Autonomous Robots (2023) 47:503–504
https://doi.org/10.1007/s10514-023-10119-3

EDITORIAL

Editorial

M. Ani Hsieh[1] · Dylan A. Shell[2]

Published online: 23 June 2023

Welcome! This editorial serves as an introduction for the special issue of Autonomous Robots (AURO) featuring papers from the 17th Robotics: Science and Systems (RSS) conference. In the midst of the pandemic, with international travel infeasible and in-person events still imprudent, the conference was held virtually from July 12–16, 2021. RSS'21 employed a combination of on-line platforms and tools to enable both asynchronous and synchronous activities, including: pre-timed release of pre-recorded talks, live video streaming, and interactive meetings. There were also engaging workshop and tutorial sessions that, in keeping with the tradition of the event, were very well attended.

The conference facilitated presentation and discussion of results at the very vanguard of robotics research—results from across many of the sub-areas of the field. Of the total 282 submissions, a double-blind review process (carried out resolutely by 26 Area Chairs) led to a selection, finally, of 92 papers. From those papers, with input from the Area Chairs, a subset were identified on the basis of their importance and expected significance to the field. The authors of those papers were invited to submit manuscripts that improved and expanded upon the original work appearing within the conference proceedings. The submissions then underwent a further stage of peer review and, in most cases, revision (in some cases, multiple revisions). The seven articles included in this collection are the result; we hope that readers will appreciate the more mature presentation and evolution of the work over the form that originally appeared at the conference.

This collection contains papers which offer advances across a wide span of robotics topics, ranging from perception and mapping, through aspects of co-design and multi-agency, to manipulation. Machine Learning plays *some* role in each article, but the variety of forms it takes is intriguing in its own way.

✉ Dylan A. Shell
dshell@tamu.edu

1 University of Pennsylvania, Philadelphia, PA, USA

2 Texas A&M University, College Station, TX, USA

The article by Liu, Bansal, Daruna, and Chernova concerns the problem of predicting properties of objects based on partial observations. Their work advocates use of a richer knowledge representation than commonly employed: specifically they pursue a relational model which captures high-arity or -degree interrelationships between various objects' semantic properties. Using this model, they introduce two transformer-based neural networks which learn such interrelationships and are shown to be effective at predicting unobserved properties when presented with novel objects; the specific utility of their advance is demonstrated in scenarios involving a mobile manipulator. The authors also contribute a rich situated object dataset that includes a large set of items, in diverse situations, with multimodal and difficult-to-observe object properties. Closely related, and adopting an active perception viewpoint more expressly, is the article entitled "Multimodal Embodied Attribute Learning by Robots for Object-Centric Action Policies." The work addresses the problem of efficiently generating a policy of exploratory behaviors to allow robots to identify attributes of objects. It shows how, even when some of the underlying actions may fail, multimodal interactions may be employed to accurately perceive objects' properties—and the demonstration in a tabletop scenario (combining visual data, grasping, shaking, and other acts) is especially compelling.

The theme of drawing upon data from different sensing modalities is apparent also in the paper by Tang, De Martini, and Newman. Their work presents a method that employs overhead imagery (such as publicly available satellite images) to improve ground LIDAR–based localization. The three authors show how to learn a transformation to "emulate" a range sensor, using overhead images to produce point-clouds as if generated by a LIDAR scan. In this setting, their paper is the first to learn place recognition *and* metric localization from overhead imagery. If internet services can provide useful data (like Google's satellite imagery), what about computing power? The contribution entitled "Co-Design of Communication and Machine Inference for Cloud Robotics" considers the question of how to produce compact representations for remote robotic data processing (processing, typically, of sensed data). The idea is that if information were to be adequately compressed, then the computational burden could be shared between a local robot and some remote server. The article introduces an encoder-decoder–based architecture to learn reduced representations tailored for particular uses through a loss function which balances task accuracy against reconstructability. Their evaluation case-studies show that the method is effective at distilling relevant information because task performance can be retained under substantial compression rates.

Under the broad banner of challenges arising from perception data when two or more parties are involved, the work presented by Gao, Zhu, and Zhang tackles the problem of finding correspondences between objects observed by multiple agents, each within their own field of view. Their formulation solves the correspondences problem by identifying matches based upon graph neural network–based features, but also quantifies uncertainty (they operate within a Bayesian learning framework). The authors add two extra penalty terms within the loss function to explicitly account for non-covisibility and uncertainty in each identified correspondence. The effectiveness of their approach is compellingly demonstrated in collaborative furniture assembly, collaborative object localization, and autonomous driving scenarios.

Another place where insight manifests as the canny design of loss functions is the contribution titled "An Empowerment-based Solution to Robotic Manipulation Tasks with Sparse Rewards" which tackles the acute problem of slow learning rates in Reinforcement Learning (RL) with sparse rewards. To do this, the authors formulate a new, modified empowerment-based approach that is particularly well suited to robotic manipulation tasks as it specifically rewards interactions with the objects being manipulated in the workspace. The technique consistently improves learning rates and has the advantage of being straightforward to integrate with algorithms for a variety of RL settings beyond standard RL (including skill discovery and curriculum learning).

One further paper considers manipulation: the contribution with the memorable acronym *DiSECt* details the first differentiable simulator for cutting soft materials. In addition to speed improvements that the differentiablity affords, the work has direct implications for the task of cutting with a robotic manipulator: the slicing motion identified reduces the needed cutting force by a substantial amount—on the order of 40% compared to a vertical chop. (Your humble guest editors have tried to resist the urge to state that the article represents cutting-edge cutting edge research.)

To conclude, we wish to thank the reviewers who devoted time to providing thorough and constructive reviews for the manuscripts; we would also like to thank Gaurav Sukhatme for permitting us to put this special issue together. We do hope that these articles will inspire fresh innovations, ignite creative thoughts that lead to subsequent developments, and spur follow-up work.

Yours,
the guest editors:
M. Ani Hsieh
Dylan A. Shell