

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

MATHIAS FASSINI MANTELLI

**Exploiting semantic information in indoor
environments**

Ph.D. Thesis Proposal

Advisor: Profa. Dra. Mariana Luderitz Kolberg
Coadvisor: Prof. Dr. Renan de Queiroz Maffei

Porto Alegre
October 2021

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Carlos André Bulhões

Vice-Reitora: Prof^a. Patricia Pranke

Pró-Reitor de Pós-Graduação: Prof. Celso Giannetti Loureiro Chaves

Diretora do Instituto de Informática: Prof^a. Carla Maria Dal Sasso Freitas

Coordenador do PPGC: Prof. Claudio Rosito Jung

Bibliotecária-chefe do Instituto de Informática: Beatriz Regina Bastos Haro

*“If I have seen farther than others,
it is because I stood on the shoulders of giants.”*

— SIR ISAAC NEWTON

AGRADECIMENTOS

Agradeço ao \LaTeX por não ter vírus de macro...

ABSTRACT

Este documento é um exemplo de como formatar documentos para o Instituto de Informática da UFRGS usando as classes L^AT_EX disponibilizadas pelo UTUG. Ao mesmo tempo, pode servir de consulta para comandos mais genéricos. *O texto do resumo não deve conter mais do que 500 palavras.*

Keywords: Formatação eletrônica de documentos. Padronização de documentos. Instituto de Informática da UFRGS. L^AT_EX. ABNT. UFRGS.

Using L^AT_EX to Prepare Documents at II/UFRGS

RESUMO

This document is an example on how to prepare documents at II/UFRGS using the L^AT_EX classes provided by the UTUG. At the same time, it may serve as a guide for general-purpose commands. *The text in the abstract should not contain more than 500 words.*

Palavras-chave: Electronic formatting of documents. Instituto de Informática da UFRGS. L^AT_EX. ABNT. UFRGS.

LIST OF FIGURES

Figure 1.1 plane.	12
------------------------	----

LIST OF TABLES

LIST OF ABBREVIATIONS AND ACRONYMS

ABNT Associação Brasileira de Normas Técnicas

NUMA Non-Uniform Memory Access

SIMD Single Instruction Multiple Data

SMP Symmetric Multi-Processor

SPMD Single Program Multiple Data

CONTENTS

1 INTRODUCTION.....	11
REFERENCES.....	15
APPENDIX A — RESUMO EXPANDIDO	16

1 INTRODUCTION

The first decades of research in Mobile Robotics, from the beginning until 2004, handled the challenges of connecting efficiency and data association. They introduced probabilistic formulations to path planning, exploration, simultaneous localization and mapping (SLAM), and many other areas. Some of the approaches from these areas are still popular nowadays, such as RaoBlackwellised Particle Filters and Extended Kalman Filters. The majority of them were based on ultrasonic or lidar sensors, as they were the most popular and robust sensors at the time. Consequently, the outcome maps were mostly 2D grid ones, in which the cells represented the free, occupied, and unknown regions.

After building a solid foundation, the research community moved forward, concentrating on improving the properties like observability, convergence, and consistency of the already proposed and the new approaches. Using visual sensors as one of the main ways to read the environment is another highlight for this period (2004-2015), given the considerable improvement in such sensors regarding the data quality and cameras' size and price. In fact, building 2D and 3D maps from the environment with a visual sensor resulted in a new term, Visual SLAM.

Simultaneously to algorithmic advances, mobile robotics shifted its focus from factory floors and assembly lines to everyday living spaces. Mobile robotics is increasingly demanded in our daily lives, whether with simple vacuum cleaners or complex autonomous cars. However, this demand for robots to perform high-level tasks in different scenarios, such as a service robot interacting with the objects within the environment or a ground robot avoiding mud terrains, revealed the geometric robots' perception weaknesses.

Despite the progress on the software and hardware fronts, the researchers realized the limitations of purely geometric maps and that the robot's perception should be improved. For example, a vacuum cleaner robot a few years ago would be asked to clean all the free spaces within the environment and avoid obstacles. A 2D grid lidar-based SLAM would be enough for this task, as the robot would map the free space and avoid the obstacles. In contrast, now it has to clean the kitchen on Mondays and the living room on Wednesdays, which brings the question: "what is a kitchen for the robot? Is there a sensor that reads house rooms and informs what each room is?". Hence, the difference between the two versions of robots in this example is the capability of going beyond basic

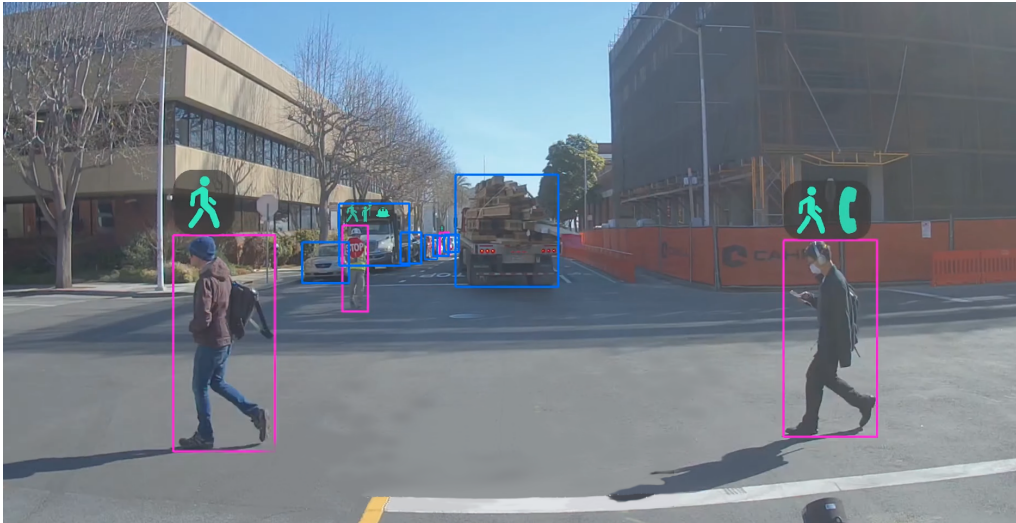


Figure 1.1 – plane.

geometry representations to obtain a high-level understanding of the environment.

The association of semantic information (or concepts) to geometric entities in the map is called semantic mapping, one of the newest topics the researchers have explored. It enhances the robot's autonomy and robustness in many ways, besides facilitating some high-level tasks. Fig. 1.1 is an image from Zoox's autonomous car, and it illustrates the advantage of using semantic information in robotics tasks. The car would probably map this scene with its geometric perception as four obstacles in its front, and two are closer than the other two. Differently, with a semantic perception, the car estimates three people and a truck within the scene. Most importantly, it estimates that one person is distracted using his phone, and another is holding a stop sign. Combining the detection of a walking person and a phone allows the car to estimate the semantic information that this person is likely distracted. Hence, the car should drive itself even more carefully. This whole process is natural for human drivers, but the same can not be said about robots.

As semantic information is more like a specific knowledge inferred from the robot's surroundings than a specific type of data from a sensor reading, several questions need to be answered before using it in a robotic task. We see the following as noteworthy challenges:

- Deciding on what type of semantic information is possible to infer or estimate from the robot's surroundings that is relevant to the task
- How to perform the inferring or estimate the semantic information
- How to use the semantic information to improve the robot's performance in a given task

The first point is frequently discussed in its geometric version, as semantic information is relatively new in the literature. Briefly, for the context of a given robotic task, what information is not explicitly in the environment but could be inferred or estimated to improve the robot's performance? This demands a deep understanding of the task and the general environment characteristics where the robot operates. An inspiration for answering this point is to consider how humans behave and solve such a task and how we connect and process the environment's information to accomplish the task efficiently.

Second, depending on the needed semantic information, it may be necessary to use methods based on machine learning to estimate it. For example, training a deep learning model for estimating terrain traversability for an outdoor ground robot may provide a suitable result. However, besides the training requirement, the solution's quality depends on the training data, and this approach does not scale well. Probabilistic-based estimations appear as a second option, as it does not require a large set of data for training, and accepts a wide range of different models.

The third and last point, the proper use of the inferred semantic information in the robot's system, is crucial for successful task completion. As the robot gains more information from the environment, it is important to keep updating the estimations, and it is even better if the estimations become more robust over time.

The exploitation of semantic information in robotics is an idea that has recently gained attention from researchers, and thus, most of the challenging problems are still unsolved. A simply way of pushing the limits further and exploring these problems is to study the advantages of semantic information in different areas. We have chosen a task with a high difficulty level that can benefit from semantic information, object search (OS) in indoor and unknown environments, a yet unsolved problem in robotics.

In OS tasks, the robot's goal is to find a target object in the environment with a visual sensor. Usually, the environment is unknown to the robot, and it only guides its moves with the clues it finds out during the search by its sensor readings.

The first years of research in the field of Mobile Robotics saw the introduction of many probabilistic formulations for SLAM, path planning, -At the beginning, Robotics was interested in estimating the obstacle's positions and the free space in the environment (Robotics has started with robots operating in assembly lines in factories, and now it is shifting to everyday living spaces)

-However, this field has evolved and expanded the varieties of places the robots op-

erate -By operating in more different environment and developing many tasks, researchers started including many sensors to the robots in order to make it more capable of acquiring data -However, that is not necessary if we can process the sensor readings and estimate more information besides the raw data

Robotics has been changing its focus from factory floors to everyday living spaces, such as offices, houses, hospitals, airports, and etc (AYDEMIR, 2012).

REFERENCES

AYDEMIR, A. **Exploiting structure in man-made environments**. Thesis (PhD) — KTH Royal Institute of Technology, 2012.

APPENDIX A — RESUMO EXPANDIDO

Resolução 02/2021 – Redação de Teses e Dissertações em Inglês Dissertações de Mestrado e Teses de Doutorado do PPGC, bem como outros trabalhos escritos tais como Proposta de Tese e PEP, poderão ser redigidas em inglês desde que contenham um título e resumo expandido redigidos em português. O resumo expandido deve conter no mínimo duas páginas inteiras, deve aparecer como apêndice e deve conter as principais contribuições e resultados do trabalho.