55] M. M. Hayhoe et al. Predictive eye movements in natural vision. *Experimental brain research* **217** (2012), 125–136. DOI: 10.1007/s00221-011-2979-2.

[56] R. Held et al. The newly sighted fail to match seen with felt. *Nature neuroscience* **14**.5 (2011), 551–553. DOI: 10.1038/nn.2795.

[57] J. M. Henderson. Gaze control as prediction. *Trends in cognitive sciences* **21**.1 (2017), 15–23. DOI: 10.1016/j.tics.2016.11.003.

[58] J. M. Henderson. Human gaze control during real-world scene perception. *Trends in Cognitive Sciences* **7**.11 (2003), 498–504. ISSN: 1364-6613. DOI: https://doi.org/10.1016/j.tics.2003.09.006. URL: https://www.sciencedirect.com/science/article/pii/S1364661303002481.

[59] J. Heutink et al. The effect of target speed on perception of visual motion direction in a patient with akinetopsia. *Cortex* **119** (2019), 511–518. ISSN: 0010-9452. DOI: https://doi.org/10.1016/j.cortex.2018.12.002. URL: https://www.sciencedirect.com/science/article/pii/S001094521830409X.

[60] S. Idrees et al. *Perceptual saccadic suppression starts in the retina*. 2020. DOI: 10.1038/s41467-020-15890-w.

[61] E. Ilg et al. "FlowNet 2.0: Evolution of Optical Flow Estimation With Deep Networks". *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. July 2017. eprint: https://openaccess.t

hecvf.com/content_cvpr_2017/papers/Ilg_FlowNet_2.0_Evolution_CVPR_2017_paper.pdf.

[62] K. Johari et al. Eye movement analysis for real-world settings using segmented linear regression. *Computers in Biology and Medicine* **174** (2024), 108364. ISSN: 0010-4825. DOI: 10.1016/j.compbiomed.2024.108364. URL: https://www.sciencedirect.com/science/article/pii/S0010482524004487.

[63] J. Kennedy and R. Eberhart. "Particle swarm optimization". *Proceedings of ICNN'95 - International Conference on Neural Networks*. Vol. 4. 1995, 1942–1948 vol.4. DOI: 10.1109/icnn.1995.488968.

[64] N.-G. Kim and M. Turvey. Eye movements and a rule for perceiving direction of heading. *Ecological Psychology* **11**.3 (1999), 233–248. DOI: 10.1207/s15326969eco1103_3.

[65] I. Koch et al. Cognitive structure, flexibility, and plasticity in human multitasking–An integrative review of dual-task and task-switching research. *Psychological bulletin* **144**.6 (2018), 557. DOI: 10.1037/bul0000144.

[66] T. Kroeger et al. "Fast Optical Flow Using Dense Inverse Search". *Computer Vision – ECCV 2016*. Ed. by B. Leibe et al. Cham: Springer International Publishing, 2016, pp. 471–488. ISBN: 978-3-319-46493-0.

[67] Y. Lamarre, G. Spidalieri, and J. P. Lund. Patterns of muscular and motor cortical activity during a simple arm movement in the monkey. *Canadian Journal of Physiology and Pharmacology* **59**.7 (1981). Pmid: 7317854, 748–756. DOI: 10.1139/y81-111. eprint: https://doi.org/10.1139/y81-111. URL: https://doi.org/10.1139/y81-111.

[68] M. F. Land and D. N. Lee. Where we look when we steer. *Nature* **369**.6483 (1994), 742–744. DOI: 10.1038/369742a0.

[69] M. Lappe, F. Bremmer, and A. van den Berg. Perception of self-motion from visual flow. *Trends in Cognitive Sciences* **3**.9 (1999), 329–336. ISSN: 1364-6613. DOI: 10.1016/s1364-6613(99)01364-9. URL: https://www.sciencedirect.com/science/article/pii/S1364661399013649.

[70] O. Lappi et al. Beyond the tangent point: Gaze targets in naturalistic driving. *Journal of Vision* **13**.13 (Nov. 2013), 11–11. ISSN: 1534-7362. DOI: 10.1167/13.13.11. eprint: https://arvojournals.org/arvo/content_public/journal/jov/933542/i1534-7362-13-13-11.pdf. URL: https://doi.org/10.1167/13.13.11.

[71] O. Lappi et al. Humans use optokinetic eye movements to track Waypoints for Steering. *Scientific reports* **10**.1 (2020), 1–14. DOI: 10.1038/s41598-020-60531-3.

[72] O. W. Layton, E. Mingolla, and N. A. Browning. A motion pooling model of visually guided navigation explains human behavior in the presence of independently moving objects. *Journal of vision* **12**.1 (2012), 20–20. DOI: 10.1167/12.1.20.

[73] D. N. LEE and R. LISHMAN. Visual control of locomotion. *Scandinavian Journal of Psychology* **18**.1 (1977), 224–230. DOI: 10.1111/j.1467-9450.1977.tb00281.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9450.1977.tb00281.x. URL: https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1467-9450.1977.tb00281.x.

[74] E. Lehtonen et al. Effect of driving experience on anticipatory look-ahead fixations in real curve driving. *Accident Analysis and Prevention* **70** (2014), 195–208. ISSN: 0001-4575. DOI: 10.1016/j.aap.2014.04.002. URL: https://www.sciencedirect.com/science/article/pii/S0001457514001079.

[75] E. Lehtonen et al. Gaze doesn't always lead steering. *Accident Analysis and Prevention* **121** (2018), 268–278. ISSN: 0001-4575. DOI: 10.1016/j.aap.2018.09.026. URL: https://www.sciencedirect.com/science/article/pii/S0001457518306833.

[76] J. Lewis and M. Fowler. *Microservices*. 2014. URL: http://martinfowler.com/articles/microservices.html.

[77] L. Li and J. C. K. Cheng. Perceiving path from optic flow. *Journal of Vision* **11**.1 (Jan. 2011), 22–22. ISSN: 1534-7362. DOI: 10.1167/11.1.22. eprint: https://arvojournals.org/arvo/content_public/journal/jov/933479/jov-11-1-22.pdf. URL: https://doi.org/10.1167/11.1.22.

[78] L. Li and J. C. Cheng. Visual strategies for the control of steering toward a goal. *Displays* **34**.2 (2013), 97–104. ISSN: 0141-9382. DOI: 10.1016/j.displa.2012.10.005. URL: https://www.sciencedirect.com/science/article/pii/S0141938212000819.

[79] H. C. Longuet-Higgins and K. Prazdny. The interpretation of a moving retinal image. *Proceedings of the Royal Society of London. Series B. Biological Sciences* **208**.1173 (1980), 385–397. DOI: 10.1098/rspb.1980.0057.

[80] B. D. Lucas and T. Kanade. "An Iterative Image Registration Technique with an Application to Stereo Vision". *IJCAI'81: 7th international joint conference on Artificial intelligence*. Vol. 2. Vancouver, Canada, Aug. 1981, pp. 674–679. URL: https://hal.science/hal-03697340.

[81] G. Markkula et al. Sustained sensorimotor control as intermittent decisions about prediction errors: Computational framework and application to ground vehicle steering. *Biological cybernetics* **112** (2018), 181–207. DOI: 10.1007/s00422-017-0743-9. URL: https://doi.org/10.1007/s00422-017-0743-9.

[82] J. S. Matthis, K. S. Muller, and M. M. Hayhoe. Retinal optic flow and the control of locomotion. *Journal of Vision* **19**.10 (2019), 179–179. DOI: 10.1167/19.10.179.

[83] J. S. Matthis, J. L. Yates, and M. M. Hayhoe. Gaze and the control of foot placement when walking in natural terrain. *Current Biology* **28**.8 (2018), 1224–1233. DOI: 10.1016/j.cub.2018.03.008.

[84] J. S. Matthis et al. Retinal optic flow during natural locomotion. *PLOS Computational Biology* **18**.2 (2022), e1009575. DOI: 10.1371/journal.pcbi.1009575. URL: https://doi.org/10.1371/journal.pcbi.1009575.

[85] M. Menze, C. Heipke, and A. Geiger. "Joint 3d Estimation Of Vehicles And Scene Flow". *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. Vol. Ii-3/w5. 2015, pp. 427–434. DOI: 10.5194/isprsannals-II-3-W5-427-2015. URL: https://isprs-annals.copernicus.org/articles/II-3-W5/427/2015/.

[86] M. Menze, C. Heipke, and A. Geiger. Object Scene Flow. *ISPRS Journal of Photogrammetry and Remote Sensing* **140** (2018). Geospatial Computer Vision, 60–76. ISSN: 0924-2716. DOI: https://doi.org/10.1016/j.isprsjprs.2017.09.013. URL: https://www.sciencedirect.com/science/article/pii/S0924271617301727.

[87] N. Messikommer et al. "Data-Driven Feature Tracking for Event Cameras". *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2023, pp. 5642–5651. eprint: https://openaccess.thecvf.com/content/CVPR2023/papers/Messikommer_Data-Driven_Feature_Tracking_for_Event_Cameras_CVPR_2023_paper.pdf.

[88] P. Morasso. Spatial control of arm movements. *Experimental brain research* **42**.2 (1981), 223–227. DOI: 10.1007/bf00236911.

[89] M. M. Morrow and L. E. Miller. Prediction of Muscle Activity by Populations of Sequentially Recorded Primary Motor Cortex Neurons. *Journal of Neurophysiology* **89**.4 (2003). Pmid: 12612022, 2279–2288. DOI: 10.1152/jn.00632.2002. eprint: https://doi.org/10.1152/jn.00632.2002. URL: https://doi.org/10.1152/jn.00632.2002.

[90] B. Nguyen. Bio-inspired retinal optic flow perception in robotic navigation. *Chalmers University of Technology* **Licentiate Thesis** (2021).

[91] B. Nguyen and O. Benderius. *Closed raw data set for: Intermittent control and retinal optic flow when maintaining a curvilinear path*. Chalmers University of Technology, 2024. DOI: 10.5878/xnax-zk75.

[93] B. Nguyen and O. Benderius. *Open data set for: Intermittent control and retinal optic flow when maintaining a curvilinear path*. Chalmers University of Technology, 2024. DOI: 10.5878/vjkp-z436.

[94] B. Nguyen and O. Benderius. *Source code for: Intermittent control and retinal optic flow when maintaining a curvilinear path*. Version 1.0.0. Sept. 2024. DOI: 10.5281/zenodo.13683463. URL: https://doi.org/10.5281/zenodo.13683463.

[97] P. Nilsson, L. Laine, and B. Jacobson. A Simulator Study Comparing Characteristics of Manual and Automated Driving During Lane Changes of Long Combination Vehicles. *IEEE Transactions on Intelligent Transportation Systems* **18**.9 (Sept. 2017), 2514–2524. ISSN: 1558-0016. DOI: 10.1109/tits.2017.2664890.

[98] R. Nishizono, N. Saijo, and M. Kashino. Highly reproducible eyeblink timing during formula car driving. *iScience* **26**.6 (2023), 106803. ISSN: 2589-0042. DOI: 10.1016/j.isci.2023.106803. URL: https://www.sciencedirect.com/science/article/pii/S2589004223008805.

[99] H. Pashler. Dual-task interference in simple tasks: data and theory. *Psychological bulletin* **116**.2 (1994), 220. DOI: 10.1037/0033-2909.116.2.220.

[100] J. Pekkanen and O. Lappi. A new and general approach to signal denoising and eye movement classification based on segmented linear regression. *Scientific reports* **7**.1 (2017), 1–13. DOI: 10.1038/s41598-017-17983-x.

[101] J. Pekkanen et al. A computational model for driver's cognitive state, visual perception and intermittent attention in a distracted car following task. *Royal Society open science* **5**.9 (2018), 180194. DOI: 10.1098/rsos.180194.

[102] S. Perfiliev. "Programming and updating of reaching movements". English. PhD thesis. Göteborgs Universitet, 2003, p. 81. ISBN: 9162856596. eprint: http://hdl.handle.net/2077/15841.

[103] S. Perfiliev et al. Reflexive Limb Selection and Control of Reach Direction to Moving Targets in Cats, Monkeys, and Humans. *Journal of Neurophysiology* **104**.5 (2010). Pmid: 20810693, 2423–2432. DOI: 10.1152/jn.01133.2009.

[104] R. Plamondon. "On the Origin of Asymmetric Bell-Shaped Velocity Profiles in Rapid-Aimed Movements". *Tutorials in Motor Neuroscience*. Ed. by J. Requin and G. E. Stelmach. Dordrecht: Springer Netherlands, 1991, pp. 283–295. ISBN: 978-94-011-3626-6. DOI: 10.1007/978-94-011-3626-6_23.

[105] K. Prazdny. Egomotion and relative depth map from optical flow. *Biological cybernetics* **36**.2 (1980), 87–102. DOI: 10.1007/bf00361077.

[106] K. Prazdny. On the information in optical flows. *Computer Vision, Graphics, and Image Processing* **22**.2 (1983), 239–259. ISSN: 0734-189x. DOI: https://doi.org/10.1016/0734-189X(83)90067-1. URL: https://www.sciencedirect.com/science/article/pii/0734189X83900671.

[107] B. A. Purcell et al. Neurally constrained modeling of perceptual decision making. *Psychological review* **117**.4 (2010), 1113. DOI: 10.1037/a0020311.

[108] F. Raudies. Optic flow. *Scholarpedia* **8**.7 (2013). revision #149632, 30724. DOI: 10.4249/scholarpedia.30724.

[109] D. Raviv. "Visual looming". *Intelligent Robots and Computer Vision XI: Algorithms, Techniques, and Active Vision*. Ed. by D. P. Casasent. Vol. 1825. International Society for Optics and Photonics. Spie, 1992, pp. 703–713. DOI: 10.1117/12.131574. URL: https://doi.org/10.1117/12.131574.

[110] C. W. Reynolds. Flocks, herds and schools: A distributed behavioral model. *SIGGRAPH Comput. Graph.* **21**.4 (Aug. 1987), 25–34. ISSN: 0097-8930. DOI: 10.1145/37402.37406. URL: https://doi.org/10.1145/37402.37406.

[111] C. Rudin. Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead. *Nature Machine Intelligence* **1**.5 (2019), 206–215. DOI: 10.1038/s42256-019-0048-x.

[112] P. Salvini, C. Laschi, and P. Dario. Design for acceptability: improving robots' coexistence in human society. *International journal of social robotics* **2**.4 (2010), 451–460. DOI: 10.1007/s12369-010-0079-2.

[113] D. D. Salvucci and R. Gray. A Two-Point Visual Control Model of Steering. *Perception* **33**.10 (2004). Pmid: 15693668, 1233–1248. DOI: 10.1068/p5343. eprint: https://doi.org/10.1068/p5343. URL: https://doi.org/10.1068/p5343.

[114] D. D. Salvucci and J. H. Goldberg. "Identifying fixations and saccades in eye-tracking protocols". *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications*. Etra '00. Palm Beach Gardens, Florida, USA: Association for Computing Machinery, 2000, pp. 71–78. ISBN: 1581132808. DOI: 10.1145/355017.355028. URL: https://doi.org/10.1145/355017.355028.

[115] J. Sánchez Pérez, N. Monzón López, and A. Salgado de la Nuez. Robust Optical Flow Estimation. *Image Processing On Line* **3** (2013), 252–270. DOI: 10.5201/ipol.2013.21.

[116] R. Schweitzer and M. Rolfs. Intrasaccadic motion streaks jump-start gaze correction. *Science Advances* **7**.30 (2021), eabf2218. DOI: 10.1126/sciadv.abf2218. eprint: https://www.science.org/doi/pdf/10.1126/sciadv.abf2218.

[117] H. A. Simon. Rational choice and the structure of the environment. *Psychological review* **63**.2 (1956), 129. DOI: 10.1037/h0042769.

[118] T. Sjöblom et al. *REEDS: Reference data and algorithms for research and development of smart ships*. 2023. URL: https://urn.kb.se/resolve?urn=urn:nbn:se:ri:diva-65755.

[119] C. L. Smith, N. Pivovarova, and T. S. Reese. Coordinated feeding behavior in Trichoplax, an animal without synapses. *PloS one* **10**.9 (2015), e0136098. DOI: 10.1371/journal.pone.0136098.

[120] K. U. Smith, R. Kaplan, and H. Kao. Experimental systems analysis of simulated vehicle steering and safety training. *Journal of Applied Psychology* **54**.4 (1970), 364. DOI: 10.1037/h0029696.

[121] D. L. Strayer, F. A. Drews, and D. J. Crouch. A Comparison of the Cell Phone Driver and the Drunk Driver. *Human Factors* **48**.2 (2006). Pmid: 16884056, 381–391. DOI: 10.1518/001872006777724471. eprint: https://doi.org/10.1518/001872006777724471.

[122] H. Summala. "Towards Understanding Motivational and Emotional Factors in Driver Behaviour: Comfort Through Satisficing". *Modelling Driver Behaviour in Automotive Environments: Critical Issues in Driver Interactions with Intelligent Transport Systems*. Ed. by P. C. Cacciabue. London: Springer London, 2007, pp. 189–207. ISBN: 978-1-84628-618-6. DOI: 10.1007/978-1-84628-618-6_11. URL: https://doi.org/10.1007/978-1-84628-618-6%5C%5F11.

[123] M. Svärd et al. Computational modeling of driver pre-crash brake response, with and without off-road glances: Parameterization using real-world crashes and near-crashes. *Accident Analysis and Prevention* **163** (2021), 106433. ISSN: 0001-4575. DOI: 10.1016/j.aap.2021.106433. URL: https://www.sciencedirect.com/science/article/pii/S0001457521004644.

[124] M. Svärd et al. Using naturalistic and driving simulator data to model driver responses to unintentional lane departures. *Transportation Research Part F: Traffic Psychology and Behaviour* **100** (2024), 361–387. ISSN: 1369-8478. DOI: 10.1016/j.trf.2023.11.021. URL: https://www.sciencedirect.com/science/article/pii/S1369847823002656.

[125] M. Tao et al. SimpleFlow: A Non-iterative, Sublinear Optical Flow Algorithm. *Computer Graphics Forum* **31**.2pt1 (2012), 345–353. DOI: 10.1111/j.1467-8659.2012.03013.x. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-8659.2012.03013.x.

[126] H. R. Terry, S. G. Charlton, and J. A. Perrone. The role of looming and attention capture in drivers' braking responses. *Accident Analysis and Prevention* **40**.4 (2008), 1375–1382. ISSN: 0001-4575. DOI: 10.1016/j.aap.2008.02.009. URL: https://www.sciencedirect.com/science/article/pii/S0001457508000390.

[127] A. Thiele et al. Neural Mechanisms of Saccadic Suppression. *Science* **295**.5564 (2002), 2460–2462. DOI: 10.1126/science.1068788. eprint: https://www.science.org/doi/pdf/10.1126/science.1068788.

[128] S. Tuhkanen et al. Can gaze control steering? *Journal of Vision* **23**.7 (July 2023), 12–12. ISSN: 1534-7362. DOI: 10.1167/jov.23.7.12. eprint: https://arvojournals.org/arvo/content_public/journal/jov/938640/i1534-7362-23-7-12_1689933247.6478.pdf. URL: https://doi.org/10.1167/jov.23.7.12.

[129] S. Tuhkanen et al. Humans use predictive gaze strategies to target waypoints for steering. *Scientific reports* **9**.1 (2019), 1–18. DOI: 10.1038/s41598-019-44723-0.

[130] A. Tustin. The nature of the operator's response in manual control, and its implications for controller design. *Journal of the Institution of Electrical Engineers - Part IIA: Automatic Regulators and Servo Mechanisms* **94** (2 1947), 190–206. DOI: 10.1049/ji-2a.1947.0025. eprint: https://digital-library.theiet.org/doi/pdf/10.1049/ji-2a.1947.0025. URL: https://digital-library.theiet.org/doi/abs/10.1049/ji-2a.1947.0025.

[131] S. Uras et al. A computational approach to motion perception. *Biological Cybernetics* **60**.2 (1988), 79–87. DOI: 10.1007/bf00202895.

[132] E. Vagnoni, S. F. Lourenco, and M. R. Longo. Threat modulates perception of looming visual stimuli. *Current Biology* **22**.19 (2012), R826–r827. ISSN:

0960-9822. DOI: 10.1016/j.cub.2012.07.053. URL: https://www.sciencedirect.com/science/article/pii/S0960982212008780.

[133] S. Vedula et al. "Three-dimensional scene flow". *Proceedings of the Seventh IEEE International Conference on Computer Vision*. Vol. 2. 1999, 722–729 vol.2. DOI: 10.1109/iccv.1999.790293.

[134] M. Vince. The intermittency of control movements and the psychological refractory period. *British Journal of Psychology. General Section* **38**.3 (1948), 149–157. DOI: 10.1111/j.2044-8295.1948.tb01150.x. eprint: https://bpspsychub.onlinelibrary.wiley.com/doi/pdf/10.1111/j.2044-8295.1948.tb01150.x. URL: https://bpspsychub.onlinelibrary.wiley.com/doi/abs/10.1111/j.2044-8295.1948.tb01150.x.

[135] M. Wahde. *Biologically inspired optimization methods: an introduction*. WIT press, 2008.

[136] M. Wahde and M. Virgolin. "The five Is: Key principles for interpretable and safe conversational AI". *Proceedings of the 2021 4th International Conference on Computational Intelligence and Intelligent Systems*. Ciis '21. Tokyo, Japan: Association for Computing Machinery, 2022, pp. 50–54. ISBN: 9781450385930. DOI: 10.1145/3507623.3507632. URL: https://doi.org/10.1145/3507623.3507632.

[137] J. Wann and M. Land. Steering with or without the flow: is the retrieval of heading necessary? *Trends in cognitive sciences* **4**.8 (2000), 319–324. DOI: 10.1016/s1364-6613(00)01513-8.

[138] J. P. Wann and D. K. Swapp. Why you should look where you are going. *Nature neuroscience* **3**.7 (2000), 647–648. DOI: 10.1038/76602.

[139] W. H. Warren et al. Optic flow is used to control human walking. *Nature neuroscience* **4**.2 (2001), 213–216. DOI: 10.1038/84054.

[140] W. H. Warren and B. R. Fajen. "From Optic Flow to Laws of Control". *Optic Flow and Beyond*. Ed. by L. M. Vaina, S. A. Beardsley, and S. K. Rushton. Dordrecht: Springer Netherlands, 2004, pp. 307–337. ISBN: 978-1-4020-2092-6. DOI: 10.1007/978-1-4020-2092-6_14. URL: https://doi.org/10.1007/978-1-4020-2092-6%5C%5F14.

[141] P. Weinzaepfel et al. "DeepFlow: Large Displacement Optical Flow with Deep Matching". *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*. Dec. 2013. eprint: https://openaccess.thecvf.com/content_iccv_2013/papers/Weinzaepfel_DeepFlow_Large_Displacement_2013_ICCV_paper.pdf.

[142] D. Whitney, D. A. Westwood, and M. A. Goodale. The influence of visual motion on fast reaching movements to a stationary object. *Nature* **423**.6942 (2003), 869–873. DOI: 10.1038/nature01693.

[143] R. P. Wildes et al. Recovering Estimates of Fluid Flow from Image Sequence Data. *Computer Vision and Image Understanding* **80**.2 (2000), 246–266. ISSN: 1077-3142. DOI: https://doi.org/10.1006/cviu.2000.0874. URL:

https://www.sciencedirect.com/science/article/pii/S1077314200908749.

[144] R. M. Wilkie and J. P. Wann. Driving as Night Falls: The Contribution of Retinal Flow and Visual Direction to the Control of Steering. *Current Biology* **12**.23 (2002), 2014–2017. ISSN: 0960-9822. DOI: 10.1016/s0960-9822(02)01337-4. URL: https://www.sciencedirect.com/science/article/pii/S0960982202013374.

[145] R. M. Wilkie, J. P. Wann, and R. S. Allison. Active gaze, visual look-ahead, and locomotor control. *Journal of experimental psychology: Human perception and performance* **34**.5 (2008), 1150. DOI: 10.1037/0096-1523.34.5.1150.

[146] R. M. Wilkie and J. P. Wann. Eye-movements aid the control of locomotion. *Journal of Vision* **3**.11 (Nov. 2003), 3–3. ISSN: 1534-7362. DOI: 10.1167/3.11.3. eprint: https://arvojournals.org/arvo/content_public/journal/jov/933554/jov-3-11-3.pdf. URL: https://doi.org/10.1167/3.11.3.

[147] J. Wulff and M. J. Black. "Efficient Sparse-to-Dense Optical Flow Estimation Using a Learned Basis and Layers". *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. June 2015. eprint: https://openaccess.thecvf.com/content_cvpr_2015/papers/Wulff_Efficient_Sparse-to-Dense_Optical_2015_CVPR_paper.pdf.

[148] H. G. Yeom, J. S. Kim, and C. K. Chung. Brain mechanisms in motor control during reaching movements: Transition of functional connectivity according to movement states. *Scientific reports* **10**.1 (2020), 567. DOI: 10.1038/s41598-020-57489-7.

[149] C. Zach, T. Pock, and H. Bischof. "A Duality Based Approach for Realtime TV-L1 Optical Flow". *Pattern Recognition*. Ed. by F. A. Hamprecht, C. Schnörr, and B. Jähne. Berlin, Heidelberg: Springer Berlin Heidelberg, 2007, pp. 214–223. ISBN: 978-3-540-74936-3.

[150] H. Zhao and W. H. Warren. On-line and model-based approaches to the visual control of action. *Vision Research* **110** (2015). On-line Visual Control of Action, 190–202. ISSN: 0042-6989. DOI: 10.1016/j.visres.2014.10.008. URL: https://www.sciencedirect.com/science/article/pii/S0042698914002417.