A Survey of Mirror-neuron Mechanism and Robot Tool Use

Xizhi Xiao
Yuanpei College
Peking University
xxz3470608107@stu.pku.edu.cn

Abstract

The tool-use ability has traditionally been regarded as an indicator of intelligence, separating humans from animals. However, recent research has discovered that animals are able to manipulate a tool or even create a tool from raw materials, questioning the previously prevailing belief. In this essay, we firstly introduce a neurophysiological mechanism—-the mirror-neuron mechanism—-which is believed to play a fundamental role in human action understanding and imatation. Then, we review several computational frameworks for robot tool affordance understanding and tool manipulation. We conclude this essay by proposing several directions for future research.

1 The mirror-neuron mechanism

Mirror neurons are a special type of visuomotor neurons, originally discovered in the area F5 of the monkey premotor cortex, that discharge not only when the monkey does a particular action, but also when it observes other individuals doing a similar action [2] [3]. Rizzolatti and Luppino [7] discovered that mirror neurons respond when the monkey sees object-directed action. Mirror neurons require an interaction between a biological effector (hand or mouth) and an object in order to be triggered by the visual stimuli. Further research reveal that presenting widely different visual stimuli, but which all represent the same action, is equally effective. Morover, it doesn't matter for neuron activation whether the observed action in done near or far from the monkey, or whether the action eventually rewarded. Kohler and his colleges [4] explored mirror neurons activity while the monkey was observing a noisy action, and found out that about 15% of mirror neurons responsive to presentation of actions accompanied by sounds responded to the presentation of the sound alone, too. Results of these experiments not only revealed basic properties of mirror neurons in monkey's brain, but also implied a relationship of mirror neurons and action perceiving and understanding.

Further research discovered the relationship between the visual and motor properties of the mirror neurons. All mirror neurons show congruence between the visual actions they respond to and the motor responses they code. According to the extent of congruence, mirror neurons are divided into "strictly congruent" and "broadly congruent" neurons [3]. "Broadly congruent" neurons can be triggered when the observed actions are not exactly congruent with the action encoded in the brain. However, "strictly congruent" neurons respond only when the observed actions are strictly congruent with the actions encoded in the brain. Despite their differences, the two types of mirror neurons imply the process of matching the observed actions with the corresponding actions encoded in the brain, or prior knowledge. Umilta et al. [9] tested whether the mental representation of an action triggers their activity. They hypothesized that if mirror neurons are involved in action understanding, they should discharge also in conditions in which monkey does not see the occurring action but has sufficient clues to create a mental representation of what the experimenter does. Based on the assumption, they designed a controlled experiment and discovered that even when the monkeys didn't not observe the final, critical action directly, they could infer it from the previous actions, therefore their mirror neurons were activated just as those who observe the actions fully. Both of the evidence above shows that mirror neurons are connected with action understanding.

2 Imitation learning and tool use

In this section, we discuss how the mirror-neuron mechanism enables imitation learning in humans and animals, and how it relates to the ability of tools using. We review computational models and robotics research of tool affordance understanding, tool manipulation and tool creation.

Qin et al. [6] gets inspiration from Shumaker et al. [8]'s definition and further clarifies the definition of robot tool use as the use of externally manufactured, unanimated tools by a robot to manipulate objects or the environment, through either dynamic or static interactions, in order to achieve a variety of goals. They divides the skills required for tool use into three categories: perception, manipulation and high-level cognition skills. Perception skills refer to the ability to identify and localize tools and manipulanda from the environment, which is also required in general robot manipulation tasks. Manipulation skills focus on how to realize the required kinematics and dynamics of tool use actions, which can be subdivided into the contact poses (e.g. grasping) and the course of the action. High-level cognition skills includes reasoning and planning tool use actions given the tasks and available tools. These definitions and classifications serve as a foundation for building an intelligent system for tool use.

Humans understand the functions of a tool not just by its appearance, but also by its physical attributes and the predicted physical dynamic applying it to a task. That explains why people utilize diverse objects as the bottle opener, including a wooden board, teeth, or even another wine bottle. Zhu et al. [10] propose task oriented modeling, learning and recognition which aims at understanding the underlying functions, physics and causality in using objects as "tools". Given a specific task, they represent each object in a generative spatiotemporal representation consisting of an affordance basis to be grasped by hand, a functional basis to act on a target object, the imagined actions with typical motion trajectories, and the underlying physical concepts, e.g. force, pressure, etc. From this new perspective, any objects can be viewed as the tool for the task, and object recognition is not merely memorizing typical appearance examples for each category but reasoning the physical mechanisms in various tasks to achieve generalization. Brawer et al. [1] introduce a method for a robot to learn an explicit model of cause-and-effect relations by constructing a structural causal model through a mix of observation and self-supervised experimentation, allowing a robot to reason from causes to effects and from effects to causes. Their model achieves good performance tool affordance learning tasks. Results suggest that after minimal training examples, the system can preferentially choose new tools based on the context, and can use these tools for goal-directed object manipulation. Myers et al. [5] assume that the geometry of a part is closely related to its possible functions, or its affordances. Therefore, they propose two approaches for learning affordances from local shape and geometry primitives: superpixel based hierarchical matching pursuit (S-HMP) and structured random forests (SRF). Moreover, since a part can be used in many ways, they introduce a large RGB-Depth dataset where tool parts are labeled with multiple affordances and their relative rankings.

3 Conclusion and outlook

In conclusion, the discovery of mirror neurons and their role in action understanding and imitation provides a neurological basis for how humans and animals can imitate and learn to use tools. Computational models that represent tools in terms of their affordances and physical properties show promise for enabling robots to understand tool functions and effects. Future research could benefits from the following directions:

1. **New tool creation**. Enabling robots to create novel tools by combining and modifying existing objects based on task goals and environmental constraints. This could lead to more flexible and creative tool use.
2. **Transfer in human demonstration and robot learning**. Improving the transfer of tool skills and knowledge from human demonstrations and previous experience to new tool-use tasks and environments. This could accelerate robot learning and generalizability.

References

[1] Jake Brawer, Meiying Qin, and Brian Scassellati. A causal approach to tool affordance learning. In *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages

8394–8399. IEEE, 2020. 2

[2] Giuseppe Di Pellegrino, Luciano Fadiga, Leonardo Fogassi, Vittorio Gallese, and Giacomo Rizzolatti. Understanding motor events: a neurophysiological study. *Experimental brain research*, 91:176–180, 1992. 1

[3] Vittorio Gallese, Luciano Fadiga, Leonardo Fogassi, and Giacomo Rizzolatti. Action recognition in the premotor cortex. *Brain*, 119(2):593–609, 1996. 1

[4] Evelyne Kohler, Christian Keysers, M Alessandra Umilta, Leonardo Fogassi, Vittorio Gallese, and Giacomo Rizzolatti. Hearing sounds, understanding actions: action representation in mirror neurons. *Science*, 297(5582):846–848, 2002. 1

[5] Austin Myers, Ching L Teo, Cornelia Fermüller, and Yiannis Aloimonos. Affordance detection of tool parts from geometric features. In *2015 IEEE International Conference on Robotics and Automation (ICRA)*, pages 1374–1381. IEEE, 2015. 2

[6] Meiying Qin, Jake Brawer, and Brian Scassellati. Robot tool use: A survey. *Frontiers in Robotics and AI*, 9:1009488, 2023. 2

[7] Giacomo Rizzolatti and Giuseppe Luppino. The cortical motor system. *Neuron*, 31(6):889–901, 2001. 1

[8] Robert W Shumaker, Kristina R Walkup, and Benjamin B Beck. *Animal tool behavior: the use and manufacture of tools by animals*. JHU Press, 2011. 2

[9] Maria Alessandra Umilta, Evelyne Kohler, Vittorio Gallese, Leonardo Fogassi, Luciano Fadiga, Christian Keysers, and Giacomo Rizzolatti. I know what you are doing: A neurophysiological study. *Neuron*, 31(1):155–165, 2001. 1

[10] Yixin Zhu, Yibiao Zhao, and Song Chun Zhu. Understanding tools: Task-oriented object modeling, learning and recognition. In *Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015. 2