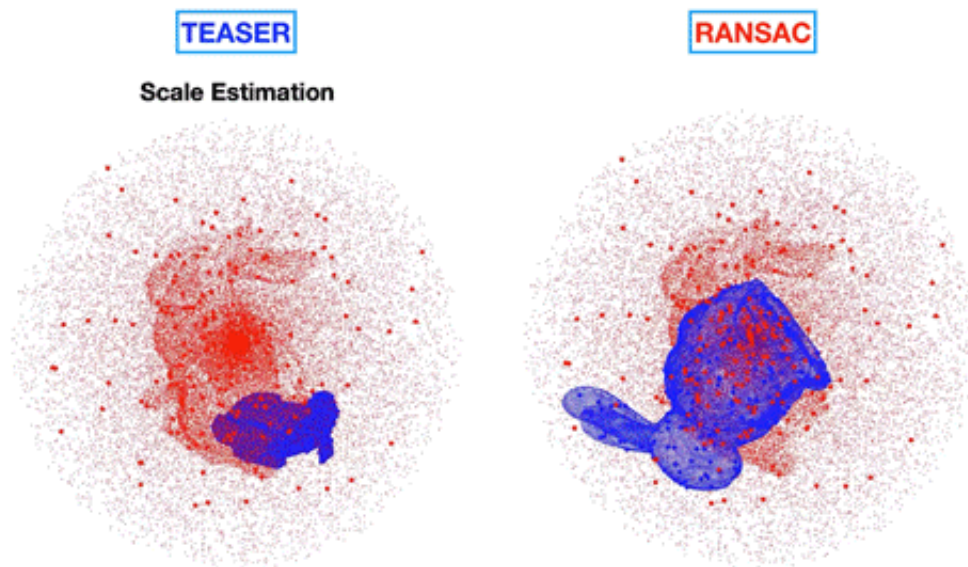# Algorithm quickly finds hidden objects in dense point clouds

June 20 2019, by Jennifer Chu

The MIT-developed technique quickly and smoothly matches objects to those hidden in dense point clouds (left), versus existing techniques (right) that produce incorrect, disjointed matches. Credit: Courtesy of the researchers

A new MIT-developed technique enables robots to quickly identify objects hidden in a three-dimensional cloud of data, reminiscent of how some people can make sense of a densely patterned "Magic Eye" image if they observe it in just the right way.

Robots typically "see" their environment through sensors that collect and translate a visual scene into a matrix of dots. Think of the world of, well,

"The Matrix," except that the 1s and 0s seen by the fictional character Neo are replaced by dots—lots of dots—whose patterns and densities outline the objects in a particular scene.

Conventional techniques that try to pick out objects from such clouds of dots, or point clouds, can do so with either speed or accuracy, but not both.

With their new technique, the researchers say a robot can accurately pick out an object, such as a small animal, that is otherwise obscured within a dense cloud of dots, within seconds of receiving the visual data. The team says the technique can be used to improve a host of situations in which machine perception must be both speedy and accurate, including driverless cars and robotic assistants in the factory and the home.

"The surprising thing about this work is, if I ask you to find a bunny in this cloud of thousands of points, there's no way you could do that," says Luca Carlone, assistant professor of aeronautics and astronautics and a member of MIT's Laboratory for Information and Decision Systems (LIDS). "But our algorithm is able to see the object through all this clutter. So we're getting to a level of superhuman performance in localizing objects."

Carlone and graduate student Heng Yang will present details of the technique later this month at the Robotics: Science and Systems conference in Germany.

"Failing without knowing"

Robots currently attempt to identify objects in a point cloud by comparing a template object—a 3-D dot representation of an object, such as a rabbit—with a point cloud representation of the real world that may contain that object. The template image includes "features," or

collections of dots that indicate characteristic curvatures or angles of that object, such the bunny's ear or tail. Existing algorithms first extract similar features from the real-life point cloud, then attempt to match those features and the template's features, and ultimately rotate and align the features to the template to determine if the point cloud contains the object in question.

But the point cloud data that streams into a robot's sensor invariably includes errors, in the form of dots that are in the wrong position or incorrectly spaced, which can significantly confuse the process of feature extraction and matching. As a consequence, robots can make a huge number of wrong associations, or what researchers call "outliers" between point clouds, and ultimately misidentify objects or miss them entirely.

Carlone says state-of-the-art algorithms are able to sift the bad associations from the good once features have been matched, but they do so in "exponential time," meaning that even a cluster of processing-heavy computers, sifting through dense point cloud data with existing algorithms, would not be able to solve the problem in a reasonable time. Such techniques, while accurate, are impractical for analyzing larger, real-life datasets containing dense point clouds.

Other algorithms that can quickly identify features and associations do so hastily, creating a huge number of outliers or misdetections in the process, without being aware of these errors.

"That's terrible if this is running on a self-driving car, or any safety-critical application," Carlone says. "Failing without knowing you're failing is the worst thing an algorithm can do."
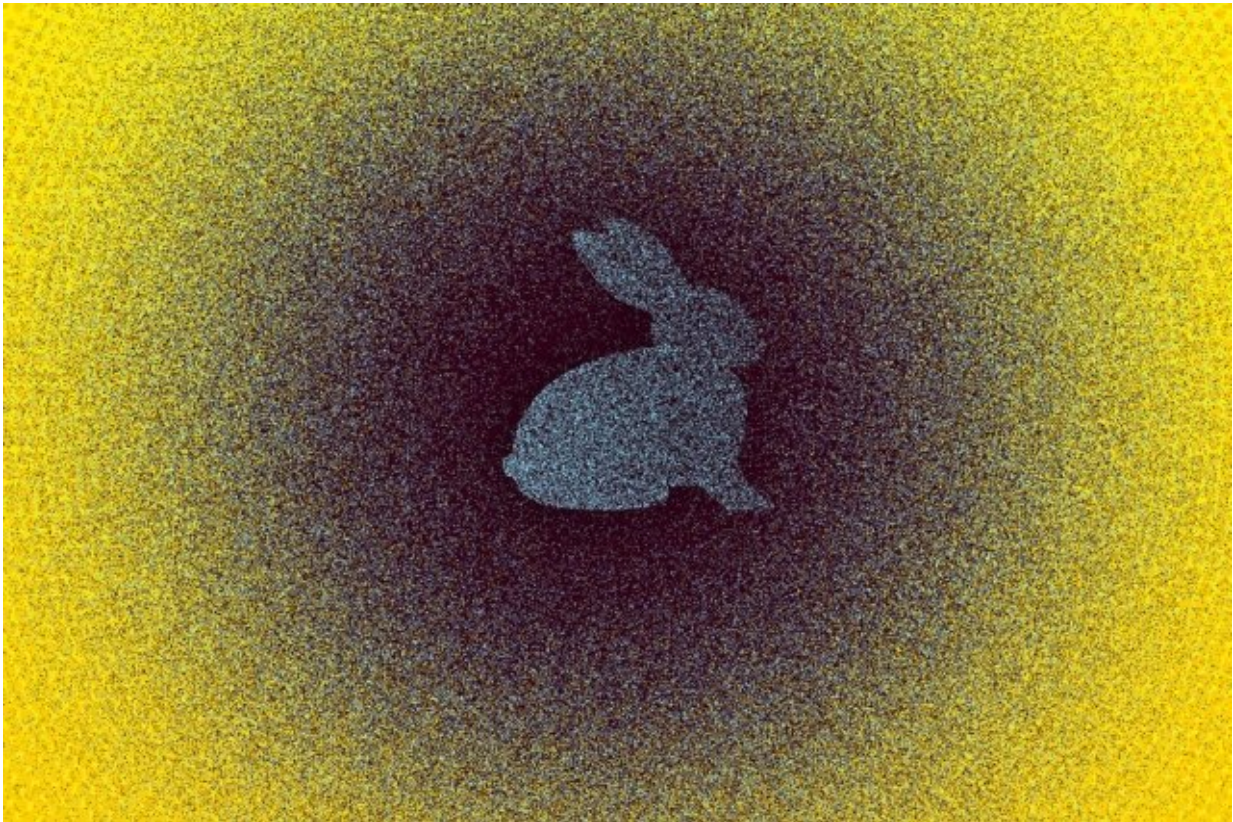
Robots currently attempt to identify objects in a point cloud by comparing a template object — a 3-D dot representation of an object, such as a rabbit — with a point cloud representation of the real world that may contain that object. Credit: Christine Daniloff, MIT

A relaxed view

Yang and Carlone instead devised a technique that prunes away outliers in "polynomial time," meaning that it can do so quickly, even for increasingly dense clouds of dots. The technique can thus quickly and accurately identify objects hidden in cluttered scenes.

The researchers first used conventional techniques to extract features of a template object from a point cloud. They then developed a three-step

process to match the size, position, and orientation of the object in a point cloud with the template object, while simultaneously identifying good from bad feature associations.

The team developed an "adaptive voting scheme" algorithm to prune outliers and match an object's size and position. For size, the algorithm makes associations between template and point cloud features, then compares the relative distance between features in a template and corresponding features in the point cloud. If, say, the distance between two features in the point cloud is five times that of the corresponding points in the template, the algorithm assigns a "vote" to the hypothesis that the object is five times larger than the template object.

The algorithm does this for every feature association. Then, the algorithm selects those associations that fall under the size hypothesis with the most votes, and identifies those as the correct associations, while pruning away the others. In this way, the technique simultaneously reveals the correct associations and the relative size of the object represented by those associations. The same process is used to determine the object's position.

The researchers developed a separate algorithm for rotation, which finds the orientation of the template object in three-dimensional space.

To do this is an incredibly tricky computational task. Imagine holding a mug and trying to tilt it just so, to match a blurry image of something that might be that same mug. There are any number of angles you could tilt that mug, and each of those angles has a certain likelihood of matching the blurry image.

Existing techniques handle this problem by considering each possible tilt or rotation of the object as a "cost"—the lower the cost, the more likely that that rotation creates an accurate match between features. Each

rotation and associated cost is represented in a topographic map of sorts, made up of multiple hills and valleys, with lower elevations associated with lower cost.

But Carlone says this can easily confuse an algorithm, especially if there are multiple valleys and no discernible lowest point representing the true, exact match between a particular rotation of an object and the object in a point cloud. Instead, the team developed a "convex relaxation" algorithm that simplifies the topographic map, with one single valley representing the optimal rotation. In this way, the algorithm is able to quickly identify the rotation that defines the orientation of the object in the point cloud.

With their approach, the team was able to quickly and accurately identify three different objects—a bunny, a dragon, and a Buddha—hidden in point clouds of increasing density. They were also able to identify objects in real-life scenes, including a living room, in which the algorithm quickly was able to spot a cereal box and a baseball hat.

Carlone says that because the approach is able to work in "polynomial time," it can be easily scaled up to analyze even denser point clouds, resembling the complexity of sensor data for driverless cars, for example.

"Navigation, collaborative manufacturing, domestic robots, search and rescue, and self-driving cars is where we hope to make an impact," Carlone says.

  More information: A Polynomial-time Solution for Robust Registration with Extreme Outlier Rates: rss2019.informatik.uni-freibur … e/papers/0013_FI.pdf