



OPEN ACCESS

EDITED AND REVIEWED BY
Giuseppe Boccignone,
University of Milan, Italy

*CORRESPONDENCE
Konstantinos A. Tsintotas,
✉ ktsintot@pme.duth.gr

RECEIVED 01 December 2023
ACCEPTED 14 February 2024
PUBLISHED 01 March 2024

CITATION
Tsintotas KA, Kansizoglou I, Pastra K,
Aloimonos Y, Gasteratos A, Sirakoulis GC and
Sandini G (2024), Editorial: Enhanced human
modeling in robotics for socially-aware place
navigation.
Front. Robot. AI 11:1348022.
doi: 10.3389/frobt.2024.1348022

COPYRIGHT
© 2024 Tsintotas, Kansizoglou, Pastra,
Aloimonos, Gasteratos, Sirakoulis and Sandini.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Editorial: Enhanced human modeling in robotics for socially-aware place navigation

Konstantinos A. Tsintotas^{1*}, Ioannis Kansizoglou¹,
Katerina Pastra², Yiannis Aloimonos³, Antonios Gasteratos¹,
Giorgios Ch. Sirakoulis¹ and Giulio Sandini⁴

¹Democritus University of Thrace, Komotini, Greece, ²Athena Research Center, Marousi, Greece,
³University of Maryland, Baltimore, MD, United States, ⁴Italian Institute of Technology (IIT), Genova,
Liguria, Italy

KEYWORDS

robotics, social navigation, AI, machine learning, language processing

Editorial on the Research Topic

[Enhanced human modeling in robotics for socially-aware place navigation](#)

1 Introduction

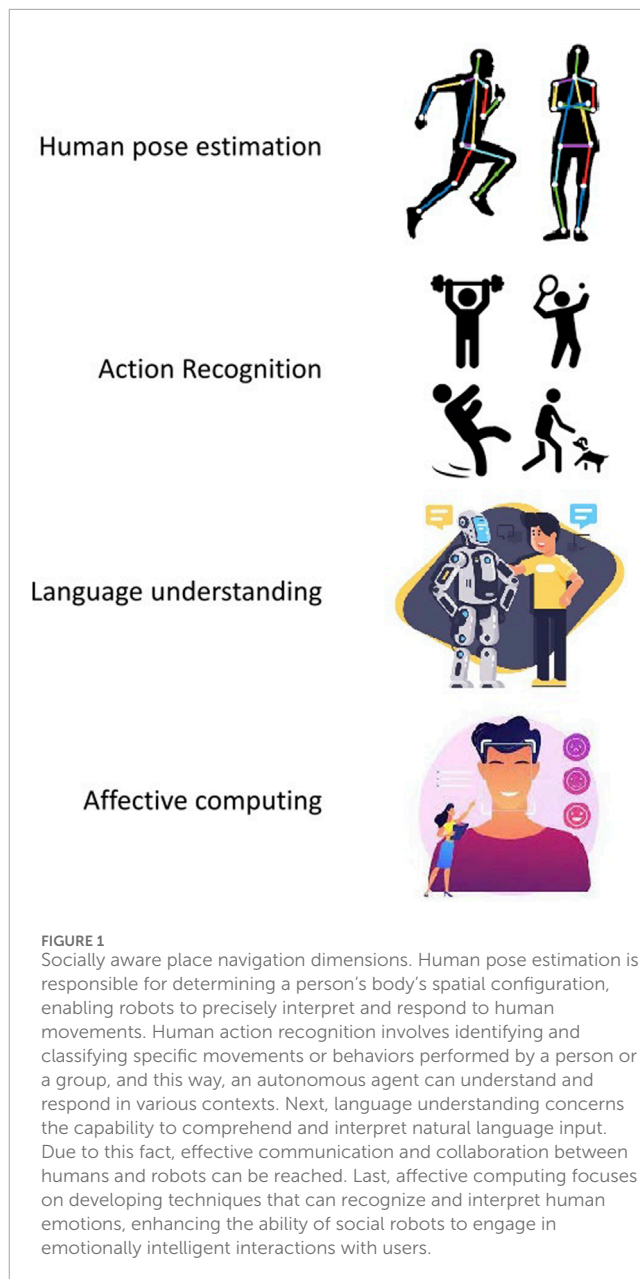
Autonomous and accurate navigation is a prerequisite for any intelligent system assigned to various missions. Yet, this task presents a higher complexity when a mobile robot navigates in an unfamiliar terrain, as it needs to move through the environment and construct a detailed map of its surroundings. At the same time, the system should estimate its pose and orientation during the incremental construction of its internal map (Tsintotas et al., 2022). This process is widely known as simultaneous localization and mapping (SLAM) and is paramount for effective and context-aware navigation. However, this challenge becomes even more intricate when robots work within human environments, as human-robot coexistence introduces variables such as human activities, intentions, and their impacts on the robot's path (Keroglou et al., 2023). At the same time, the integration necessitates adherence to stringent safety and security requirements. Consequently, the robotic community tries to tackle these challenges through several techniques that collectively shape the field into a demanding, interdisciplinary pursuit known as socially aware navigation. This involves technical considerations and a deep understanding of the social dynamics between humans and robots, marking a crucial intersection of robotics, artificial intelligence, and human-computer interaction. Should we understand human activities, intentions, or social dynamics via intelligent pipelines, robots can navigate spaces shared with humans, fostering a harmonious coexistence, e.g., healthcare, or assistive technologies to smart homes and public space. Last, socially aware robot navigation aims to bridge the gap between artificial intelligence and human interaction, paving the way for a more integrated and socially intelligent future.

2 Analysis of the Research Topic

The paradigm of socially aware place navigation is situated within the intricate domain of human modeling, systematically examining various dimensions such as human pose estimation (Wei et al., 2022), action recognition (Charalampous et al., 2017; Dessalene et al., 2021), language understanding (Vatakis and Pastra, 2016), and affective computing (Kansizoglou et al., 2022) (see Figure 1). The first is the discernment of the spatial configuration of an individual's body, a pivotal facet enabling a robotic system to comprehend humans' physical presence and movements within its proximate environment (An et al., 2022). At the same time, action recognition further augments this comprehension by interpreting the activities in which individuals are engaged (Dessalene et al., 2023), thereby contributing to a nuanced understanding of the contextual environment (Moutsis et al., 2023). Language understanding, a fundamental component of this multifaceted paradigm, empowers the robot to discern verbal cues and commands (Pastra and Aloimonos, 2012), thereby facilitating seamless communication with human counterparts. At the same time, affective computing introduces an emotional dimension, endowing the robot to discern and appropriately respond to human emotions, enhancing its adaptability to intricate social contexts (Kansizoglou et al., 2019). Last, the amalgamation of these human-centric capacities within the purview of the navigation task epitomizes a sophisticated methodological approach, and consequently, such frameworks are poised to excel in scenarios characterized by adversity, dynamism, and heightened interactivity.

2.1 Contributing articles

Although user-centered approaches are essential to create a comfortable and safe human-robot interaction, they are still rare in industrial settings. Aiming to close this research gap, in Bernotat et al., two user studies with large heterogeneous samples were conducted. In particular, in User Study 1, the participants' ideas about robot memory were explored, as well as what aspects of the robot's movements were found positive, and what they would change. The effects of participants' demographic backgrounds and attitudes were controlled for. Next, it is self-evident that even in such an elementary and minimal environment compared to the real world, home agents require guidance from dense reward functions to learn to carry out complex tasks. As task decomposition is an easy-to-use approach for introducing those dense rewards, in Petsanis et al., a method that can be used to improve training in embodied AI environments by harnessing the task decomposition capabilities of TextWord is presented. On the other hand, Karasoulas et al. examined how to detect the presence or absence of individuals indoors by analyzing the ambient air's CO₂ concentration using simple Markov Chain Models. While this study focused on employing 1-h window testing sets, there exists significant potential for accurately assessing occupancy profiles within shorter minute intervals. At last, the authors in Arapis et al. focus on localizing humans in the world and predicting the free space around them by incorporating other static and dynamic obstacles. Their research



is based on a multitasking learning strategy to handle both tasks, achieving this goal with minimal computational demands when employed in difficult industrial environments, such as human instances at a close distance or the limits of the field of view of the capturing sensor.

3 Discussion and conclusion

Overall, the main objective of a human-aware navigation pipeline is to facilitate human-robot coexistence in a shared environment. Such a scenario requires the efficient parallel realization of each member's goals without needless external interceptions or delays and the successful completion of specific

everyday tasks. On top of that, the robotic agent is expected to inspire a sense of trust and friendliness in humans, mainly realized when the agent operates concisely, adaptively, transparently, and naturally. Thus, robot navigation techniques shall employ enhanced human understanding and modeling techniques, capturing those features that mainly affect the efficiency of the task. As a result, it becomes increasingly vital to develop robust, lightweight action and affect estimation solutions based on robotics sensory data and capacities, like active vision (Aloimonos et al., 1988). Finally, computational efficiency and real-time operation capacities always limit the introduced solutions.

Author contributions

KT: Conceptualization, Writing–original draft, Writing–review and editing. IK: Conceptualization, Writing–original draft, Writing–review and editing. KP: Supervision, Writing–review and editing. YA: Supervision, Writing–review and editing. AG: Supervision, Writing–review and editing. GoS: Supervision, Writing–review and editing. GuS: Supervision, Writing–review and editing.

References

- Aloimonos, J., Weiss, I., and Bandyopadhyay, A. (1988). Active vision. *Act. Vis. Int. J. Comput. Vis.* 1, 333–356. doi:10.1007/bf00133571
- An, S., Zhang, X., Wei, D., Zhu, H., Yang, J., and Tsintotas, K. A. (2022). Fasthand: fast monocular hand pose estimation on embedded systems. *J. Syst. Archit.* 122, 102361. doi:10.1016/j.sysarc.2021.102361
- Charalampous, K., Kostavelis, I., and Gasteratos, A. (2017). Recent trends in social aware robot navigation: a survey. *Robotics Aut. Syst.* 93, 85–104. doi:10.1016/j.robot.2017.03.002
- Dessalene, E., Devaraj, C., Maynard, M., Fermuller, C., and Aloimonos, Y. (2021). “Forecasting action through contact representations from first person video,” in IEEE Transactions on Pattern Analysis and Machine Intelligence, 1 June 2023 (IEEE), 6703–6714. doi:10.1109/TPAMI.2021.3055233
- Dessalene, E., Maynard, M., Fermüller, C., and Aloimonos, Y. (2023). “Therbligs in action: video understanding through motion primitives,” in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada, 17–24 June 2023, 10618–10626. doi:10.1109/CVPR52729.2023.01023
- Kansizoglou, I., Bampis, L., and Gasteratos, A. (2019). An active learning paradigm for online audio-visual emotion recognition. *IEEE Trans. Affect. Comput.* 13, 756–768. doi:10.1109/taffc.2019.2961089
- Kansizoglou, I., Misirlis, E., Tsintotas, K., and Gasteratos, A. (2022). Continuous emotion recognition for long-term behavior modeling through recurrent neural networks. *Technologies* 10, 59. doi:10.3390/technologies10030059
- Keroglou, C., Kansizoglou, I., Michailidis, P., Oikonomou, K. M., Papapetros, I. T., Dragkola, P., et al. (2023). A survey on technical challenges of assistive robotics for elder people in domestic environments: the aspidra concept. *IEEE Trans. Med. Robotics Bionics* 5, 196–205. doi:10.1109/tmrb.2023.3261342
- Moutsis, S. N., Tsintotas, K. A., Kansizoglou, I., Shan, A., Aloimonos, Y., and Gasteratos, A. (2023). “Fall detection paradigm for embedded devices based on yolov8,” in IEEE International Conference on Imaging Systems and Techniques (IST), Copenhagen, Denmark, 17–19 Oct. 2023, 1–6. doi:10.1109/IST59124.2023.10355696
- Pastra, K., and Aloimonos, Y. (2012). The minimalist grammar of action. *Philosophical Trans. R. Soc. B Biol. Sci.* 367, 103–117. doi:10.1098/rstb.2011.0123
- Tsintotas, K. A., Bampis, L., and Gasteratos, A. (2022). *Online appearance-based place recognition and mapping: their role in autonomous navigation*, 133. Springer Nature.
- Vatakis, A., and Pastra, K. (2016). A multimodal dataset of spontaneous speech and movement production on object affordances. *Sci. Data* 3, 150078–150086. doi:10.1038/sdata.2015.78
- Wei, D., An, S., Zhang, X., Tian, J., Tsintotas, K. A., Gasteratos, A., et al. (2022). “Dual regression for efficient hand pose estimation,” in 2022 International Conference on Robotics and Automation (ICRA), Philadelphia, PA, USA, 23–27 May 2022 (IEEE), 6423–6429. doi:10.1109/ICRA46639.2022.9812217

Funding

The author(s) declare that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.