

Proceeding Paper

Assessment of SLAM Methods Applied in Monochromatic Environments [†]

Rudolf Krecht * and Áron Ballagi

Department of Automation and Mechatronics, Széchenyi István University, Egyetem tér 1, H-9026 Győr, Hungary; ballagi@ga.sze.hu

* Correspondence: krecht.rudolf@ga.sze.hu

[†] Presented at the Sustainable Mobility and Transportation Symposium 2024, Győr, Hungary, 14–16 October 2024.

Abstract: One of the most significant challenges in sustainable autonomous mobile robot and vehicle development is the perception of stochastic environments. Various environmental perception methods have been proposed to address these challenges; however, these methods often lack general applicability. Many of these methods rely on environmental feature extraction, which can fail in specific scenarios, such as monochromatic environments. This article aims to evaluate existing SLAM (Simultaneous Localization and Mapping) methods that utilize camera or combined camera and LiDAR input data in predominantly monochromatic environments. Additionally, this study seeks to identify performance issues in such applications.

Keywords: SLAM; monochromatic environments; LiDAR; camera

1. Introduction

The development of different autonomous mobility solutions, including self-driving passenger vehicles (such as robotaxis and autonomous shuttle buses), has significant potential to increase the sustainability of everyday transportation. However, autonomous transportation has not been implemented with a sufficient number of vehicles, and in addition to financial and legislative reasons, there are still many technological reasons behind this metric. One of the most challenging technology-related questions regarding self-driving transportation is environmental perception in stochastic environments.

Simultaneous Localization and Mapping (SLAM) is crucial for autonomous environmental perception, enabling robots to navigate and build maps for future use in completely unknown environments. SLAM methods are essential for various applications, including almost all kinds of autonomous vehicles from small ground-based or aerial delivery robots to passenger vehicles such as taxis and shuttle buses.

The spreading of solutions based on autonomous transportation often begins with the introduction of self-driving vehicles in closed, highly manageable areas, such as parks or campuses. Despite reducing the stochasticity of the environment, these locations are often uniform, lacking diverse environmental features. This phenomenon raises new questions, since many SLAM methods rely on feature extraction. One particular environment type that raises questions regarding the applicability of camera-based SLAM methods consists of locations containing objects that manifest a limited number of colors, or different shades of only one color, often referred to as monochromatic environments.

The aim of this paper is to list, compare, and evaluate vision-based SLAM methods regarding their applicability in monochromatic environments based on a literature review. The main focus of the comparison lies on SLAM methods that can be implemented using input from exteroceptive sensors that gather colored image information with or without additional depth data (monocular color cameras or stereo/depth color cameras). This paper is the precursor of future work based on the implementation of a camera-based SLAM method in a monochromatic environment, and since there are numerous available



Citation: Krecht, R.; Ballagi, Á. Assessment of SLAM Methods Applied in Monochromatic Environments. *Eng. Proc.* **2024**, *79*, 51. <https://doi.org/10.3390/engproc2024079051>

Academic Editors: András Lajos Nagy, Boglárka Eisinger Balassa, László Lendvai and Szabolcs Kocsis-Szürke

Published: 6 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

SLAM methods, it is considered a necessary first step to assess them based on the available literature before their implementation.

2. Methods

2.1. Comparison Based on Literature Review

Environmental perception for autonomous vehicles is one of the most important fields, making methods such as SLAM extensively researched and varied. This variety leads to many approaches from different perspectives, suited for the needs of individual use cases. Thus, research and development based on SLAM methods requires the comparative assessment of existent possible approaches [1]. This literature review-based comparison procedure requires the creation of a system of criteria then, the selected SLAM methods are assessed based on this system. Finally, a comparison can be set up, enabling a final decision to be reached.

In the case of this study, the system of criteria is constructed based on the key requirements of autonomous transportation systems' environmental perception: the type of input data, determined by the available exteroceptive on-board sensors; the expected accuracy of the output; robustness, i.e., the handling of lighting changes and dynamic obstacles; the hardware requirements (some autonomous vehicles feature limited computational power); the type of output map; and, finally, the results in monochromatic, feature-poor environments.

The four selected methods for comparison all use camera images as their input data. These methods encompass both feature-based and direct techniques, allowing for a comprehensive evaluation of their performance in feature-poor and monochromatic environments. Their widespread recognition in research and varying scalability make them suitable for diverse use cases, from small-scale indoor navigation to large-scale outdoor mapping. This selection ensures a thorough assessment of the methods' accuracy, robustness, and computational efficiency, particularly in challenging environments.

2.2. Assessed SLAM Methods

2.2.1. ORB-SLAM2

ORB-SLAM2 is a feature-based visual SLAM method that utilizes Oriented FAST and Rotated BRIEF (ORB) features [2]. It supports the usage of monocular, stereo, and RGB-D cameras as input data. ORB-SLAM2 is known for its robustness and accuracy, and it has functionalities such as tracking, mapping, and loop closure. Its reliance on visual features makes it effective in feature-rich environments but challenging to use successfully in monochromatic settings [3]. Because of its high reliance on features, it is generally expected to perform poorly in monochromatic environments.

2.2.2. DSO

DSO is a direct visual odometry method that uses pixel intensities rather than basing the SLAM process on feature extraction. Its relatively high accuracy and efficiency originate from the method on which this approach is based. It operates on sparse sets of pixel patches. DSO can be successfully implemented using a monocular camera. Its main drawback manifests in environments with changing lighting conditions, as it may show dropping accuracy in the event of drastic lighting changes [4].

2.2.3. LSD-SLAM

LSD-SLAM (Large-Scale Direct Monocular SLAM) is a direct SLAM method that, similarly to DSO, processes pixel intensities directly rather than relying on feature extraction. It is capable of real-time, semi-dense 3D reconstruction using a monocular camera. LSD-SLAM is advantageous in large-scale environments but can be sensitive to lighting conditions and computationally demanding [5]. Furthermore, amongst many feature-based approaches, LSD-SLAM might be a feasible option in case of feature-poor environments.

2.2.4. RTAB-Map

RTAB-Map (Real-Time Appearance-Based Mapping) is a graph-based visual SLAM method that is intended to perform real-time processing. It supports all three of the usual camera types for robotic applications—RGB-D, stereo, and monocular cameras. Processing is based on visual and spatial data. Its most significant advantage is that it can handle large environments and can perform loop-closure detection with high efficiency [6].

2.2.5. PTAM

PTAM (Parallel Tracking and Mapping), as its name suggests, separates tracking and mapping tasks into two different, parallel threads [7].

3. Results

The results of the comparison carried out based on relevant literature is presented according to the criteria defined before. A comparison is set up based on this system of criteria, presented in Table 1.

Table 1. Comparative assessment of the ORB-SLAM2, DSO, LSD-SLAM, RTAB-Map, and PTAM methods.

Features, Characteristics	ORB-SLAM2	DSO	LSD-SLAM	RTAB-Map	PTAM
Input data	Images from monocular, stereo, or RGB-D cameras	Images from monocular camera	Images from monocular camera	Images from monocular, stereo, or RGB-D cameras	Images from monocular camera
Expected accuracy	High, but only in the case of feature-rich environments	High in texture-rich, well-lit environments	Moderate in texture-rich environments	High in large environments	High in small-scale, static environments
Robustness	Low in dynamic environments; high in static, consistent environments	Acceptable in texture-poor environments, but sensitive to lighting changes	High in large environments; low in case of changes in lighting conditions	Robust in large-scale environments	Limited to small-scale, static environments
Hardware requirements	High (feature extraction and matching)	Moderate for real-time 2D mapping	High (direct pixel intensity processing)	Efficient; well suited for real-time applications	Efficient in small-scale, real-time applications
Map type	Sparse 2D or 3D	Sparse 3D map	Semi-dense 3D	Dense 3D map	Sparse 2D map
Implementation	Moderate; requires calibration and related setup	Easy, with careful setup	Moderate, highly relies on dependencies	Easily adaptable	Easy in case of small, simple environments
Results in monochromatic environments	Reduced performance	Reduced	Good performance with proper lighting conditions	Moderate performance depending on features and lighting	Low in feature-poor, poorly lit environments

3.1. Accuracy

ORB-SLAM2 demonstrates high accuracy in environments with distinct visual features, but this performance can significantly drop due to the lack of detectable features.

DSO, since it is a direct method relying on pixel intensity values, manifests high accuracy in environments where feature extraction might not lead to acceptable results. Furthermore, it can provide precise odometry. It is important to mention that its accuracy can only be expected if proper lighting conditions are provided.

LSD-SLAM provides semi-dense 3D maps with good accuracy in textured environments. Its direct approach allows it to work reasonably well in monochromatic settings, although it can struggle with lighting variations.

RTAB-Map usually tends to be less accurate in general environments than methods similar to ORB-SLAM2, but it works well in large-scale environments and provides correct loop-closure detection.

PTAM shows its best accuracy in small-scale environments without the presence of dynamic objects. Its accuracy drops if it is applied in larger or highly complex environments.

3.2. Robustness

ORB-SLAM2 can be highly robust in environments with consistent and rich features, but it loses this characteristic in dynamic or low-texture settings.

DSO is highly robust in problematic environments such as areas with a limited number of features, colors or textures. However, it is nearly impossible to implement this method in environments with illumination that presents rapid changes in color or brightness.

LSD-SLAM is considered less robust because of its sensitivity to lighting changes and low-texture environments, but it can be still effective in large-scale environments with adequate lighting.

RTAB-Map, when applied in large-scale environments, is a highly robust algorithm. However, from the main point of view of this article, it lacks robustness, since it cannot perform well in feature-poor or monochromatic environments.

PTAM is less robust than ORB-SLAM2 or RTAB-Map. Its applicability is limited to small-scale environments.

3.3. Computational Efficiency

ORB-SLAM2 is computationally intensive due to the processes involved in feature extraction and matching. It requires significant processing power, especially in real-time applications.

DSO's computational efficiency depends heavily on its application. It can be considered computationally efficient compared to fully dense methods, but its real-time applicability depends on the optimization level of the given implementation. Furthermore, its efficiency is significantly compromised by large changes in the lighting conditions.

LSD-SLAM is, again, computationally demanding because it processes large quantities of image data directly. It requires careful management of computational resources.

RTAB-Map was designed for real-time performance even in large-scale environments. Because of techniques such as visual bag-of-words for appearance-based loop closure, it usually performs well from the point of view of efficiency.

PTAM is computationally efficient in small-scale environments. In these environments, it manifests easy real-time applicability. However, this efficiency does not scale well for large-scale environments.

3.4. Implementation Workflow

ORB-SLAM2 is moderately easy to set up. It requires careful calibration and tuning of parameters. It is supported by comprehensive documentation and active community support.

DSO is generally straightforward to implement, but because of sensitivity to lighting conditions, its input sensor requires careful installation.

LSD-SLAM is generally easy to set up. It requires handling of dependencies and management of computational load, but it has active support in academic research with various forks and improvements available. It has had less industrial adoption than ORB-SLAM2.

RTAB-Map has well-documented, modular implementation that supports monocular, stereo, and RGB-D cameras. It is supported by a large user community.

PTAM is very simple to set up for simple, small-scale applications but highly difficult to implement in larger, more complex environments.

4. Discussion

The comparison presented in this paper highlights the strengths and weaknesses of each SLAM method, focusing on general criteria important from the point of view of environmental perception by autonomous vehicles. Furthermore, an assessment of behavior in monochromatic environments is included. Among the four compared SLAM methods, ORB-SLAM2 excels in feature-rich environments but faces challenges in low-texture settings. DSO, with its direct sparse odometry approach, performs well in low-texture environments but is sensitive to drastic lighting changes. LSD-SLAM offers an alternative with its direct approach, handling large-scale environments effectively. Its most significant drawback is that it requires the careful and proper management of lighting conditions and computational resources. RTAB-Map, designed for real-time appearance-based mapping, is particularly robust in large-scale environments with efficient loop-closure detection but may struggle in feature-poor environments. Lastly, PTAM, though primarily used in small-scale augmented reality applications, provides high accuracy in controlled environments but is limited in its robustness and scalability for more complex settings.

5. Conclusions and Future Work

In conclusion, if applicability in monochromatic, feature-poor environments is an important factor, DSO and LSD-SLAM can be considered viable options amongst the assessed SLAM methods. DSO performs well in texture-poor environments but may struggle with rapidly changing lighting conditions. LSD-SLAM, with its direct approach, is suitable for large-scale, feature-poor environments, though careful management of lighting and computational resources is necessary. In the case of 3D outdoor environments that lack distinct features and textures, the application of LSD-SLAM is recommended.

Future work will include the assessment of LSD-SLAM with real environmental data acquired in specific monochromatic environments, using the same sensor set and similar lighting conditions.

Author Contributions: Literature review, Á.B. and R.K.; methodology, Á.B.; comparison, R.K.; writing—original draft preparation, R.K. All authors have read and agreed to the published version of the manuscript.

Funding: The publication was created in the framework of the Széchenyi István University's VHFO/416/2023-EM_SZERZ project entitled 'Preparation of digital and self-driving environmental infrastructure developments and related research to reduce carbon emissions and environmental impact' (Green Traffic Cloud).

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Rauf, A.; Irshad, M.J.; Wasif, M.; Syed, U.R.; Aziz, N.; Taj, H. Comparative Study of SLAM Techniques for UAV. *Eng. Proc.* **2021**, *12*, 67.
2. Mur-Artal, R.; Montiel, J.M.; Tardos, J.D. ORB-SLAM: A versatile and accurate monocular SLAM system. *IEEE Trans. Robot.* **2015**, *31*, 1147–1163. [[CrossRef](#)]
3. Mur-Artal, R.; Tardós, J.D. ORB-SLAM2: An open-source SLAM system for monocular, stereo, and RGB-D cameras. *IEEE Trans. Robot.* **2017**, *33*, 1255–1262. [[CrossRef](#)]
4. Engel, J.; Koltun, V.; Cremers, D. Direct Sparse Odometry. *IEEE Trans. Pattern Anal. Mach. Intell.* **2018**, *40*, 611–625. [[CrossRef](#)] [[PubMed](#)]
5. Engel, J.; Schöps, T.; Cremers, D. LSD-SLAM: Large-scale direct monocular SLAM. In Proceedings of the European Conference on Computer Vision, Zurich, Switzerland, 6–12 September 2014.

6. Labbé, M.; Michuad, F. RTAB-Map as an open-source lidar and visual simultaneous localization and mapping library for large-scale and long-term online operation. *J. Field Robot.* **2019**, *36*, 416–446. [[CrossRef](#)]
7. Klein, G.; Murray, D. Parallel tracking and mapping for small AR workspaces. In Proceedings of the 6th IEEE and ACM International Symposium on Mixed and Augmented Reality, Nara, Japan, 12–15 November 2007.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.