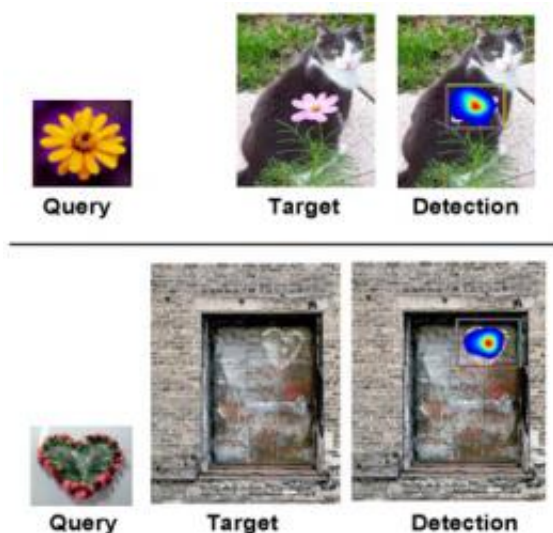# New search technique for images and videos has broad applications

November 10 2009, By Daniel Strain

Using a single image as a template, computer software can find similar images in a large database of photos, as shown in these examples. Images courtesy of P. Milanfar.

(PhysOrg.com) -- Engineers at the University of California, Santa Cruz, have developed a powerful new approach to a fundamental problem in computer vision: how to program a computer to recognize or categorize what it "sees" in an image or video. Their software could change the way people search the Web for photos and videos, and it may have applications in many other areas as well, such as video surveillance and security systems.

Peyman Milanfar, a professor of electrical engineering in the Baskin School of Engineering at UCSC, and graduate student Hae Jong Seo were able to overcome a major drawback of existing methods for computer recognition of objects in images--the need for an extensive "training" phase using a large number of examples. With a single photograph or video clip as a template, their software can sift through thousands of images or videos to pull out the ones that look like the template.

"When you search Google Images, you type in a term and it gives you returns from pages that have that text in them. We want to be able to upload an image and use it as a model for finding similar images," Milanfar said.

Milanfar and Seo developed an algorithm that enables automated recognition of both objects in images and actions in videos. The software analyzes an image or short movie and characterizes the most important constituents of the object or action represented. It can then search for those constituents in image and video databases. The researchers presented their new methods at the IEEE International Conference on Computer Vision in September and in a recent paper published by the *IEEE Transcripts on Pattern Analysis and Machine Intelligence*.

"When it comes to recognizing things in the visual world, humans have some uncanny abilities which, at least until now, well exceed the limits of what could be done by computer," Milanfar said. "In particular, we have the capacity to recognize an object after having seen it only once."

Existing technology can search for and distinguish individual objects in a database of images only after running through a time-consuming training phase. "If you're looking for images of bicycles, for instance, current algorithms have to be shown pictures of hundreds, if not thousands, of

bicycles in order to be able to recognize a bicycle," Milanfar said.

With his new software, a single photo of a bicycle at night can be used as a template to locate pictures of bicycles in full sunlight, in the foreground or the background. It works under a wide range of image qualities and lighting discrepancies. The template image or the target image can be sharp or out-of-focus, clean or noisy. To Milanfar's software, a bicycle is a bicycle.

Similarly, a person riding a bicycle is a person riding a bicycle. Video of Lance Armstrong in the Tour de France can be used to find clips of men and women riding along an ordinary street.

But the potential applications for Milanfar's work go well beyond browsing for cyclists on YouTube. By using videos of aggressive behavior as templates, the technology could help surveillance systems learn to recognize potentially dangerous situations. If a man reached for a weapon on camera and that action matched a template of such behavior, surveillance software could alert a busy security guard.

A picture is a composite of thousands of pixels. Milanfar's software examines these pixels and their relation to one another. In other words, how similar is a central pixel to adjacent pixels in orientation, coloring, and shading? To find actions within videos, like a man riding a bicycle, Milanfar's software completes the same procedures but incorporates the manner in which those pixel relationships move over time.

The software analyzes the map of pixel relationships and determines the salient geometric features of the object or action. These components remain perceptually constant within an object regardless of image quality.

"The geometry of the bicycle is recognizable by the shape of the wheels

and the way they are connected to the body, for example," Milanfar said. "We compute features from an image that are very stable. They are there even if we make the object bigger or smaller, change the background, or add noise."

Search engines can use this algorithm to detect similar patterns of pixel relationships in a whole database of photos. The software calculates the statistical likelihood that a candidate image contains the queried object. If the template is a bicycle, the outcome consists of a series of photographs containing bicycles of all shapes and sizes, ranked in order of similarity.

"This has been an area of research that has entertained people for many years, but the big successes have been few and far between," Milanfar said. "Our work is showing state-of-the-art performance with an accuracy as good as or better than any algorithm out there."

Provided by University of California, Santa Cruz

Citation: New search technique for images and videos has broad applications (2009, November 10) retrieved 19 April 2024 from https://phys.org/news/2009-11-technique-images-videos-broad-applications.html