The Economist Technology Quarterly December 2nd 200

It all depends on your point of view

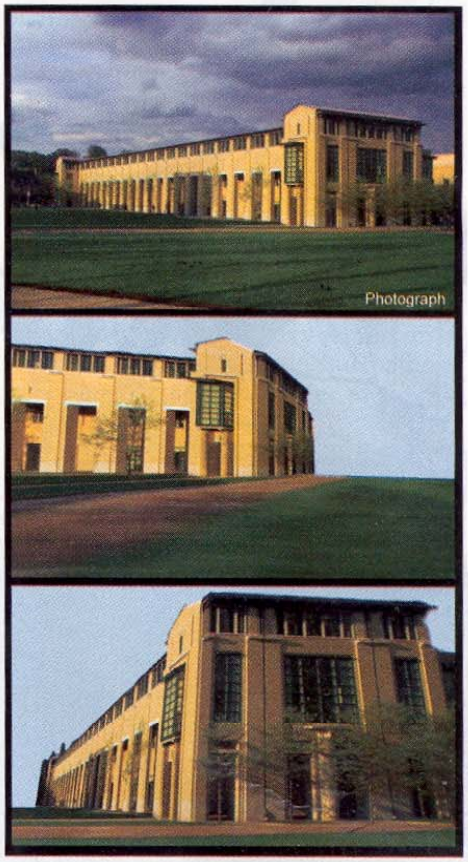
Computing: New techniques analyse two-dimensional pictures to produce detailed three-dimensional models of the world

COMPUTERS can look, but they cannot see. Cheap digital sensors can act as their eyes, but programming machines to make sense of what they see is extremely difficult. Even when they can identify faces and vehicles, computers' inability to understand context results in ludicrous mistakes, such as finding faces in clouds or cars halfway up trees. Humans, by contrast, are able to construct a mental model of a scene from a photograph by taking into account the relative sizes of recognised objects, the laws of physics and some basic common sense. Now several research groups are building new computer-vision systems to enable computers to do the same thing.

Researchers at Carnegie Mellon University believe they have achieved a breakthrough in the reconstruction of three-dimensional models from two-dimensional images (pictured). Their system analyses photographs of outdoor scenes, identifies "sky" and "ground" regions, and looks for visual cues that distinguish horizontal surfaces from vertical ones. It then reconstructs the scene by cutting and folding the original image, taking into account the constraints that apply in the real world: skies are blue, horizons are horizontal and most objects sit on the ground. "In our world things don't just float," says Martial Hebert, who co-developed the software with his colleagues Alexei Efros and Derek Hoiem.

Using multiple images of a particular object or scene reduces ambiguity and makes possible more accurate three-dimensional reconstructions. That is the approach taken by Photosynth, a system being developed by researchers at Microsoft's Live Labs, a joint venture between the software giant's research arm and its MSN portal. Photosynth trawls the internet for digital photos of a place or object. Each photo is analysed to extract hundreds of distinctive features, and images that share particular features are linked together. The software then works out the relationship between the features to generate a three-dimensional model through which users can navigate.

What might such systems be used for? Pascal Fua of the Ecole Polytechnique Fédérale de Lausanne, in Switzerland, has built a system that analyses video footage from a single camera to reconstruct how an object moves in three dimensions. He and his team are using the technology with the yacht *Alinghi*, the present holder of the America's Cup. The idea is to improve the design of the yacht's sails by analysing how they behave under actual sailing conditions, thus dispensing with the need for expensive wind-tunnel tests that might not have been able to reproduce racing conditions accurately.



No special glasses required

Dr Fua's technology is also being developed for surveillance applications. The aim is to combine video and still images from a network of video cameras, on the ground and in the air, to generate a constantly updated three-dimensional representation of an area under surveillance, tracking and analysing individuals and groups and triggering alarms if appropriate. This project, called Dynamic Visual Networks, involves a consortium of European firms and universities.

Richard Radke of the Rensselaer Polytechnic Institute in New York is working in a similar vein. He is developing software to allow hundreds of cameras, working together in small groups, to analyse their surroundings. Wireless camera "nodes" could be randomly sprinkled in large numbers over the area of interest, which might be a battlefield or the scene of a natural disaster. Each one would compose a list of distinctive features it could see and then quiz its neighbours to see if they could see any of the same features from their vantage points. By combining the results gathered, it would then be possible to reconstruct the scene. In future, some of the cameras might even be able to crawl or fly to fill in further detail.

Constructing detailed, real-time three-dimensional models of places from swarms of tiny cameras—a virtual model of a real scene—would have mind-boggling applications. Relief workers could fly through a model of a disaster area to look for survivors and guide rescue helicopters. Soldiers could look around corners or inside buildings before launching an attack. And security guards could patrol a wide area by whizzing around it in virtual rather than real space. What of the implications for privacy and civil liberties? Like the technology itself, it all depends on your point of view. ■