

# A Review of Topological Map Construction Methods for Indoor Robot Localization and Navigation

Wen Liu<sup>1\*</sup>, Ran Li<sup>1\*</sup> and Zhongliang Deng<sup>1</sup>

<sup>1</sup> School of Electronic Engineering, Beijing University of Posts and Telecommunications, Beijing 100876, China

## Abstract

In indoor scenes, robots typically rely on prior maps to obtain environmental information for localization and navigation. The review first compares the characteristics of metric maps, semantic maps, and topological maps, which are commonly used in indoor scenes, emphasizing the advantages and potential of topological maps as a high-level representation of environmental structure in the application of indoor robot localization and navigation. It introduces methods for constructing topological maps based on visibility graphs, graph partitioning, landmark features, and graph appearance, briefly analyzing their respective advantages and challenges. The review further explores multi-level expression methods that combine topological maps with other environmental information, focusing on the multi-dimensional information representation that combines topological structure with semantic mapping, and looks forward to its future development direction of intelligent interactive applications for indoor robots.

## Keywords

Indoor positioning and navigation, topological maps, semantic information, multi-level map

## 1. Introduction

Indoor positioning and navigation for robots can be summarized into three key questions: the robot needs to know "Where am I?", "Where do I want to go?", and "How do I get there?" [1]. This involves how the robot perceive, explore, acquire, and understand the surrounding environmental model, as well as how to plan and execute motion strategies effectively. First and foremost, this requires the construction of corresponding indoor maps as a priori basic reference, enabling the robot to complete complex tasks effectively.

Initially, researchers used metric maps to construct environmental models. But over time, issues such as high computational demand, and difficulty in maintenance have increasingly become apparent, making them insufficient to meet the needs of intelligent applications for indoor robots. Therefore, finding more advanced ways of environmental expression and developing navigation behaviors that are closer to those of humans has become a new direction for challenges. By leveraging the concept of graph theory, topological maps describe the environment based on topological structures, focusing only on the connectivity between nodes rather than precise geographical coordinates [2], eliminating the need for fine map construction, and thus being more lightweight, providing high-level positioning and navigation information for indoor robots.

This review will compare the commonly used types of indoor maps to highlight the unique advantages of topological maps in robot applications. Based on the summary of indoor topological map construction methods, it will explore the multi-level map representation with topological maps as the main part, aiming to provide better map services for indoor robots to achieve positioning and navigation, and to perform complex and advanced tasks.

---

*Proceedings of the Work-in-Progress Papers at the 14th International Conference on Indoor Positioning and Indoor Navigation (IPIN-WiP 2024), October 14 - 17, 2024, Hong Kong, China*

\* Corresponding author.

✉ liuwen@bupt.edu.cn (W. Liu); liran@bupt.edu.cn (R. Li); dengzhl@bupt.edu.cn (Z. Deng)

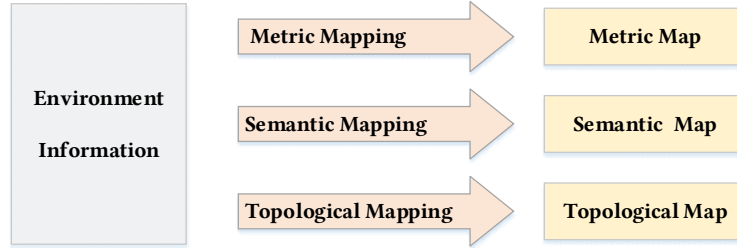
ORCID 0000-0002-6450-1969 (W. Liu); 0009-0005-2934-9358 (R. Li)



© 2024 Copyright for this paper by its authors. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

## 2. Types of Indoor Map

For indoor scenes, different mapping methods can be used to construct various types of maps. Fig. 1 illustrates the mapping relationship between the environmental information and the map. This chapter will compare and introduce these three types of indoor maps.



**Figure 1:** Environment Information and the Mapping Relationship with Maps

**Metric maps** refer to maps that describe the environment through mathematical metric mapping, representing the environment with real physical dimensions. Metric maps can be constructed based on point clouds [5][6], or more typically, by discretizing the environment into grid cells [3][4], which encode certain metrics to represent environmental information (such as occupancy rate, distance from obstacles, etc.). However, due to the need for rich information to accurately represent the environment, the construction of metric maps requires a significant consumption of computational resources and storage space, and for most indoor robot tasks, much of this information is unnecessary [22]. Secondly, because the construction of metric maps demands a high level of granularity, their quality is greatly affected by the precision and noise of the sensors, and it is difficult to update and maintain them in dynamic environments.

**Semantic maps** map the semantic information of objects, areas, etc., in the environment, providing robots with more in-depth conditions for understanding the environment. Semantic maps provide an environmental representation with elements of high-level abstraction [36], furthering the logical representation of the environment. It means that robots can infer new information through additional knowledge of global elements [3], greatly enhancing their capabilities. However, due to the construction of semantic maps relies on computer vision algorithms[7][8][9][10], the accuracy and robustness of the algorithms and models used will directly affect the quality of the semantic maps. In addition, in actual indoor scenes, changes in the position and state of objects may make the updating and maintenance of semantic maps difficult [40].

**Topological maps** use topological mapping to represent spatial relationships, simplifying the detailed description of actual geographical environmental information, focusing on the relative positions and connection relationships between nodes in the geographical space. The mathematical relationship of topological maps can generally be expressed by Eq:

$$G = (V, E), \quad (1)$$

Where G refer to the topological map, V refer to the set of vertex information in the topological graph, and E refer to the matrix of edges. In practical applications, nodes can represent key positions in the geographical environment, such as rooms, doors, corridors, or other feature points, and edges represent the connection relationships between nodes, such as adjacency, direction, accessibility, etc.

In the field of positioning and navigation for indoor robot, topological maps are a low-granularity structural representation method that can reduce the complexity of maps by abstracting the environment, thereby reducing information storage and lowering computational complexity. Since topology itself is a branch of mathematics that studies shapes and spaces, especially the properties that remain unchanged under continuous deformation of space, topological maps have a certain degree of flexibility and adaptability. And It naturally have a certain tolerance for dynamic environments. Even in the face of environmental changes or the appearance of dynamic obstacles, researchers can more easily maintain and update them.

### **3. Construction Methods of Indoor Topological Map**

Constructing a topological map begins with using various sensors to acquire data from the environment, followed by utilizing different algorithms to extract the topological structure from the data and represent it as a graph or network. This chapter mainly introduces several commonly used methods for constructing topological maps.

#### **3.1. Constructing Topological Map Based on Voronoi Diagram**

The Voronoi diagram is a graphical representation method based on the concept of geometric distance. It effectively represents the spatial structure and relationships by dividing space into multiple regions associated with a set of input points. The boundaries of Voronoi regions are composed of the perpendicular bisectors between adjacent input points. The Voronoi diagram provides positional information of points in space and their relationships with the nearest neighboring points [12][13], making it suitable for robot motion planning and obstacle avoidance and supporting fast and efficient nearest neighbor searches. As in references [12], during the process of constructing a topological map, the junction points or endpoints of pruned Voronoi edges can be used as topological nodes, with Voronoi paths between nodes considered as edges, thus converting the Voronoi diagram into a topological map. In three-dimensional space, reference [15] uses the three-dimensional generalized Voronoi diagram (GVD) extracted from the Euclidean signed distance field [5] to represent the topological structure of the spatial environment and generate a thin skeleton graph [14]. The sparse topological map can resist changes in noise and resolution and is suitable for micro aerial robots to perform indoor navigation tasks. The application of these methods has made the Voronoi diagram a powerful tool in the field of robot navigation from the beginning.

#### **3.2. Constructing Topological Map Based on Graph Partitioning**

Constructing topological map based on graph partitioning methods typically involves dividing acquired RGB images or metric maps into different subgraphs as nodes, with the connections between subgraphs serving as edges. The aim is to maximize connectivity within the same group while minimizing connectivity between different cluster nodes. Spectral clustering [16][17][18] is one of the commonly used graph partitioning algorithms, which shares the characteristic of using an affinity matrix as input [17]. This matrix typically uses Euclidean distance or other metrics to describe the similarity between data points. There are inherent disadvantages to using spectral clustering algorithms to generate topological maps [19], such as high computational costs when the input affinity matrix is large, and issues with excessive nodes and non-repetitive results. In Reference [21], the environment is modeled with "appearance graphs" using visual camera poses as low-level map nodes. The normalized cut criterion [20], one of the graph partitioning method, is then used to cluster nodes and construct higher-level mappings. This approach can be considered a precursor to constructing topological maps based on graph appearance.

Methods for constructing indoor topological maps based on Voronoi diagrams and graph partitioning, although more structured and hierarchical compared to contrast metric maps, are still not concise enough, and have many constraints during construction and a relatively fixed range of applicable environments. Therefore, in the subsequent research on topological map construction methods, they serve more as auxiliary tools.

#### **3.3. Constructing Topological Map Based on Landmark Features**

The method based on landmark features utilizes feature points in the environment, such as corner points, doors, rooms, etc., as nodes, and the distance or direction between feature points as edges to construct a topological map. Feature points can be obtained from the environment through sensors; for example, Reference [23] uses a 3D sensor to acquire depth information and develops a progressive Bayesian classifier for directly identifying different types of corridors (such as dead ends, T-junctions, crossroads, etc.). It abstracts the environment into a topological map with rooms or intersections as

nodes and corridors as edges, integrating information from multiple observations to extract features. Reference [26] obtains stable visual landmarks from videos as nodes (such as doors, fire extinguishers, elevators, etc.) and constructs a topological map using continuous sequences as connectivity information. In addition, feature points can be extracted from prior environmental information, such as Reference [24] which is based on a 3D indoor map model from Building Information Modeling (BIM), extracting elements like doors, windows, facilities, rooms, etc., as nodes, with edges including corridors and connection relationships. It also proposes step nodes for assisting indoor positioning and navigation, adaptable to complex and open indoor environments. Reference [25], on the other hand, extracts elements from CAD drawings and analyzes their topological relationships to construct an object-oriented topological structure.

### 3.4. Constructing Topological Map Based on Graph Appearance

Methods based on graph appearance primarily use visual information to construct topological maps. One such method represents the robot's world environment as a collection of linked waypoint images, that is, using images as nodes, creating edges between consecutive images and uses image matching methods for localization and navigation. Reference [27] establishes nodes based on positional visibility and assigns edges with spatial distance information and the navigability probability between two nodes, allowing intelligent agents to form long-term plans and navigate in new environments without prior knowledge of specific environments.

Another more typical method is the planning method based on topological memory [29][30][31][33][34][35]. Topological memory is a memory map where each node corresponds to a past observation of the robot. The SPTM [28], as a representative of topological memory, establishes nodes by interacting with the environment at discrete time steps. SPTM builds a dense topological map using image similarity as accessibility. Reference [36] specifically learns an accessibility estimator to predict the probability of reachability and sparsely reduces dense trajectories to anchor observation sequences, using anchor observation values as nodes and assigning edge weights based on reachability probability to construct a sparser topological map. Furthermore, Reference [37] proposes a graph maintenance strategy to improve lifelong navigation performance by eliminating incorrect edges and expanding the graph as needed. Reference [48] merges multiple sparse trajectories into a single topological map suitable for localization and navigation planning, using RGB-D panoramic images as nodes and additionally attaching rough geometric information to the directed edges in the map, enhancing the robot's global navigation capabilities.

**Table 1**  
Comparison of Indoor Topological Map Construction Methods

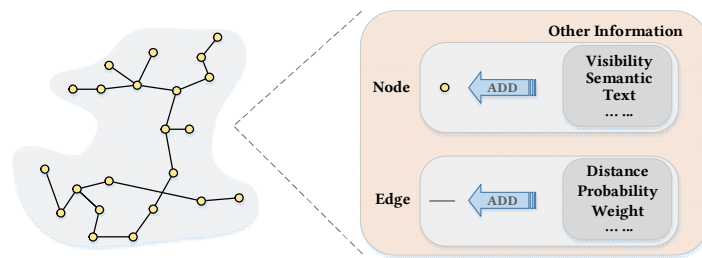
Method	Features	Applicable Scenarios
3.1	Emphasizes geometric distance and spatial division	Static environments / Obstacle avoidance tasks
3.2	Focuses on maximizing regional connectivity	Existing foundational maps
3.3	Concentrates on prominent feature points in the environment	Structured scenarios
3.4	Utilizes visual information, capable of real-time exploration and map construction	Relying on visual information / Real-time construction

Table 1. compares the four methods of topological map construction for indoor scenes mentioned earlier. These methods utilize different sensor data and algorithms, each with its own characteristics and suitable for various application scenarios, together forming a diversified framework for the construction of topological maps in indoor scenes. Although topological maps have unique advantages among the three types of indoor maps introduced in Chapter 2, under the backdrop of the rapid development of artificial intelligence, a single topological map still falls short when facing

the demands of intelligent real-time interactive tasks. Since topological maps are constructed with a structure of nodes and edges, extended constraints can be added to them [38], using other environmental mapping information as auxiliary to provide multi-level map information for indoor robot localization and navigation.

## 4. Multi-level Map Representation Methods

The "multi-level" refers to the use of multiple sensors to perceive and map the environment using various mathematical expression methods to obtain multi-dimensional information, with the aim of better serving future intelligent interactive applications for indoor robots [11]. As shown in Fig. 2, this chapter will focus on topological maps, combined with other information of the environment, to study indoor multi-level map expression methods.



**Figure 2:** Adding additional information to the topological structure

### 4.1. Metrical-Topological Methods

The metrical-topological methods involve combining metric mapping with topological mapping to construct multi-layered topological maps. It integrates the basic geometric information of space on the basis of topological structures to provide a multi-dimensional representation of environmental structures. The Spatial Semantic Hierarchy (SSH) proposed by Kuipers [41] describes the knowledge of large-scale space using four dimensions: metric, topological, causal, and control, which is a meaningful pioneering attempt to describe the environment by integrating multi-dimensional information. Subsequent research [42] expanded the basic SSH, using metric mapping to create and store local perceptual maps of position neighborhoods as small sealed space ontologies, mapping them into the large-scale spatial ontology of cognitive maps, constructing a global topological relationship mapping, which enables robots to perform global topological inference and local motion planning effectively. In the metrical-topological methods, some researchers study how to extract topological maps from metric maps and represent the environment together with both [43][44][45], while others study how to enrich the topological structure by adding metric mapping information to topological maps. These attempts have a significant enlightening effect on the development of indoor topological map construction. In this process, many researchers have discovered the importance of semantic concepts in the positioning and navigation of indoor robots, making semantic topological methods a new research hotspot.

### 4.2. Semantic-Topological Methods

The semantic-topological methods combine semantic mapping with topological mapping to construct multi-layered topological maps. They integrate semantic features such as objects and locations on the basis of topological structures, helping robots understand their surroundings from the perspective of human spatial concepts. Some researchers combine semantic mapping with topological mapping in a layer-by-layer manner [15][52][53][54], but these approaches is merely "multi-layered" in a literal sense. It does not deeply integrate information across dimensions and is not lightweight enough, containing too much redundant information.

Another approaches are to integrate semantic information with the topological structure at the level of the structure, constructing multi-layered maps that are more suitable for positioning and navigation tasks indoor robots. For example, based on the SLAM algorithm, conference [32] proposed a semantic-topological method based on ORB-SLAM2, which uses the YOLOv5 network for object detection to obtain semantic features, and constructs a topological map based on the spatial position information of static objects. Based on the topological memory method SPTM [28], conference [47] proposed the Topological Semantic Graph Memory (TSGM), where image nodes represent different locations, and object nodes point to unique semantic objects using their visual representations. Object nodes within the neighborhood are connected to the corresponding image nodes as contextual auxiliary information according to visual rules, to eliminate the ambiguity of similar but different objects. There are also studies [11][48][49][50] based on modular methods, using cross-modal encoders to fuse topological maps with natural language instructions, effectively integrating semantic information to generate navigation plans, enabling robots to better achieve visual and language navigation tasks. The latest research [51] combines topological mapping with large language model, capturing the spatial structure and connectivity of the environment to build topological maps online and convert them into text prompts, and using visual models to convert the visual information of the scene into feature information rich in semantic content. These semantic-topological methods combine the global navigation advantages of topological structures with the rich semantic information of visual observations, helping robots understand their navigation paths in the environment, and enabling effective global exploration and robust and efficient navigation in complex environments.

The combination of topological maps with semantic information shows great application potential and development prospects in the field of indoor robots. Future research can further explore how to integrate rich semantic information more deeply into topological maps, not only including the recognition of objects and landmarks and other points of interest, but also involving the understanding of functional areas of the environment and divergent cognition of dynamic changes, enhancing generalization ability, enabling robots to adapt to various scales and complexities of indoor scenes. For example, developing advanced cross-modal learning algorithms and combining them with large language models to achieve effective integration of visual, language information, and topological structure, providing flexible and reliable support for human-computer interaction and intelligent strategy-making.

## 5. Conclusion

This review emphasizes the superiority of topological maps as an abstract representation method of environmental structure, in providing high-level navigation information, handling dynamic environments, reducing computation and storage and other dimensions. It points out their potential in positioning and navigation for indoor robots, compares and outlines some construction methods of indoor topological maps. On this basis, this review further proposes the concept of multi-level map expression, using the structural flexibility and scalability of topological maps, combined with other environmental information, to construct multi-level maps that enrich the robot's cognitive understanding of the environment, thereby better supporting indoor positioning and navigation of robots. In particular, semantic-topological methods aim to enhance the autonomy and intelligence of robots by combining advanced semantic understanding and topological representation technologies. The development of this field in the future is expected to be multifaceted, and with the continuous advancement in semantic-topological research, it is anticipated to bring breakthrough progress to indoor robotics technology.

## Acknowledgements

This work was financially supported by the National Natural Science Foundation of China under Grant No.62372049.

## References

- [1] J.J. Leonard, H. F. Durrant-Whyte, Mobile robot localization by tracking geometric beacons, *IEEE Transactions on Robotics and Automation* 7 (1991) 376-382. doi:10.1109/70.88147.
- [2] X. Meng, N. Ratliff, Y. Xiang, and D. Fox, Scaling local control to large-scale topological navigation, in: *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Paris, France, 2020, pp. 672-678. doi:10.1109/ICRA40945.2020.9196644.
- [3] P. Racinkis, J. Arents, M. Greitans, Constructing Maps for Autonomous Robotics: An Introductory Conceptual Overview, *Electronics* 12 (2023). doi:10.3390/electronics12132925.
- [4] T. Collins, J. J. Collins and D. Ryan, Occupancy grid mapping: An empirical evaluation, in: *Mediterranean Conference on Control & Automation*, Athens, Greece, 2007, pp. 1-6. doi:10.1109/MED.2007.4433772.
- [5] H. Oleynikova, Z. Taylor, et al., Voxblox: Incremental 3D Euclidean signed distance fields for on-board MAV planning, in: *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Vancouver, BC, Canada, 2017, pp. 1366-1373. doi:10.1109/IROS.2017.8202315.
- [6] V. Reijgwart, A. Millane, H. Oleynikova, R. Siegwart, C. Cadena, and J. Nieto, Voxgraph: Globally consistent, volumetric mapping using signed distance function submaps, *IEEE Robot. Autom. Lett.* 5 (2020) 227-234. doi:10.1109/LRA.2019.2953859.
- [7] C. Case, B. Suresh, A. Coates and A. Y. Ng, Autonomous sign reading for semantic mapping, in: *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Shanghai, China, 2011, pp. 3297-3303. doi:10.1109/ICRA.2011.5980523.
- [8] N. Sünderhauf, F. Dayoub, et al, Place categorization and semantic mapping on a mobile robot, in: *Proc. IEEE Int. Conf. Robot. Autom. (ICRA)*, Stockholm, Sweden, 2016, pp. 5729-5736. doi: 10.1109/ICRA.2016.7487796.
- [9] G. Narita, T. Seno, T. Ishikawa and Y. Kaji, PanopticFusion: Online Volumetric Semantic Mapping at the Level of Stuff and Things, in: *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, Macau, China, 2019, pp. 4205-4212. doi:10.1109/IROS40897.2019.8967890.
- [10] Qi, Xianyu et al., Building semantic grid maps for domestic robot navigation, *Int. J. Adv. Robot. Syst.* 17 (2020). doi:10.1177/1729881419900066.
- [11] S. Chen, P. -L. Guhur, M. Tapaswi, C. Schmid and I. Laptev, Think Global, Act Local: Dual-scale Graph Transformer for Vision-and-language Navigation, in: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, New Orleans, LA, USA, 2022, pp. 16516-16526. doi:10.1109/CVPR52688.2022.01604.
- [12] V. Setalaphruk, A. Ueno, I. Kume, Y. Kono and M. Kidode, Robot navigation in corridor environments using a sketch floor map, in: *Proceedings 2003 IEEE International Symposium on Computational Intelligence in Robotics and Automation*, Kobe, Japan, 2003, pp. 552-557. doi:10.1109/CIRA.2003.1222240.
- [13] S. Friedman, H. Pasula, D. Fox, Voronoi Random Fields: Extracting the Topological Structure of Indoor Environments via Place Labeling, in: *Proceedings of the 20th. International Joint Conference on Artificial Intelligence, IJCAI'07*, San Francisco, CA, USA, 2006, pp. 2109-2114. doi/10.5555/1625275.1625616.
- [14] P. Beeson, N. K. Jong and B. Kuipers, Towards Autonomous Topological Place Detection Using the Extended Voronoi Graph, in: *Proc. IEEE Int. Conf. Robot. Autom.*, Barcelona, Spain, 2005, pp. 4373-4379. doi: 10.1109/ROBOT.2005.1570793.
- [15] H. Oleynikova, Z. Taylor, R. Siegwart, J. Nieto, 3D Topological Graphs for Micro-Aerial Vehicle Planning, in: *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Madrid, Spain, 2018, pp. 1-9. doi:10.1109/IROS.2018.8594152.
- [16] U. Luxburg, A tutorial on spectral clustering, *Statistics and computing* 17 (2007) 395-416. doi:10.1007/s11222-007-9033-z.
- [17] C. Valgren, T. Duckett, A. Lilienthal, Incremental Spectral Clustering and Its Application to Topological Mapping, in: *Proc. IEEE Int. Conf. Robot. Autom.*, Rome, Italy, 2007, pp. 4283-4288. doi:10.1109/ROBOT.2007.364138.

- [18] E. Brunskill, T. Kollar and N. Roy, Topological mapping using spectral clustering and classification, in: 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems, San Diego, CA, USA, 2007, pp. 3491-3496. doi:10.1109/IROS.2007.4399611.
- [19] M. Liu, F. Colas and R. Siegwart, Regional topological segmentation based on mutual information graphs. in: Proc. IEEE Int. Conf. Robot. Autom., Shanghai, China, 2011, pp. 369-3274. doi:10.1109/ICRA.2011.5979672.
- [20] Jianbo Shi and J. Malik, Normalized cuts and image segmentation, in IEEE Transactions on Pattern Analysis and Machine Intelligence 22 (2000) pp. 888-905. doi: 10.1109/34.868688.
- [21] Z. Zivkovic, B. Bakker, B. Krose, Hierarchical map building using visual landmarks and geometric constraints, in: IEEE/RSJ International Conference on Intelligent Robots and Systems, Edmonton, AB, Canada, 2005, pp. 2480-2485. doi:10.1109/IROS.2005.1544951.
- [22] M. Liu, F. Colas, L. Oth, R. Siegwart, Incremental topological segmentation for semi-structured environments using discretized GVG, Autonomous Robots 38 (2015) 143-160. doi:10.1007/s10514-014-9398-8.
- [23] H. Cheng, H. Chen and Y. Liu, Topological Indoor Localization and Navigation for Autonomous Mobile Robot, IEEE Transactions on Automation Science and Engineering 12 (2015), pp. 729-738. doi:10.1109/TASE.2014.2351814.
- [24] J. Liu, J. Luo, J. Hou, D. Wen, G. Feng, X. Zhang, A BIM Based Hybrid 3D Indoor Map Model for Indoor Positioning and Navigation, ISPRS Int. J. Geo-Inf. 9 (2020). doi:10.3390/ijgi9120747
- [25] Z. Lin, C. Xiu, W. Yang, D. Yang, A Graph-Based Topological Maps Generation Method for Indoor Localization, in: Ubiquitous Positioning, Indoor Navigation and Location-Based Services (UPINLBS), Wuhan, China, 2018, pp. 1-8. doi:10.1109/UPINLBS.2018.8559830.
- [26] J. Zhu, Q. Li, R. Cao, K. Sun, T. Liu, J.M. Garibaldi et al., Indoor Topological Localization Using a Visual Landmark Sequence, Remote Sensing 11 (2019). doi:10.3390/rs11010073
- [27] E. Beeching, J. Dibangoye, O. Simonin, C. Wolf, Learning to plan with uncertain topological maps, in: Vedaldi, A., Bischof, H., Brox, T., Frahm, JM. (eds) Computer Vision-ECCV 2020. Volume 12348 of Lecture Notes in Computer Science, Springer, Cham, pp. 473-490. doi:10.1007/978-3-030-58580-8\_28.
- [28] N. Savinov, A. Dosovitskiy, V. Koltun, Semi-parametric Topological Memory for Navigation, in: International Conference on Learning Presentations, Vancouver, Canada, 2018.
- [29] K. Chen, J.P.d. Vicente, G. Sepulveda, F. Xia, A. Soto, M. Vázquez, and S. Savarese. A behavioral approach to visual navigation with graph localization networks, Robotics: Science and Systems 2 (2019).
- [30] Z. Huang, F. Liu, and H. Su, Mapping state space using landmarks for universal goal reaching, in: Proceedings of the 33rd. International Conference on Neural Information Processing Systems, Curran Associates Inc., Red Hook, NY, USA, 2019, pp. 1942-1952.
- [31] M. Laskin, S. Emmons, A. Jain, T. Kurutach, P. Abbeel, and D. Pathak, Sparse graphical memory for robust planning, Advances in Neural Information Processing Systems 33 (NeurIPS 2020) (2020).
- [32] Y. Wang, Y. Zhang, L. Hu, Wei. Wang, C. Ge, S. Tan, A Semantic Topology Graph to Detect Re-Localization and Loop Closure of the Visual Simultaneous Localization and Mapping System in a Dynamic Environment, Sensors 23 (2023). doi:10.3390/s23208445
- [33] K. Liu, T. Kurutach, C. Tung, P. Abbeel, A. Tamar, Hallucinative topological memory for ZeroShot visual planning, in: Proceedings of the 37th. International Conference on Machine Learning, ICML'20, 2020, pp. 6259-6270.
- [34] T. -H. Wang, H. -J. Huang, J. -T. Lin, C. -W. Hu, K. -H. Zeng and M. Sun, Omnidirectional CNN for visual place recognition and navigation, in: Proc. IEEE Int. Conf. Robot. Autom. (ICRA), Brisbane, QLD, Australia, 2018, pp. 2341-2348. doi:10.1109/ICRA.2018.8463173
- [35] A. Taniguchi, F. Sasaki and R. Yamashina, Pose Invariant Topological Memory for Visual Navigation, in: Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Montreal, QC, Canada, 2021, pp. 15364-15373. doi:10.1109/ICCV48922.2021.01510.



- [36] X. Meng, N. Ratliff, Y. Xiang and D. Fox, Scaling Local Control to Large-Scale Topological Navigation, in: Proc. IEEE Int. Conf. Robot. Autom. (ICRA), Paris, France, 2020, pp. 672-678. doi:10.1109/ICRA40945.2020.9196644.
- [37] R. R. Wiyatno, A. Xu and L. Paull, Lifelong Topological Visual Navigation, IEEE Robotics and Automation Letters 7 (2022) 9271-9278. doi:10.1109/LRA.2022.3189164.
- [38] F. Fraundorfer, C. Engels and D. Nister, Topological mapping, localization and navigation using image collections, in: IEEE/RSJ International Conference on Intelligent Robots and Systems, San Diego, CA, USA, 2007, pp. 3872-3877. doi:10.1109/IROS.2007.4399123.
- [39] J. Crespo, J. Carlos, O.M. Mozos, R. Barber, Semantic Information for Robot Navigation: A Survey, Applied Sciences 10 (2020). doi:10.3390/app10020497.
- [40] X. Han, S. Li, X. Wang, W. Zhou, Semantic Mapping for Mobile Robots in Indoor Scenes: A survey, Information 12 (2021). doi:10.3390/info12020092
- [41] B. J. Kuipers, The Spatial Semantic Hierarchy, Artificial Intelligence 119 (2000) 191-233. doi:10.1016/S0004-3702(00)00017-5.
- [42] B. Kuipers, J. Modayil, P. Beeson, M. MacMahon and F. Savelli, Local metrical and global topological maps in the hybrid spatial semantic hierarchy, in: Proc. IEEE Int. Conf. Robot. Autom. (ICRA), ICRA '04, New Orleans, LA, USA, 2004, pp. 4845-4851. doi:10.1109/ROBOT.2004.1302485.
- [43] B. Kaleci, O. Parlaktuna, U. Gürel, A comparative study for topological map construction methods from metric map, in: Proceedings of the 26th Signal Processing and Communications Applications Conference (SIU), Izmir, Turkey, 2018, pp. 1-4. doi:10.1109/SIU.2018.8404845.
- [44] F. Blochliger, M. Fehr, M. Dymczyk, T. Schneider and R. Siegwart, Topomap: Topological Mapping and Navigation Based on Visual SLAM Maps, in: Proc. IEEE Int. Conf. Robot. Autom. (ICRA), Brisbane, QLD, Australia, 2018, pp. 1-9. doi:10.1109/ICRA.2018.8460641.
- [45] F. Wang, Y. Liu, C. Wu, H. Chu, Topological Map Construction Based on Region Dynamic Growing and Map Representation Method, Applied Sciences 9 (2019). doi:10.3390/app9050816
- [46] K. He, G. Gkioxari, P. Dollár, and R. Girshick, Mask R-CNN, in: Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Venice, Italy, 2017, pp. 2961-2969. doi:10.1109/ICCV.2017.322.
- [47] N. Kim, O. Kwon, H. Yoo, Y. Choi, J. Park, S. Oh, Topological Semantic Graph Memory for Image-Goal Navigation, in: 6th. Annual Conference on Robot Learning, Auckland, New Zealand, 2022.
- [48] Kevin Chen, Junshen K. Chen, Jo Chuang, Marynel Vazquez, Silvio Savarese, Topological Planning with Transformers for Vision-and-Language Navigation, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021, pp. 11276-11286.
- [49] D. An, H. Wang, W. Wang, Z. Wang, Y. Huang, K. He, L. Wang, ETPNav: Evolving Topological Planning for Vision-Language Navigation in Continuous Environments, IEEE Transactions on Pattern Analysis and Machine Intelligence (2024). doi: arxiv-2304.03047.
- [50] M. Hwang, J. Jeong, M. Kim, Y. Oh, S. Oh, Meta-Explore Exploratory Hierarchical Vision-and-Language Navigation Using Scene Object Spectrum Grounding, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, Canada, 2023, pp. 6683-6693. doi:arXiv.2303.04077
- [51] J. Chen, B. Lin, R. Xu, Z. Chai, X. Liang, K. Wong, MapGPT: Map-Guided Prompting with Adaptive Path Planning for Vision-and-Language Navigation, in: Proceedings of the 62nd Annual Meeting of the Association for Computational Linguistics, Bangkok, Thailand, 2024.
- [52] D.An, Y. Qi, et al., BEVBert: Multimodal Map Pre-training for Language-guided Navigation, in: Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Paris, France, 2023. doi: arXiv.2212.04385.
- [53] A. Rosinol, A. Gupta, M. Abate, J. Shi, L. Carlone, 3D Dynamic Scene Graphs: Actionable Spatial Perception with Places, Objects, and Humans, Robotics: Science and Systems (RSS) (2020). doi: arXiv.2002.06289.
- [54] N. Hughes, Y. Chang, L. Carlone, Hydra: A Real-time Spatial Perception System for 3D Scene Graph Construction and Optimization, Robotics: Science and Systems (RSS) (2022).