Qeios

Peer Review

Review of: "What's the Move? Hybrid Imitation Learning via Salient Points"

Ryota Takamido[1]

1. The University of Tokyo, Tokyo, Japan

This study proposes a novel imitation learning (IL) method called SPHINX for learning and generalizing policies in complex, long-horizon robotic tasks that also require precise object manipulation. In the SPHINX process, the method infers "salient points," which represent task-relevant parts within the input 3D point cloud information (e.g., the handle of a cup), and utilizes them to estimate waypoints for switching between gross long-range movements and precise manipulation. The training data were collected using a customized teleoperation system that allows human demonstrators to specify salient points and waypoints with simple clicks and drags. The evaluation was conducted on four different complex robotic tasks in the physical world, and the advantages of SPHINX over existing IL algorithms were clearly demonstrated.

Overall, the proposed method introduces many innovative and noteworthy contributions that enhance the performance of existing imitation learning algorithms. Therefore, I have given this paper a high rating. More specifically, if my understanding of the paper is correct, its novelties and advantages are as follows:

(1) The salient points-based mode-switching approach introduced in this study overcomes the limitations of existing IL methods.
(2) The developed teleoperation system for demonstration data collection is well designed and has the potential for use in various contexts.
(3) A rigorous comparison with state-of-the-art methods has been conducted in the experiments.
(4) SPHINX demonstrated significantly high generalizability with small datasets, utilizing only a few dozen demonstrations to solve complex and highly varied benchmark tests.

On the other hand, I have several concerns about this study, as outlined below. These are not critical

 doi.org/10.32388/VNR6IZ

issues, but I hope the authors may find them useful for improving their research.

(1) The paper may lack discussion from a practical perspective. For example, what kinds of real-world robotic applications do the authors envision (e.g., picking, assembling, etc.), and what further improvements would be needed to achieve them?

(2) Related to the above, the study could be further strengthened by including evaluations based on tasks that resemble practical applications, in addition to general benchmark scenarios.

(3) The experimental environment appears relatively simple, with a smaller number of obstacles and objects than some practical situations. If the number of obstacles and objects were increased to create a more “cluttered” environment, would the proposed method still perform well?

(4) Readers may wonder how to determine an appropriate value for the hyperparameter r. Providing more insight into this aspect would enhance the paper’s clarity.

Declarations

Potential competing interests: No potential competing interests to declare.