Knowledge-Based Robot Manipulation and Knowledge-Graph Embedding

Subjects: Computer Science, Artificial Intelligence

Contributor: Runqing Miao , Qingxuan Jia , Fuchun Sun , Gang Chen , Haiming Huang , Shengyi Miao

Autonomous indoor service robots are affected by multiple factors when they are directly involved in manipulation tasks in daily life, such as scenes, objects, and actions. It is of self-evident importance to properly parse these factors and interpret intentions according to human cognition and semantics.

robot manipulation   knowledge graph   representation learning   graph neural network

1. Introduction

It is crucial to apply a strong understanding of relevant skills and task-planning capabilities to service robots, whether in daily life or in industrial-assembly scenarios. Classic task and motion planning (TAMP) [1] relies heavily, often entirely, on predefined planning domains, symbolic rules, and complex strategy searches. This leads to high costs and an inability to process tasks in a way that reflects human cognition and semantics. In general, all the limitations and constraints must be known before starting a task; otherwise, failures in efficient transfer and in the ability to adapt to changing task scenarios may occur

In relation to this problem, the past decade has witnessed rapid developments in the field of robot manipulation, especially in knowledge-based methods of representation and task planning. These enlightening paradigms enable robots to acquire manipulative task-related knowledge from human knowledge. However, knowledge is a higher-dimensional form of organization than data; discrete and structured characteristics also suffer from difficulties in directly describing continuous manipulation data. Therefore, most of the existing knowledge-based robot-manipulation-representation methods focus on static-object information and usually fail to elucidate reasonable decoupling between different factors. Concretely, the descriptions of tasks, actions, and skills are flat and chaotic. Furthermore, only rule-based symbolic calculations are considered during the processes of querying and reasoning.

Efficient and reasonable representations of complex manipulation knowledge require the consideration of both continuous and discrete data, as well as static and dynamic factors. Real-time responses and continuous interactive updates are also necessary for knowledge systems to handle new tasks. Additionally, the reasonable modeling of manipulation processes and the extraction of semantic information are prerequisites for precise inferences and planning.

2. Knowledge-Based Robot Manipulation

The application of service robots in manipulation tasks is usually divided into three levels: the planning level, which includes task analysis and task planning; the strategy level, which includes the motion planning of an action primitive; and the control level, which includes the execution process of the robot's hardware. Specifically, the knowledge-based method mainly focuses on the planning level, and fruitful results have been achieved with knowledge-graph-based robot-knowledge-representation systems. RoboEarth [2][3] was the first attempt to explore a knowledge-sharing system among robots, which mainly involved storing point clouds, CAD models, and object images in robot manipulation, which can enable robots to build semantic maps, which are required for daily tasks. On this basis, a robot can fully utilize cyberspace in its living space [4]. KnowRob [5] acquires an obtaining object and environmental information from the web and stores it in a knowledge-processing system that can be queried using Prolog, providing the knowledge required for service robots to perform daily manipulation tasks. KnowRob2 [6] improves the ability to acquire and reason knowledge based on KnowRob. RoboBrain [7] focuses on data collection. The knowledge engine stores different modalities of data, including symbols, natural language, tactile senses, robot trajectories, and visual features. Subsequently, they are connected to produce rich, heterogeneous graph representations. Perception and Manipulation Knowledge (PMK) [8] involves the formalization and implementation of a standardized ontology framework to extend robots' abilities related to manipulation tasks requiring task and motion planning (TAMP). Centering around scalability and responsiveness, TRPO [9] designs corresponding task-planning algorithms, starting from three categories of ontology knowledge: task, environment, and robot. Furthermore, AKB-48 [10] constructs a large-scale knowledge graph of articulated objects, which includes 2037 3D articulated object models from 48 categories in the real world. RoboKG [11] constructs a knowledge base specifically for performing grasping-based manipulation tasks, assisting the robot in the prediction of factors related to grasping, such as which component of the object to grasp, which end effector to use, and how much force to apply during grasping.

These knowledge-representation systems each have at their core a large-scale static-knowledge graph, which takes static objects in robot manipulation as the main description objects, and the description of manipulation behavior remains at the level of the simple recording of high-level tasks or low-level motion parameters. In addition, the architectures of these knowledge-representation systems are either excessively singular and flat, or overly chaotic and disorderly, resulting in high query complexity.

3. Knowledge-Graph Embedding

The knowledge graph originated in semantic networks [12] and is a typical form of structured knowledge representation. It represents structures of facts, consisting of entities, relations, and semantic descriptions [13]. Entities can be real-world objects and abstract concepts; relations represent the relations between entities. The semantic descriptions of entities and their relations contain categories and properties with clear meanings. Mature knowledge-graph schemes include the language-knowledge base WordNet [14], the world-knowledge base Freebase [15], Wikidata [16], DBpedia [17], and ConceptNet [18][19].

The goal of knowledge-graph embedding is to represent the semantic knowledge of the research object as a dense, low-dimensional, real-valued vector using machine learning. The low-dimensional vector representation

obtained by embedding is a distributed representation [20]. The embedding of knowledge graphs is focused on entities and relations in the knowledge base, in contrast to mapping, which considers spatial, temporal, and logical dimensions in the Internet of Things [21]. By mapping entities or relations into a low-dimensional vector space, the semantic information can be represented, and the complex semantic associations between entities and relations can be efficiently calculated. This is critical for both knowledge updates and reasoning.

The current mainstream knowledge-graph-embedding method is the translation model. TransE [22] regards relations as translation operations from the head entity to the tail entity in low-dimensional vector space. TransH [23] solves the problem of TransE's inability to handle 1–n, n–1, and n–n relations well. TransR [24] replaces the projection vector in TransH with a projection matrix.

Graph neural networks mainly solve the problem of non-Euclidean space, which is naturally suitable for the graph structures of knowledge graphs. Nevertheless, classical graph convolutional networks [25] are isomorphic graphs, in which all the edges in the network share the same weight and cannot adapt to different types of relation in knowledge graphs. Therefore, a heterogeneous graph-embedding method [26] is proposed to calculate the weights of different relations separately to deal with multi-relation graph data. In the researchers' work, the knowledge base that describes robot manipulation tasks has a hierarchical structure that differs from those of other knowledge bases.

References

1. Kaelbling, L.P.; Lozano-Pérez, T. Hierarchical task and motion planning in the now. In Proceedings of the 2011 IEEE ICRA, Shanghai, China, 9–13 May 2011.

2. Waibel, M.; Beetz, M.; Civera, J.; d'Andrea, R.; Elfring, J.; Galvez-Lopez, D.; Häussermann, K.; Janssen, R.; Montiel, J.; Perzylo, A.J.I.R.; et al. Roboearth. IEEE Robot. Autom. Mag. 2011, 18, 69–82.

3. Riazuelo, L.; Tenorth, M.; Di Marco, D.; Salas, M.; Gálvez-López, D.; Mösenlechner, L.; Kunze, L.; Beetz, M.; Tardós, J.D.; Montano, L.; et al. RoboEarth semantic mapping: A cloud enabled knowledge-based approach. IEEE Trans. Autom. Sci. Eng. 2015, 12, 432–443.

4. Cai, X.; Ning, H.; Dhelim, S.; Zhou, R.; Zhang, T.; Xu, Y.; Wan, Y.J.D.C. Robot and its living space: A roadmap for robot development based on the view of living space. Digit. Commun. Netw. 2021, 7, 505–517.

5. Tenorth, M.; Beetz, M. KnowRob: A knowledge processing infrastructure for cognition-enabled robots. Int. J. Robot. Res. 2013, 32, 566–590.

6. Beetz, M.; Beßler, D.; Haidu, A.; Pomarlan, M.; Bozcuoğlu, A.K.; Bartels, G. Know rob 2.0—A 2nd generation knowledge processing framework for cognition-enabled robotic agents. In Proceedings

of the 2018 IEEE International Conference on Robotics and Automation (ICRA), Brisbane, Australia, 21–25 May 2018; pp. 512–519.

7. Saxena, A.; Jain, A.; Sener, O.; Jami, A.; Misra, D.K.; Koppula, H.S. Robobrain: Large-scale knowledge engine for robots. arXiv 2014, arXiv:1412.0691.

8. Diab, M.; Akbari, A.; Ud Din, M.; Rosell, J.J.S. PMK—A knowledge processing framework for autonomous robotics perception and manipulation. Sensors 2019, 19, 1166.

9. Sun, X.; Zhang, Y.; Chen, J.J.E. RTPO: A domain knowledge base for robot task planning. Electronics 2019, 8, 1105.

10. Liu, L.; Xu, W.; Fu, H.; Qian, S.; Han, Y.; Lu, C.J. AKB-48: A Real-World Articulated Object Knowledge Base. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022.

11. Kwak, J.H.; Lee, J.; Whang, J.J.; Jo, S.J.I.R.; Letters, A. Semantic Grasping Via a Knowledge Graph of Robotic Manipulation: A Graph Representation Learning Approach. IEEE Robot. Autom. Lett. 2022, 7, 9397–9404.

12. Sowa, J.F.J.E. Semantic networks. Encycl. Artif. Intell. 1992, 2, 1493–1511.

13. Ji, S.; Pan, S.; Cambria, E.; Marttinen, P.; Yu, P.S. A survey on knowledge graphs: Representation, acquisition, and applications. IEEE Trans. Neural Netw. Learn. Syst. 2021, 33, 494–514.

14. Miller, G.A. WordNet: A lexical database for English. Commun. ACM 1995, 38, 39–41.

15. Bollacker, K.; Evans, C.; Paritosh, P.; Sturge, T.; Taylor, J. Freebase: A collaboratively created graph database for structuring human knowledge. In Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data, Vancouver, BC, Canada, 9–12 June 2008; pp. 1247–1250.

16. Vrandečić, D.; Krötzsch, M. Wikidata: A free collaborative knowledgebase. Commun. ACM 2014, 57, 78–85.

17. Auer, S.; Bizer, C.; Kobilarov, G.; Lehmann, J.; Cyganiak, R.; Ives, Z. Dbpedia: A nucleus for a web of open data. In Proceedings of the Semantic Web: 6th International Semantic Web Conference, 2nd Asian Semantic Web Conference, ISWC 2007 + ASWC 2007, Busan, Republic of Korea, 11–15 November 2007; pp. 722–735.

18. Liu, H.; Singh, P. ConceptNet—A practical commonsense reasoning tool-kit. BT Technol. J. 2004, 22, 211–226.

19. Speer, R.; Chin, J.; Havasi, C. Conceptnet 5.5: An open multilingual graph of general knowledge. In Proceedings of the Thirty-first AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.

20. Cai, H.; Zheng, V.W.; Chang, K.C.-C. A comprehensive survey of graph embedding: Problems, techniques, and applications. IEEE Trans. Knowl. Data Eng. 2018, 30, 1616–1637.

21. Dhelim, S.; Ning, H.; Zhu, T. STLF: Spatial-temporal-logical knowledge representation and object mapping framework. In Proceedings of the 2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC), Budapest, Hungary, 9–12 October 2016; pp. 1550–1554.

22. Bordes, A.; Usunier, N.; Garcia-Duran, A.; Weston, J.; Yakhnenko, O. Translating embeddings for modeling multi-relational data. Adv. Neural Inf. Process. Syst. 2013, 2, 2787–2795.

23. Wang, Z.; Zhang, J.; Feng, J.; Chen, Z. Knowledge graph embedding by translating on hyperplanes. In Proceedings of the AAAI conference on artificial intelligence, Washington, DC, USA, 7–14 February 2023.

24. Lin, Y.; Liu, Z.; Sun, M.; Liu, Y.; Zhu, X. Learning entity and relation embeddings for knowledge graph completion. In Proceedings of the AAAI conference on artificial intelligence, Washington, DC, USA, 7–14 February 2023.

25. Kipf, T.N.; Welling, M. Semi-supervised classification with graph convolutional networks. arXiv 2016, arXiv:1609.02907.

26. Schlichtkrull, M.; Kipf, T.N.; Bloem, P.; Van Den Berg, R.; Titov, I.; Welling, M. Modeling relational data with graph convolutional networks. In Proceedings of the Semantic Web: 15th International Conference, ESWC 2018, Heraklion, Crete, Greece, 3–7 June 2018; pp. 593–607.