

Seeing things: Researchers teach computers to recognize objects

October 13 2009, by Larry Hardesty



A new object recognition algorithm treats unrelated images as if they were consecutive frames of video. Since it assumes that the objects in both images are the same, it tries to deform the objects in the first image until they map onto the objects in the second. If the objects in one of the images have already been outlined and labeled, the algorithm can simply transfer the labels to the other image. Courtesy of Ce Liu

(PhysOrg.com) -- If computers could recognize objects, they could automatically search through hours of video footage for a particular two-minute scene. A tourist strolling down a street in a strange city could take a cell-phone photo of an unmarked monument and immediately find out what it was. And an Internet image search on, say,

"Shakespeare" would pull up pictures of Shakespeare, not pictures of Gwyneth Paltrow in the movie Shakespeare in Love. Though object recognition is one of the major research topics in computer vision, MIT researchers may have found a way to make it much more practical.

Typically, object recognition algorithms need to be "trained" using [digital images](#) in which objects have been outlined and labeled by hand. By looking at a million pictures of cars labeled "car," an [algorithm](#) can learn to recognize features shared by images of cars. The problem is that for every new class of objects — trees, buildings, telephone poles — the algorithm has to be trained all over again.

But Esther and Harold E. Edgerton Associate Professor of Electrical Engineering and Computer Science Antonio Torralba and Computer Science and [Artificial Intelligence](#) Lab graduate students Ce Liu, PhD '09, and Jenny Yuen have developed an object recognition system that doesn't require any training. Nonetheless, it still identifies objects with 50 percent greater accuracy than the best prior algorithm.



Credits - Images courtesy of Ce Liu

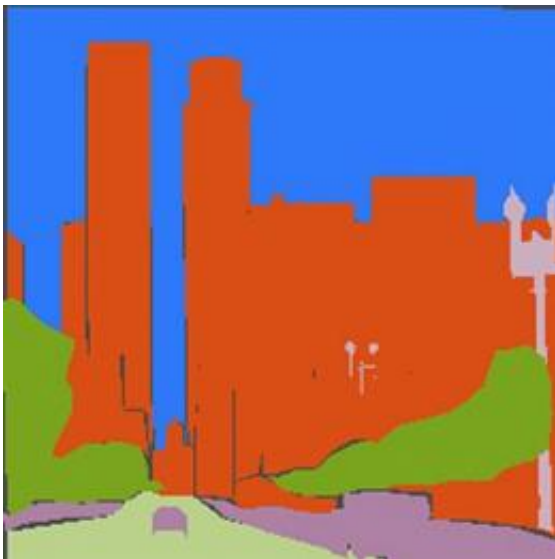
The system uses a modified version of a so-called motion estimation algorithm, a type of algorithm common in video processing. Since consecutive frames of video usually change very little, data compression schemes often store the unchanging aspects of a scene once, updating only the positions of moving objects. The motion estimation algorithm determines which objects have moved from one frame to the next. In a video, that's usually fairly easy to do: most objects don't move very far in one-30th of a second. Nor does the algorithm need to know what the object is; it just has to recognize, say, corners and edges, and how their appearance typically changes under different perspectives.

The MIT researchers' new system essentially treats unrelated images as if they were consecutive frames in a video sequence. When the modified motion estimation algorithm tries to determine which objects have "moved" between one image and the next, it usually picks out objects of the same type: it will guess, for instance, that the 2006 Infiniti in image two is the same object as the 1965 Chevy in image one.

If the first image comes from the type of database used to train computer vision systems, the Infiniti will already be labeled "car." The new system simply transfers the label to the Chevy.



Credits - Courtesy of Ce Liu



Credits - Courtesy of Ce Liu

The greater the resemblance of the labeled and unlabeled images, the better the algorithm works. Fortunately, Torralba's earlier work was

largely directed toward amassing a huge database of labeled images. Torralba and his colleagues have developed a simple web-based system called LabelMe that lets online volunteers tag objects in digital images, and they also created a web site called 80 Million Tiny Images that sorts the images according to subject matter. When confronted with an unlabeled image, the new object recognition algorithm is likely to find something similar in Torralba's database. And as the database grows larger, that likelihood will only increase.

"It's a real commonsense solution to a fundamental problem in computer vision," says Marshall Tappen, a [computer vision](#) researcher at the University of Central Florida. "The results are great and better than you can get with much more complicated methods." Tappen adds that "a large database makes it possible to do lots of really interesting thing that no one's even envisioned. There are lots of interesting things it can do beyond just standard object recognition, so I think it's really going to enable a lot of innovation." Tappen points in particular to recent work on image editing and image completion done by Alyosha Efros at Carnegie Mellon University. "If you look at his last few Siggraph papers" — that is, papers presented at Siggraph, the major conference in the field of computer graphics — "they're all using LabelMe," Tappen says.

Provided by Massachusetts Institute of Technology ([news](#) : [web](#))

Citation: Seeing things: Researchers teach computers to recognize objects (2009, October 13) retrieved 28 April 2024 from <https://phys.org/news/2009-10-seeing-things-researchers-teach-computers.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.
