

Photo software creates 3-D world



Photo tourism is a system for browsing large collections of photographs in 3D. Researcher's approach takes as input large collections of images from either personal photo collections or Internet photo sharing sites, and automatically computes each photo's viewpoint and a sparse 3D model of the scene. Photo explorer interface enables the viewer to interactively move about the 3D space by seamlessly transitioning between photographs, based on user control. Credit: University of Washington

In the digital age, organizing a photo collection has gone from bad to worse. The saying used to be that a picture is worth a thousand words. Now the question arises: what are a thousand pictures worth? In a word, a headache.

"Anyone who has a digital camera has the problem that they have more photos than they can possibly navigate," says Steve Seitz, associate professor of computer science & engineering. "And it's always a problem to find the photo that you're looking for."

Now experimental software developed by UW and Microsoft computer scientists, called Photo Tourism, turns the surfeit of images into a benefit. Hundreds of photos of a single scene can be mapped into a 3-D virtual world. The technology has potential not just for organizing photo collections, but for capturing scenes and, perhaps someday, creating a visual map of all the photos on the Internet.

Over the past year the research has catapulted to the marketplace. Early work attracted attention in March at Microsoft's TechFest meeting. The project again made headlines in August when it was presented at a major graphics conference. Microsoft Live Labs signed a commercial license for the prototype software last August. Within a few months the company shipped a technology preview of a product that it called Photosynth.

"It's been great to see a lot of people excited about it, and it's also been a thrill to just have something, especially so quickly, that people could look at and use," says doctoral student Noah Snively. While Photosynth follows its own trajectory, Snively will continue to develop Photo Tourism for his doctoral thesis, in collaboration with Seitz, an expert in computer vision, and Rick Szeliski, an employee at Microsoft Research and affiliate professor at the UW.

Snively arrived from the University of Arizona three years ago interested in researching computer graphics. His target was not just personal photos collections but massive online collections, such as on the popular photo-sharing Web site Flickr. Members' contributions to Flickr now total more than 200 million images.

"I was kind of inspired by that," Snively says.

If you type "Trevi Fountain" in Flickr's search box, you will find more than 11,000 photos. Browsing through these photos means clicking through page after page of miniature pictures. Anyone who's performed an image search on Google can appreciate the frustration. Finding a photo similar to what you need still won't bring you any closer to the perfect shot.

"You might look at a photo and say I wonder what's just to the left of it, or I wonder what's just to the right of it, or I wish I could expand the field of view," Snively explains. It's a challenge just to find the same scene taken at different times of the day. Trevi Fountain was the test case. (Snively has never been there, though by now he's seen it from almost every angle.) Later experiments used scenes of Notre Dame

Cathedral in Paris and Half Dome mountain in Yosemite National Park.

To solve the problem, the researchers harnessed recent advances in computer vision research. They wrote computer software that analyzes each image and calculates where it was taken. To do this, the software looks for small details shared between different photos that can be used to compare them and stitch them together in three dimensions. Each photo is then represented by a small square placed in the appropriate position in a sketch of the original scene.

The effect is that you're sifting through hundreds or thousands of photos, but it feels more like a video game. By moving right or left, or zooming in and out, the computer will fade to an appropriate shot. Highlighting a feature, like Neptune statue at the center of Trevi Fountain, brings up a high-resolution photo of that object.

This software goes beyond simply organizing a photo collection, Seitz says. It recreates a particular scene or location at the resolution of the photos. Real estate agencies, museums and hotels might find it a useful way to present a virtual tour because viewers could zoom in to read a restaurant menu or to view a painting. Archaeologists and biologists have expressed interest in creating realistic visual representations of their research sites. Military and surveillance organizations also would like to organize photographs in an intuitive way. Sports enthusiasts could even recreate their favorite game by combining all the photos taken at an event.

The current interface presents each photo as a little box, and photos fade into one another to give the impression of a 3-D zoom. Current research will create an even "more fluid, game-like interface," Seitz says. Users will feel as if they are navigating a 3-D world.

Companies such as Google and Microsoft recently have begun to create 3-D models of cities by painstakingly gathering photos taken from different angles and then stitching them together. Photo Tourism doesn't feel as smooth -- there are gaps, and people sometimes pop up in the photos -- but in the long term this ad-hoc method for combining photos taken at varying scales may offer advantages.

"I think it has the possibility to be much, much richer than just a static 3-D model," Snavelly says.

The most promising application for Photo Tourism, he believes, may be organizing the millions of photos that exist on the Internet. Snavelly describes the concept as a "visual Wikipedia." Contributors could upload photos and the program would combine them to create an increasingly comprehensive picture of the world. Combining the photos with a digital map like Google Earth would mean users could keep zooming in closer without the image ever going fuzzy.

But scaling up to handle millions of photos is still a ways off, Seitz says. "That's another major research project."

For more information and a trial version of Photo Tourism, go to <http://phototour.cs.washington.edu/>

Source: University of Washington

This document is subject to copyright. Apart from any fair dealing for the purpose of private study, research, no part may be reproduced without the written permission. The content is provided for information purposes only.