Three Software Tricks for Sorting Through the Barrage of UAV Footage

By Ariel Bleicher

Posted 1 Sep 2010 | 14:24 GMT

PHOTO CREDIT: Lt. Col. Leslie Pratt/USAF. *An MQ-9 Reaper drone flies a combat mission over southern Afghanistan.*

In 2009 alone, the U.S. Air Force shot 24 years' worth of video over Iraq and Afghanistan using spy drones (UAVs). With so many planes in the air, and more and more cameras being attached to each plane, the Air Force is generating more footage than analysts can sift through, the *New York Times* reported (http://www.nytimes.com/2010/01/11/business/11drone.html) back in January. "We're going to find ourselves in the not too distant future swimming in sensors and drowning in data," Lt. Gen. David A. Deptula the U.S. Air Force's top intelligence official, told *National Defense Magazine*

(http://www.nationaldefensemagazine.org/archive/2010/January/Pages/Military%E2%80%98SwimmingInSensorsandDrowninginData%E2%80%99.aspx) the same month.

Of course, the best way to reach the surface of such a large data pool is to get a computer to show you which way to swim. As one might guess, programming a system to automatically search video and pick out noteworthy information is not an easy problem. And so far, no one has developed software that can keep up with the Air Force's high-tech hardware. But Mubarak Shah (http://longwood.cs.ucf.edu/~vision/faculty/shah.html), who founded and now directs the Computer Visions Lab at the University of Central Florida, recently presented a few ideas about how to make improvements.

Shah, who spoke Monday at the *7th IEEE International Conference on Advanced Video and Signal-Based Surveillance* (http://www.avss2010.org/) in Boston, Massachusetts, is a graying, deep-eyed fellow. According to the computer engineer perched next to me, Shah is "the guy you've heard about if you know anything about anyone in video surveillance."

Shah focused on three problems for surveillance software that are notoriously difficult, particularly when you're trying to analyze video shot thousands of meters above the ground from a drone flying hundreds of kilometers per hour.

Follow the Dots

The first problem he addressed was how to track big swarms of objects, such as cars, traveling over a wide area, such as an expressway. The difficulty of this task lies in the fact that, when shot from above, cars traveling on an expressway are exceedingly small (no more than 30 pixels), and there are thousands of them. Plus, the plane that's shooting footage is moving faster than the cars it's capturing, so you've only got a few frames to work with for each car.

Shah's solution to this problem depends on keeping track of all the possible paths a vehicle may have taken, then weeding out the poor choices based on common sense (Two vehicles probably didn't cross paths at the same time...unless, of course, they crashed) and a bit of modern transportation theory (If one car is behind another car, it's probably accelerating at a similar rate). It looked like he got pretty good results: about 80-90 percent accuracy.

Wait... Is That a Pedestrian or a Palm Tree?

If you think antlike cars are hard for a computer to follow, you can imagine the difficulty in trying to program it to find a person. Some of the best people-detection systems rely on histograms (http://en.wikipedia.org/wiki/Histogram_of_oriented_gradients)—statistical distributions that determine whether an object is a person based on probabilities. But because people are so very tiny in aerial images, the histogram method mistakes quite a lot of things (trees, mailboxes, stoplights) for people. "There can be thousands of those that are completely wrong," Shah says. He proposes using some basic 8th-grade geometry tricks to find a person based on the relationship between the height of an object and the length of its shadow. He admits that this strategy wouldn't work so well on video shot on cloudy days or using infrared light.

Mapping by Motion

The third—and in my opinion, coolest—tool Shah presented was a method for determining movement patterns. Say, for example, you have some aerial footage of an Afghan city and you want to automatically know how it's laid out—where the roads are, the bridges, the intersections, where people regularly travel, the areas they avoid, where they gather. What a computer sees, however, is "very noisy optical flow," Shah says—lots of motion but not much order. Using a mathematical noise-reducing tool known as a "mixture of Gaussians," Shah can find order in the static-like mess of optical data and get his software to draw a picture of the city in motion. "Using this, we can basically discover the road networks," he says.

Automating video search seems an ambitious project, but one that needs to be done. The newer Reaper drones (http://www.airforce-technology.com/projects/predator/) now shoot video in 10 directions at once, and the Air Force plans to eventually upgrade that number to 65. That's 65 video streams coming from one spy plane. The deluge of data isn't stopping, and there just aren't enough eyeballs to sort through it all.