

Computer vision researchers use motion to discover objects in videos

July 27 2023, by Byron Spice



Credit: Carnegie Mellon University

Researchers at Carnegie Mellon University's Robotics Institute have shown that computer vision systems can more easily detect objects in motion—like a car driving down the street or a person walking in a crosswalk—than stationary objects.

Martial Hebert, dean of CMU's School of Computer Science and a professor in the Robotics Institute, and robotics Ph.D. student Zhipeng

Bao collaborated on the project with Toyota Research Institute, which sponsored the work. The research could help computers and robots better automatically detect objects in videos.

Object recognition is fundamental to understanding real-world scenes, so developing motion-guided methods for discovering objects could improve autonomous driving. It could also prove useful for retail robotics, robotic manipulation and robots in the home.

Working with colleagues from Toyota, the University of California, Berkeley, and the University of Illinois Urbana-Champaign, the CMU researchers developed a framework called MoTok that enables the computer to identify features of things it sees moving on its own. MoTok then uses these features to reconstruct the object, allowing the computer to discover the object in a way that enables it to find that same object again.

The researchers have since extended the work so the computer can depict these features in a simplified, virtualized fashion. This development enables the computer to better identify high-level features, making it possible for the computer to categorize objects rather than just identifying a particular object. The paper is currently available on the *arXiv* preprint server.

Visualizing objects comes naturally to people—so naturally, in fact, that vision is hard to introspect.

"We have no awareness of how we do it," Hebert said.

Machine learning advances have helped improve computers' ability to recognize objects, albeit in a way much different than humans. Those methods, however, require tens of thousands of hours of video containing labeled objects. It is laborious, expensive and prone to

failures outside the lab.

"Obviously, that doesn't scale," Hebert said.

What is needed is a generalized method that enables computer programs to discover objects in videos on their own, without the need for labels or supervision. As MoTok demonstrates, using motion to guide object discovery is one way of achieving this goal.

"Objects that move are easy to differentiate from static backgrounds," said Bao, who completed the research while interning at Toyota Research Institute. "Movement also can help define an object that has multiple moving parts. A car door might open and close and wheels might spin, but all the parts moving together as the car travels down a street can help [computer](#) programs better understand the concept of a car."

The team presented its paper on MoTok in June at the [Conference on Vision and Pattern Recognition](#). More information about MoTok is available on the [project's website](#).

More information: Zhipeng Bao et al, Discovering Objects that Can Move, *arXiv* (2022). [DOI: 10.48550/arxiv.2203.10159](https://doi.org/10.48550/arxiv.2203.10159)

Provided by Carnegie Mellon University

Citation: Computer vision researchers use motion to discover objects in videos (2023, July 27) retrieved 27 April 2024 from <https://techxplore.com/news/2023-07-vision-motion-videos.html>

This document is subject to copyright. Apart from any fair dealing for the purpose of private study or research, no part may be reproduced without the written permission. The content is

provided for information purposes only.