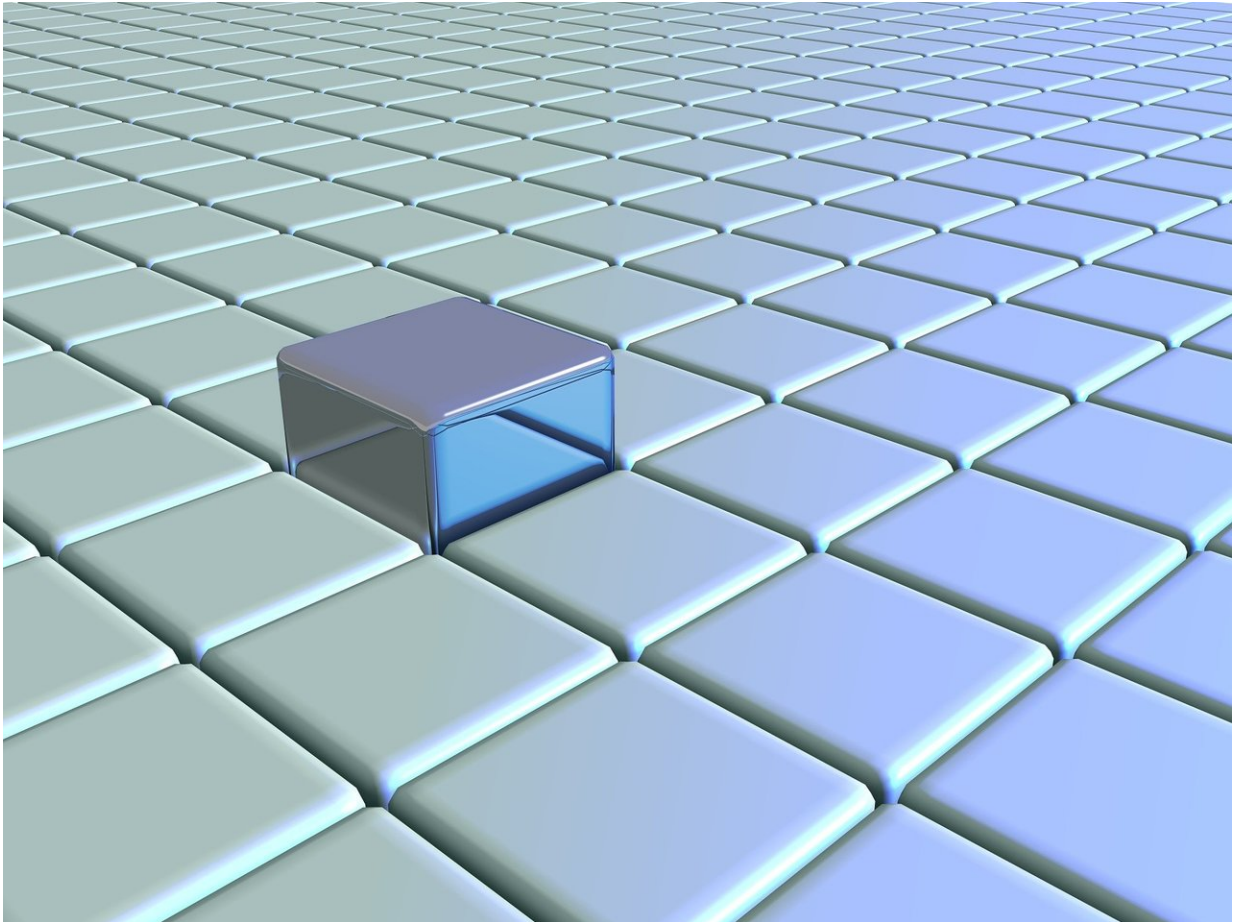


Helping computers to see 3-D structures

November 30 2018



Credit: CC0 Public Domain

If you can recognize structures around you while walking down a city street, you have your eyes to thank. Humans can automatically perceive 3-D structure in the world by identifying lines, shapes, symmetries and

the patterns and relationships between them in things like buildings, sidewalks and everyday objects. But can a computer be taught to do the same?

Zihan Zhou, assistant professor of information sciences and technology at Penn State, is setting out to explore that question thanks to a recent grant from the National Science Foundation.

"We want a [computer](#) to see 3-D space as humans do," said Zhou. "This particular award and project is about structure perception, which has been largely ignored in 3-D vision. This is something that has not been done before."

Structure perception is the ability of a human's eyes to organize data or patterns and group them in certain ways. For example, a human can look at a line drawing of a building and visualize doors, windows and walls.

"There are many types of these relationships in the [real world](#), and humans make use of those relationships to sense the 3-D space," he said. "Human eyes can easily perceive these kinds of things. The question now is: Can the computer have the ability to sense these things as a human does?"

To answer that question, Zhou plans to develop a new data-driven framework for structure discovery, leveraging the availability of massive visual data and recent advances in machine learning techniques.

These techniques could then be applied to a wide spectrum of real-world [computer vision](#) problems, including 3-D modeling of urban environments, virtual and augmented reality, and autonomous driving. The research could also impact cognitive sciences, by suggesting new computational mechanisms for image understanding; and [human-robot interaction](#), by enabling robots to reason in terms of geometric shape,

physics and dynamics.

"If a robot recognizes something as a specific type of structure, then it knows how to interact with it," said Zhou. "For example, if a robot is able to recognize a [structure](#) with a flat top, it would know that it could put an object like a cup on it."

Additionally, the framework may impact the work of architects, designers and engineers.

"If you think of those architects, they are working with 3-D models every day," said Zhou. "If they build something, they first create line drawings. So if a computer can understand doors and windows in the drawings, it would be very useful for architectural design and engineering."

Zhou developed an interest in this topic while a graduate intern at Adobe. In his internship, he studied the relationship between camera motion and the environment, which could help the movie industry to analyze scenes.

"I tried to extract some kinds of structures from the videos and the sequence of the camera," he said. "At that point it was to analyze camera trajectory for the movie industry, but later we realized it was more systematic."

Now, at Penn State, Zhou hopes to leverage the interdisciplinary network to advance his work.

"IST has people working in diverse areas, and many of them can be impacted by this kind of work," he said. "This has generated a lot of interest in different areas. We are looking to extend this beyond and to find applications to make this more collaborative."

"About 70 percent of information we obtain is from visual cues from our eyes," he concluded. "Obviously we have areas like natural language processing to help understand speaking and sounds, but human vision is the dominating factor in how we understand this world. To make the computer see the world as we do is one of the most exciting areas in artificial intelligence and computer science."

Provided by Pennsylvania State University

Citation: Helping computers to see 3-D structures (2018, November 30) retrieved 26 June 2024 from <https://phys.org/news/2018-11-d.html>

<p>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.</p>
--