Team creates computerized method for matching images in photos, paintings, sketches

December 6 2011

Computers can mimic the human ability to find visually similar images, such as photographs of a fountain in summer and in winter, or a photograph and a painting of the same cathedral, by using a technique that analyzes the uniqueness of images, say researchers at Carnegie Mellon University's School of Computer Science.

The research team, led by Alexei Efros, associate professor of computer science and robotics, and Abhinav Gupta, assistant research professor of robotics, found that their surprisingly simple technique performed well on a number of visual tasks that normally stump computers, including matching sketches of automobiles with [photographs](#) of cars.

The team from the Robotics Institute and [Computer Science](#) Department will present its findings on "data-driven uniqueness" on Dec. 14 at [SIGGRAPH](#) Asia, a [computer graphics](#) and interactive techniques conference in Hong Kong. Their research paper is available online [here](#).

Most computerized methods for matching [images](#) — in contrast to image searches based on keywords — focus on similarities in shapes, colors and composition. That approach has proven effective for finding exact or very close image matches and enabled successful applications such as Google Goggles.

But those methods can fail miserably when applied across different

domains — photographs taken in different seasons or under different lighting conditions, or in different media, such as photographs, color paintings or black-and-white sketches. "The language of a painting is different than the language of a photograph," Efros explained. "Most computer methods latch onto the language, not on what's being said."

One problem, Gupta said, is that many images have strong elements, such as a cloud-filled sky, that may have superficial similarities to other images, but really only distract from what makes the image interesting to people. He and his collaborators hypothesized that it is instead the unique aspects of an image, in relation to other images being analyzed, that sets it apart and it is those elements that should be used to match it with similar images.

On the pixel level, a photo of a garden statue in the summer or fall will look very different than the same statue photographed in winter, said Abhinav Shrivastava, a master's degree student in robotics and first author of the research paper. But the unique aspects of the statue will carry over from a summer image to a winter image, or from a color photo to a sketch.

Estimating uniqueness is no simple task. The team computes uniqueness based on a very large data set of randomly selected images. Features that are unique are those that best discriminate one image from the rest of the random images. In a photo of a person in front of the Arc de Triomphe in Paris, for instance, the person likely is similar to people in other photos and thus would be given little weight in calculating uniqueness. The Arc itself, however, would be given greater weight because few photos include anything like it.

"We didn't expect this approach to work as well as it did," Efros acknowledged. "We don't know if this is anything like how humans compare images, but it's the best approximation we've been able to

achieve."

In addition to automated image searches, this technique has applications to computational rephotography — the combination of historic photographs with modern-day photos taken from the same perspective. By using the new technique, it may be possible in many cases to eliminate the need for rephotography by simply matching the historic photo with an existing online photo that matches its perspective. Likewise, the technique can be combined with large GPS-tagged photo collections to determine the location where a particular painting of a landmark was painted.

The technique also can be used to assemble a "visual memex" — a data set that explores the visual similarities and contexts of a set of photos. For instance, the researchers downloaded 200 images of the Medici Fountain in Paris — paintings, historic photographs and recent snapshots from various seasons and taken from various distances and angles — and assembled them into a graph, as well as a YouTube video that shows a particular path through the data.

Future work includes using the technique to enhance object detection for computer vision and investigating ways to speed up the computationally intensive matching process.

Provided by Carnegie Mellon University