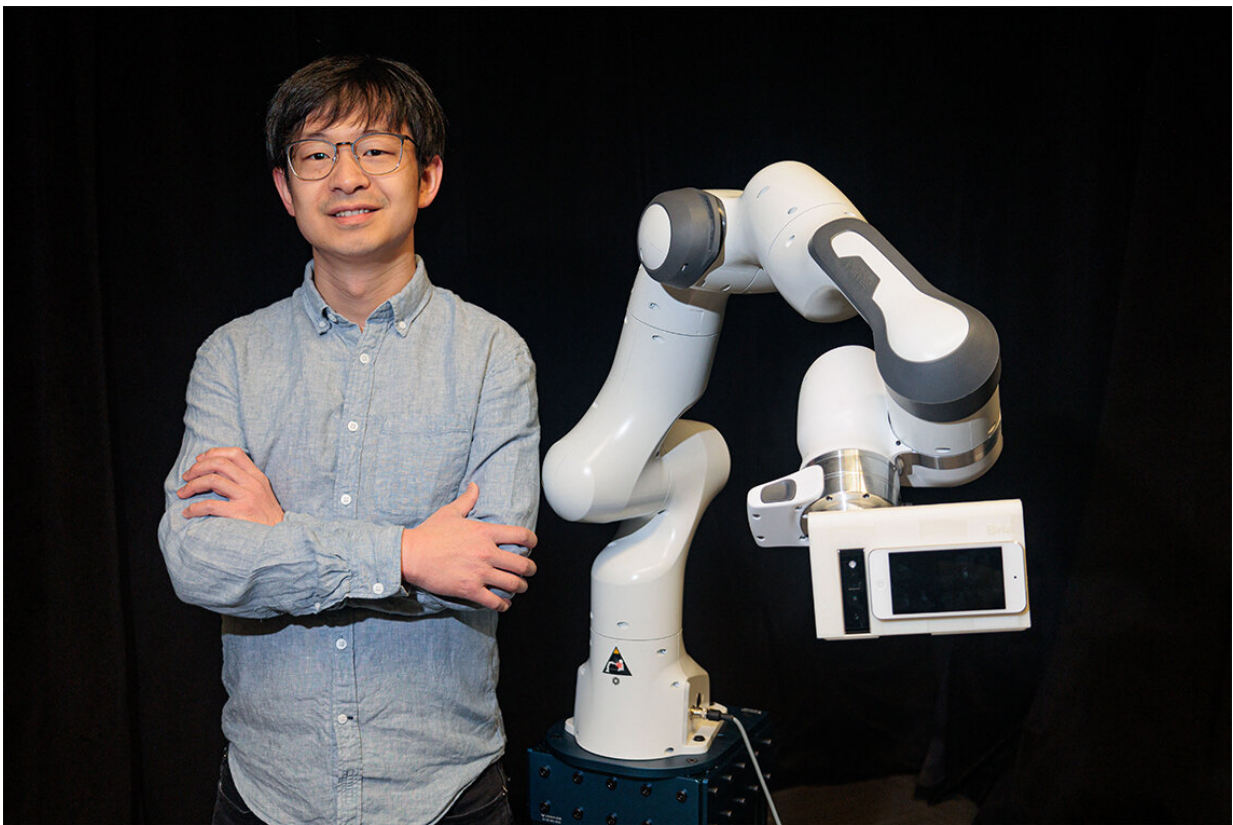


Researchers develop interactive 'Stargazer' camera robot that can help film tutorial videos

May 22 2023, by Krystle Hewitt



Research led by U of T computer science PhD candidate Jiannan Li explores how an interactive camera robot can assist instructors and others in making how-to videos. Credit: Matt Hints

A group of computer scientists from the University of Toronto wants to make it easier to film how-to videos.

The team of researchers have developed Stargazer, an interactive [camera](#) robot that helps university instructors and other content creators create engaging tutorial videos demonstrating physical skills.

For those without access to a cameraperson, Stargazer can capture dynamic instructional videos and address the constraints of working with static cameras.

"The robot is there to help humans, but not to replace humans," explains lead researcher Jiannan Li, a Ph.D. candidate in U of T's department of computer science in the Faculty of Arts & Science. "The instructors are here to teach. The robot's role is to help with filming—the heavy-lifting work."

The Stargazer work is outlined in a paper published in *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. The international conference on [human-computer interaction](#) was held in Hamburg, Germany, April 23–28.

Li's co-authors include fellow members of U of T's Dynamic Graphics Project (dgp) lab: postdoctoral researcher Mauricio Sousa, Ph.D. students Karthik Mahadevan and Bryan Wang, Professor Ravin Balakrishnan and Associate Professor Tovi Grossman; as well as Associate Professor Anthony Tang (cross-appointed with the Faculty of Information); recent U of T Faculty of Information graduates Paula Akemi Aoyai and Nicole Yu; and third-year computer engineering student Angela Yang.

Stargazer uses a single camera on a robot arm, with seven independent motors that can move along with the video subject by autonomously

tracking regions of interest. The system's camera behaviors can be adjusted based on subtle cues from instructors, such as body movements, gestures and speech that are detected by the prototype's sensors.

The instructor's voice is recorded with a wireless microphone and sent to Microsoft Azure Speech-to-Text, a speech-recognition software. The transcribed text, along with a custom prompt, is then sent to the GPT-3 program, a large language model which labels the instructor's intention for the camera—such as a standard versus high angle and normal versus tighter framing.

These camera control commands are cues naturally used by instructors to guide the attention of their audience and are not disruptive to instruction delivery, the researchers say.

For example, the instructor can have Stargazer adjust its view to look at each of the tools they will be using during a tutorial by pointing to each one, prompting the camera to pan around. The instructor can also say to viewers, "If you look at how I put 'A' into 'B' from the top," Stargazer will respond by framing the action with a high angle to give the audience a better view.

In designing the interaction vocabulary, the team wanted to identify signals that are subtle and avoid the need for the instructor to communicate separately to the robot while speaking to their students or audience.

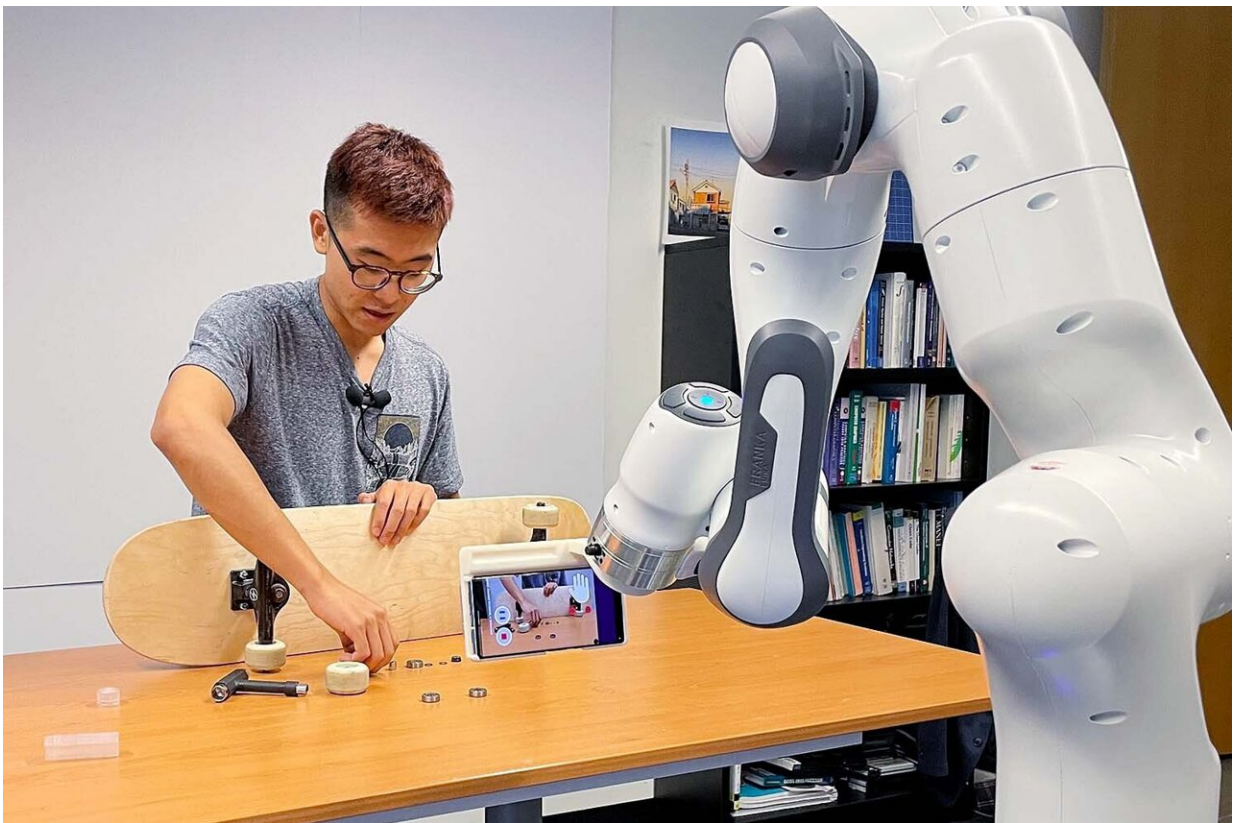
"The goal is to have the robot understand in real time what kind of shot the instructor wants," Li says. "The important part of this goal is that we want these vocabularies to be non-disruptive. It should feel like they fit into the tutorial."

Stargazer's abilities were put to the test in a study involving six

instructors, each teaching a distinct skill to create dynamic tutorial videos.

Using the robot, they were able to produce videos demonstrating physical tasks on a diverse range of subjects, from skateboard maintenance to interactive sculpture-making and setting up virtual-reality headsets, while relying on the robot for subject tracking, camera framing and camera angle combinations.

The participants were each given a practice session and completed their tutorials within two takes. The researchers reported all of the participants were able to create videos without needing any additional controls than what was provided by the robotic camera and were satisfied with the quality of the videos produced.



A study participant uses the interactive camera robot Stargazer to record a how-to video on skateboard maintenance. Credit: University of Toronto

While Stargazer's range of camera positions is sufficient for tabletop activities, the team is interested in exploring the potential of camera drones and robots on wheels to help with filming tasks in larger environments from a wider variety of angles.

They also found some study participants attempted to trigger object shots by giving or showing objects to the camera, which were not among the cues that Stargazer currently recognizes. Future research could investigate methods to detect diverse and subtle intents by combining simultaneous signals from an instructor's gaze, posture and speech, which Li says is a long-term goal the team is making progress on.

While the team presents Stargazer as an option for those who do not have access to professional film crews, the researchers admit the robotic camera prototype relies on an expensive robot arm and a suite of external sensors. Li notes, however, that the Stargazer concept is not necessarily limited by costly technology.

"I think there's a real market for robotic filming equipment, even at the consumer level. Stargazer is expanding that realm, but looking farther ahead with a bit more autonomy and a little bit more interaction. So realistically, it could be available to consumers," he says.

Li says the team is excited by the possibilities Stargazer presents for greater human-[robot](#) collaboration.

"For robots to work together with humans, the key is for robots to

understand humans better. Here, we are looking at these vocabularies, these typically human communication behaviors," he explains.

"We hope to inspire others to look at understanding how humans communicate ... and how robots can pick that up and have the proper reaction, like assistive behaviors."

More information: Jiannan Li et al, Stargazer: An Interactive Camera Robot for Capturing How-To Videos Based on Subtle Instructor Cues, *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems* (2023). [DOI: 10.1145/3544548.3580896](https://doi.org/10.1145/3544548.3580896)

Provided by University of Toronto

Citation: Researchers develop interactive 'Stargazer' camera robot that can help film tutorial videos (2023, May 22) retrieved 25 April 2024 from <https://techxplore.com/news/2023-05-interactive-stargazer-camera-robot-tutorial.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.