Text

Place a plain-coloured vase on a table. Take a photo and print it in black and white. Now smash the vase into pieces, gather them up, and try to rebuild it based entirely on the picture.

Easy? Not according to Edwin Hancock, Professor of Computer Vision at York, who works on the so-called 'shape-from-shading' problem in York's Department of Computer Science. "We humans can glance at a flat image and immediately interpret it correctly in three dimensions. But in its simplest form, that task – generating a 3D shape from a 2D picture – is mathematically unsolvable," Professor Hancock explains. "So the fact that we can do it is very interesting. It means our brains must be doing some work in the background."

Think about the vase photograph again for a moment. For each pixel on the photo, you only have one piece of information, namely its brightness. But a light-coloured pixel might be a dull patch on a raised surface, or a shiny patch on a recessed one, or it might simply be close to a light source. Somehow, your brain has to resolve multiple variables to make sense of the whole picture – and when you get right down to it, a set of pixels could be interpreted according to any one of infinite number of possible values for those variables, with each interpretation representing a different 3D shape.

According to Professor Hancock and his team, our brains get around this problem by relying on some automatic assumptions: that areas of similar lightness are likely to represent smooth surfaces, that light sources generally come from above, and so on. By making educated guesses about some of the variables, the brain can eliminate uncertainties and reduce an unsolvable problem to an approachable one. The Gregory illusion [refer to illusion.tif] shows these assumptions at work.

This technique mean the brain can be fooled, of course. The use of 'forced perspective', where a shallow stage backdrop is made to look like a long corridor leading into the distance, has been a favourite theatrical trick since Victorian times. And the York researchers are quick to point out that even the Gregory illusion allows for many different interpretations [refer to ambiguity.jpg]. But the ability to visualise 3D objects reliably has been crucial to our success as a species, and the fact that we fall into the occasional trap is a small price to pay.

Now the challenge facing the Computer Vision researchers is to teach a computer program these same tricks – and they're adding a few of their own. One novel approach measures the polarisation state of the light reflected from an object's surface and then uses this information to calculate the likely position of the light source. Another combines simultaneous stereo images of an object – similar to the images captured by a pair of human eyes – and then factors in estimates about the object's surface based on shading and texture.

[Sub-heading] Face-off

One particular focus of the team's work is modelling faces. "There are good reasons for giving facial recognition special treatment," says Computer Science lecturer Will Smith. "If I show you an obviously facial image, your brain will invoke all kinds of new assumptions to process it."

Working with colleagues from the Department of Psychology, Dr Smith has identified some of these unique assumptions. It turns out our internal face-spotting software is so powerful that it can even override other, more basic ideas about light source direction and perspective in order to make sense of what it's seeing. Put simply, people are face-spotting specialists.

Convincing a computer to develop the same kind of expertise is a difficult task, but the team are making good progress, and the key weapon in their armoury is statistics. It helps that most faces are broadly similar, so by collecting a large number of images using a Hollywood-style 3D scanner, the program can derive some basic ground rules about faces in general (they always have eyes here, nose there). These rules can then be converted into guiding principles for analysing new images. Another technique involves writing a program to compare various 3D models with their 2D counterparts and make correlations which can then be applied across the board.

Combined with more technical approaches based on principles borrowed from cartography and geometry, the researchers have developed a sophisticated tool that can take a straightforward facial image, model it in 3D, and then process the model to produce a rotated image in a new pose. The results are impressive, and instantly recognisable.

So how far are we from being able to use these new tools in real life? "Day-to-day applications are still some way off," Dr Smith admits, "but the value of being able to derive an accurate model from a single facial image such as a CCTV snapshot is obvious. And constructing a biometric recognition system is very much within our sights."

Pull quotes

"In its simplest form, the task of generating a 3D shape from a 2D picture is mathematically unsolvable."

"Working with colleagues from the Department of Psychology, researchers have identified some of the unique assumptions used by our brains' inbuilt face-spotting software."

Images and captions

- illusion.tif: An illusion devised by Professor Richard Gregory. These markings could equally be interpreted as mounds or dimples, but the brain prefers one interpretation over the other, working on the assumption that light tends to come from above. Try turning the page upside-down and see how your perception changes.
- ambiguity.jpg: "Viewed from above, all four of these shapes would give rise to the same 2D image."
- [One of the existing media images from the 'shape-from-shading' story.]