

# Researchers develop new image-recognition software

May 21 2008

---



Question: What do you see in the red circles? A bottle, a cell phone, a person, a shoe? The answer: They're all the same. Professor Antonio Torralba created these low-resolution images, in which the circled shapes were inserted and are all identical, to demonstrate how context affects our recognition of objects. Even the 'car' in the lower left image is the same object. Photo / Antonio Torralba

It takes surprisingly few pixels of information to be able to identify the subject of an image, a team led by an MIT researcher has found. The discovery could lead to great advances in the automated identification of online images and, ultimately, provide a basis for computers to see like

humans do.

Antonio Torralba, assistant professor in MIT's Computer Science and Artificial Intelligence Laboratory, and colleagues have been trying to find out what is the smallest amount of information--that is, the shortest numerical representation--that can be derived from an image that will provide a useful indication of its content.

Deriving such a short representation would be an important step toward making it possible to catalog the billions of images on the Internet automatically. At present, the only ways to search for images are based on text captions that people have entered by hand for each picture, and many images lack such information. Automatic identification would also provide a way to index pictures people download from digital cameras onto their computers, without having to go through and caption each one by hand. And ultimately it could lead to true machine vision, which could someday allow robots to make sense of the data coming from their cameras and figure out where they are.

"We're trying to find very short codes for images," says Torralba, "so that if two images have a similar sequence [of numbers], they are probably similar--composed of roughly the same object, in roughly the same configuration." If one image has been identified with a caption or title, then other images that match its numerical code would likely show the same object (such as a car, tree, or person) and so the name associated with one picture can be transferred to the others.

"With very large amounts of images, even relatively simple algorithms are able to perform fairly well" in identifying images this way, says Torralba. He will be presenting his latest findings this June in Alaska at a conference on Computer Vision and Pattern Recognition. The work was done in collaboration with Rob Fergus at the Courant Institute in New York University and Yair Weiss of Hebrew University in Jerusalem.

To find out how little image information is needed for people to recognize the subject of a picture, Torralba and his co-authors tried reducing images to lower and lower resolution, and seeing how many images at each level people could identify.

"We are able to recognize what is in images, even if the resolution is very low, because we know so much about images," he says. "The amount of information you need to identify most images is about 32 by 32." By contrast, even the small "thumbnail" images shown in a Google search are typically 100 by 100.

Even an inexpensive current digital camera produces images consisting of several megapixels of data--and each pixel typically consists of 24 bits (zero or one) of data. But Torralba and his collaborators devised a mathematical system that can reduce the data from each picture even further, and it turns out that many images are recognizable even when coded into a numerical representation containing as little as 256 to 1024 bits of data.

Using such small amounts of data per image makes it possible to search for similar pictures through millions of images in a database, using an ordinary PC, in less than a second, Torralba says. And unlike other methods that require first breaking down an image into sections containing different objects, this method uses the entire image, making it simple to apply to large datasets without human intervention.

For example, using the coding system they developed, Torralba and his colleagues were able to represent a set of 12.9 million images from the Internet with just 600 megabytes of data--small enough to fit in the RAM memory of most current PCs, and to be stored on a memory stick. The image database and software to enable searches of the database, are being made publicly available on the web.

Of course, a system using drastically reduced amounts of information can't come close to perfect identification. At present, the matching works for the most common kinds of images. "Not all images are created equal," he says. The more complex or unusual an image is, the less likely it is to be correctly matched. But for the most common objects in pictures--people, cars, flowers, buildings--the results are quite impressive.

The work is part of research being carried out by hundreds of teams around the world, aimed at analyzing the content of visual information. Torralba has also collaborated on related work with other MIT researchers including William Freeman, a professor in the Department of Electrical Engineering and Computer Science; Aude Oliva, professor in the Department of Brain and Cognitive Sciences; and graduate students Bryan Russell and Ce Liu, in CSAIL. Torralba's work is supported in part by a grant from the National Science Foundation.

Torralba stresses that the research is still preliminary and that there will always be problems with identifying the more-unusual subjects. It's similar to the way we recognize language, Torralba says. "There are many words you hear very often, but no matter how long you have been living, there will always be one that you haven't heard before. You always need to be able to understand [something new] from one example."

Source: MIT

Citation: Researchers develop new image-recognition software (2008, May 21) retrieved 19 April 2024 from <https://phys.org/news/2008-05-image-recognition-software.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.