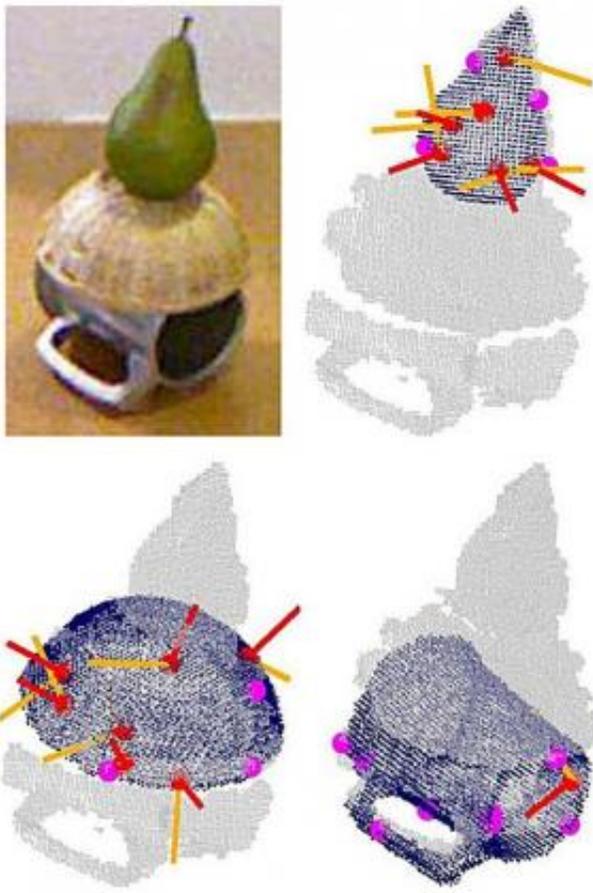


# Better robot vision: Neglected statistical tool could help robots better understand the objects in the world around them

October 7 2013, by Larry Hardesty

---



Object recognition is one of the most widely studied problems in

computer vision. But a robot that manipulates objects in the world needs to do more than just recognize them; it also needs to understand their orientation. Is that mug right-side up or upside-down? And which direction is its handle facing?

To improve robots' ability to gauge [object](#) orientation, Jared Glover, a graduate student in MIT's Department of Electrical Engineering and Computer Science, is exploiting a statistical construct called the Bingham distribution. In a [paper](#) they're presenting in November at the International Conference on Intelligent Robots and Systems, Glover and MIT alumna Sanja Popovic '12, MEng '13, who is now at Google, describes a new robot-vision [algorithm](#), based on the Bingham distribution, that is 15 percent better than its best competitor at identifying familiar objects in cluttered scenes.

That algorithm, however, is for analyzing high-quality visual data in familiar settings. Because the Bingham distribution is a tool for reasoning probabilistically, it promises even greater advantages in contexts where information is patchy or unreliable. In [ongoing work](#), Glover is using Bingham distributions to analyze the orientation of pingpong balls in flight, as part of a broader project to teach robots to play pingpong. In cases where visual information is particularly poor, his algorithm offers an improvement of more than 50 percent over the best alternatives.

"Alignment is key to many problems in robotics, from object-detection and tracking to mapping," Glover says. "And ambiguity is really the central challenge to getting good alignments in highly cluttered scenes, like inside a refrigerator or in a drawer. That's why the Bingham distribution seems to be a useful tool, because it allows the algorithm to get more information out of each ambiguous, local feature."

Because Bingham distributions are so central to his work, Glover has

also developed a suite of software tools that greatly speed up calculations involving them. The software is [freely available online](#), for other researchers to use.

## In the rotation

One reason the Bingham distribution is so useful for robot vision is that it provides a way to combine information from different sources.

Generally, determining an object's orientation entails trying to superimpose a geometric [model](#) of the object over visual data captured by a camera—in the case of Glover's work, a Microsoft Kinect camera, which captures a 2-D color image together with information about the distance of the color patches.

For simplicity's sake, imagine that the object is a tetrahedron, and the geometric model consists of four points marking the tetrahedron's four corners. Imagine, too, that software has identified four locations in an image where color or depth values change abruptly—likely to be the corners of an object. Is it a tetrahedron?

The problem, then, boils down to taking two sets of points—the model and the object—and determining whether one can be superimposed on the other. Most algorithms, Glover's included, will take a first stab at aligning the points. In the case of the tetrahedron, assume that, after that provisional alignment, every point in the model is near a point in the object, but not perfectly coincident with it.

If both sets of points in fact describe the same object, then they can be aligned by rotating one of them around the right axis. For any given pair of points—one from the model and one from the object—it's possible to calculate the probability that rotating one point by a particular angle around a particular axis will align it with the other. The problem is that the same rotation might move other pairs of points farther away from

each other.

Glover was able to show, however, that the rotation probabilities for any given pair of points can be described as a Bingham distribution, which means that they can be combined into a single, cumulative Bingham distribution. That allows Glover and Popovic's algorithm to explore possible rotations in a principled way, quickly converging on the one that provides the best fit between points.

## **Big umbrella**

Moreover, in the same way that the Bingham distribution can combine the probabilities for each pair of points into a single probability, it can also incorporate probabilities from other sources of information—such as estimates of the curvature of objects' surfaces. The current version of Glover and Popovic's algorithm integrates point-rotation probabilities with several other such [probabilities](#).

In experiments involving visual data about particularly cluttered scenes—depicting the kinds of environments in which a household robot would operate—Glover's algorithm had about the same false-positive rate as the best existing algorithm: About 84 percent of its object identifications were correct, versus 83 percent for the competition. But it was able to identify a significantly higher percentage of the objects in the scenes—73 percent versus 64 percent. Glover argues that that difference is because of his algorithm's better ability to determine object orientations.

He also believes that additional sources of information could improve the algorithm's performance even further. For instance, the Bingham distribution could also incorporate statistical information about particular objects—that, say, a coffee cup may be upside-down or right-side up, but it will very rarely be found at a diagonal angle.

Indeed, it's because of the Bingham distribution's flexibility that Glover considers it such a promising tool for robotics research. "You can spend your whole PhD programming a robot to find tables and chairs and cups and things like that, but there aren't really a lot of general-purpose tools," Glover says. "With bigger problems, like estimating relationships between objects and their attributes and dealing with things that are somewhat ambiguous, we're really not anywhere near where we need to be. And until we can do that, I really think that robots are going to be very limited."

Gary Bradski, vice president of computer vision and machine learning at Magic Leap and president and CEO of OpenCV, the nonprofit that oversees the most widely used open-source computer-vision software library, believes that the Bingham distribution will eventually become the standard way in which roboticists represent object orientation. "The Bingham distribution lives on a hypersphere," Bradski says—the higher-dimensional equivalent of a circle or sphere. "We're trying to represent 3-D objects, and the spherical representation fits naturally with the 3-D space. It's just kind of a recoding of the features that has more natural properties."

"It isn't really as hard as the math looks," Bradski adds. "It's a better representation, so I think once it's understood, this'll just kind of become one of the things that is built in when you're doing the 3-D fits. [Glover] found something that was obscure, but once people are familiar with it, it will just be a no-brainer."

The paper is titled "Bingham Procrustean Alignment for Object Detection in Clutter."

*This story is republished courtesy of MIT News ([web.mit.edu/newsoffice/](http://web.mit.edu/newsoffice/)), a popular site that covers news about MIT research, innovation and teaching.*

Provided by Massachusetts Institute of Technology

Citation: Better robot vision: Neglected statistical tool could help robots better understand the objects in the world around them (2013, October 7) retrieved 7 May 2024 from <https://phys.org/news/2013-10-robot-vision-neglected-statistical-tool.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.