- CBS Interactive
- **BNET UK**
- **silicon.com**
- **ZDNet UK**

silicon.com - Technology insight. Business leadership.

Photos: 3D replicas to help sell your home? That's computer vision

Seeing through the eyes of a computer at Toshiba's Cambridge R&D Lab

By Natasha Lomas, **30 March 2010 16:38**

FEATURE

The human brain processes the gigabytes of data per second delivered through our eyes, enabling us to see and interact with our surroundings. Computer vision technology seeks to give machines the same ability - to process and act on visual data. For example, using a camera as its eye, a computer vision system could perform tasks such as recognising and following objects - something humans can do very easily but that computers have struggled to do until now.

It is also conjuring up a world of digitised objects - from everyday items you might sell on eBay to the human face and body, all perfectly captured in three dimensions where they can take on a whole new life of their own.

Professor Roberto Cipolla, the MD of Toshiba's Cambridge Research Lab and professor of information engineering at Cambridge University, has spent more than two decades working in the field of computer vision - and believes the technology is on the brink of going commercial in a big way.

Want to sell your house or car online? Imagine uploading a 3D replica of it, faithful in every detail down to the dent on the bumper and the scratch on the front door. Fancy a radical new hairstyle but too scared to go straight for the snip? With a digital replica of your head you could experiment with a Hoxton fin without setting foot in the barber shop.

Cipolla imagines just such commercial applications arriving in the next few years. "This is going to be the decade of computer vision," he tells silicon.com.

Computer vision technology has already been used for digitising museum collections including collections at the Louvre, and as an aid to design processes. Cipolla has worked with sculptor Antony Gormley, digitising photographs of casts of the artist's body and creating 3D models which are in the process of being turned into metres-high geometric giants. A small statuette of Gormley created from one of these 3D models is pictured above.

But the professor's research is aimed at enabling the technology to be applied with low cost hardware such as handheld digital cameras, rather than expensive and cumbersome alternatives such as laser scanners.

Technological challenges such as reconstructing shiny, non-patterned objects - such as the statuette pictured below - and textureless and mobile surfaces such as the human face and body have now been overcome,

according to Cipolla.

Old and new: A traditional porcelain statuette and a small geometric model of Antony Gormley's head, printed via 3D printer

Photo credits: Natasha Lomas/silicon.com

The Computer Vision Group at Toshiba's lab works on modelling shapes from images and video to build accurate three-dimensional models of objects. Pictured above is a prototype real-time 3D modelling system at the lab that can digitally reconstruct an object in real-time.

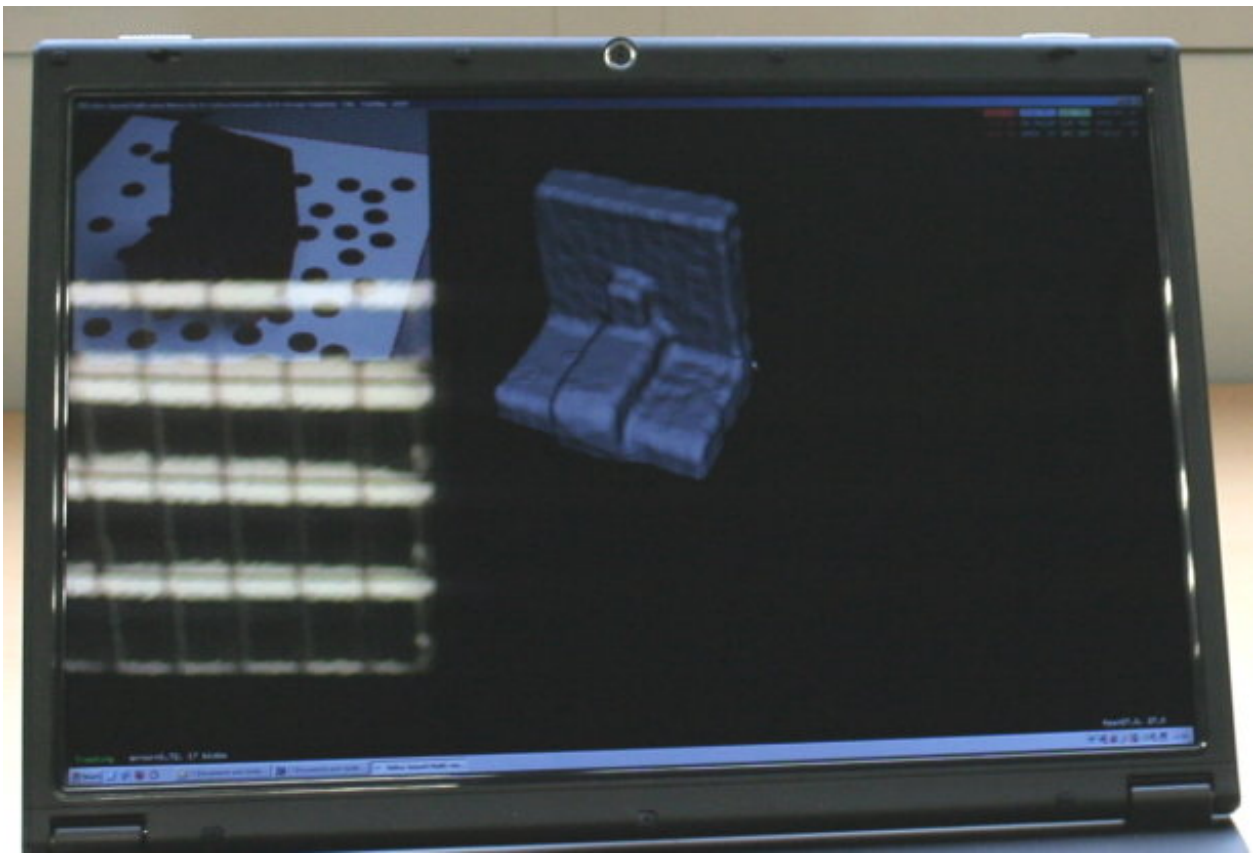
The multiview stereo 3D modelling system uses a single camera connected to a laptop, plus a turntable where the object is placed. The ultimate aim is to have a system that works with just a handheld video camera - and without the need for a patterned surface beneath the object. The team also wants to be able to texture-map the reconstructed objects.

"It's still a work in progress," says Dr Björn Stenger, senior research engineer. "[In future we want to] use it on handheld cameras and reconstruct more densely so it's similar to augmented reality systems where you also can superimpose things.

"We want to actually capture more information about the objects that we're looking at so there you can take it,

build a 3D model and do other things with it - like put it in a game, for example, or also if you want to sell things on the internet there's some ideas to do that. Of course for artists too, for engineers too to get measurements [to manufacture mechanical or design parts].

"Once you have 3D information you can also reproduce the object for example - using 3D printers to make copies."



The polystyrene object (shown above) digitally recreated with a complete surface but without texture mapping

Photo credits: Natasha Lomas/silicon.com

Stenger said another future application for the technology could be selling houses - enabling a digitally accurate model of the outside and inside of a house to be created by walking around it with a camera.

Another use could be putting video cameras into cars and using the system as a driver aid - by reconstructing road scenes as the car travels along to help with driver safety and parking, and ultimately enable **driverless cars**.

An in-car computer vision system could detect when a driver hasn't seen a car stopped in front of it, or when they are in danger of falling asleep, according to Cipolla. "It can look at you when you're driving and see if you're blinking and falling asleep, so warn you. It can actually look outside the lanes and see your driving is very erratic, you seem to be crossing over frequently and correcting sharply - a very strong sign you're about to fall asleep," he says.

"But these are future applications," adds Stenger. "At the moment we're looking at solving basic technological problems."

A toy car being digitally reconstructed in real-time at Toshiba's Research Lab

Photo credits: Natasha Lomas/silicon.com

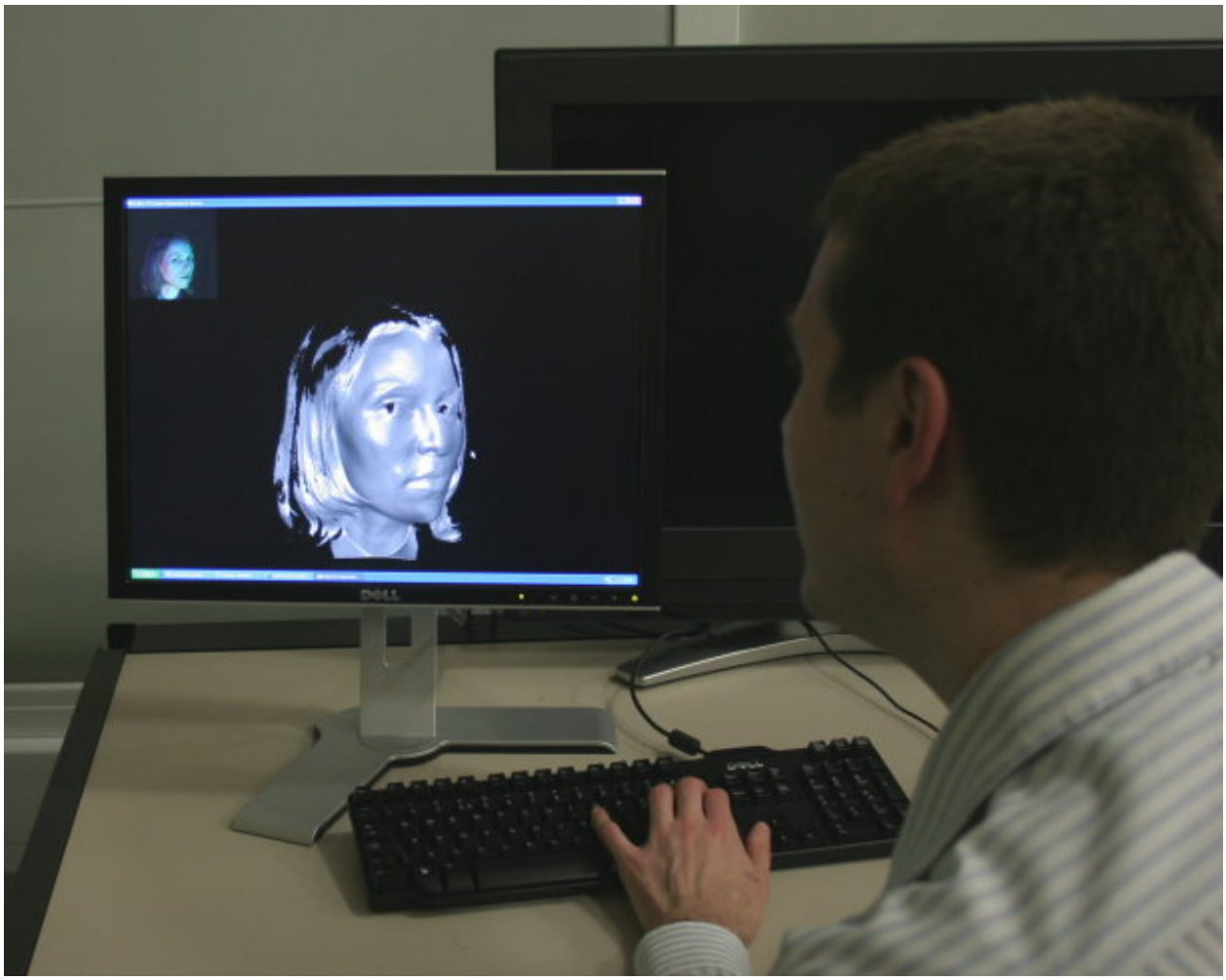
When it comes to capturing the raw shape of the human body and face in real-time the multiview stereo system is no good - humans move and expressions are, by nature, mobile. However, pictured above is another 3D modelling technology developed at Toshiba's labs that has been designed to capture the human body and face moving in real-time - yet is still faithful to every individual lump and bump.

The system works by using three coloured lights - red, green and blue - cast from different fixed positions in a dark room onto the subject (in this case Stenger, pictured above and below). "All you're doing is illuminating the object with RGB light," notes professor Cipolla. "One camera. And you get the most accurate 3D surface with all its blemishes."

Toshiba's RGB system captures facial expressions in real-time

One potential usage for the tech is to create avatars that are not just cartoonesque versions of the computer user but an exact copy. Gamers would then be able to upload their digital double into their favourite games.

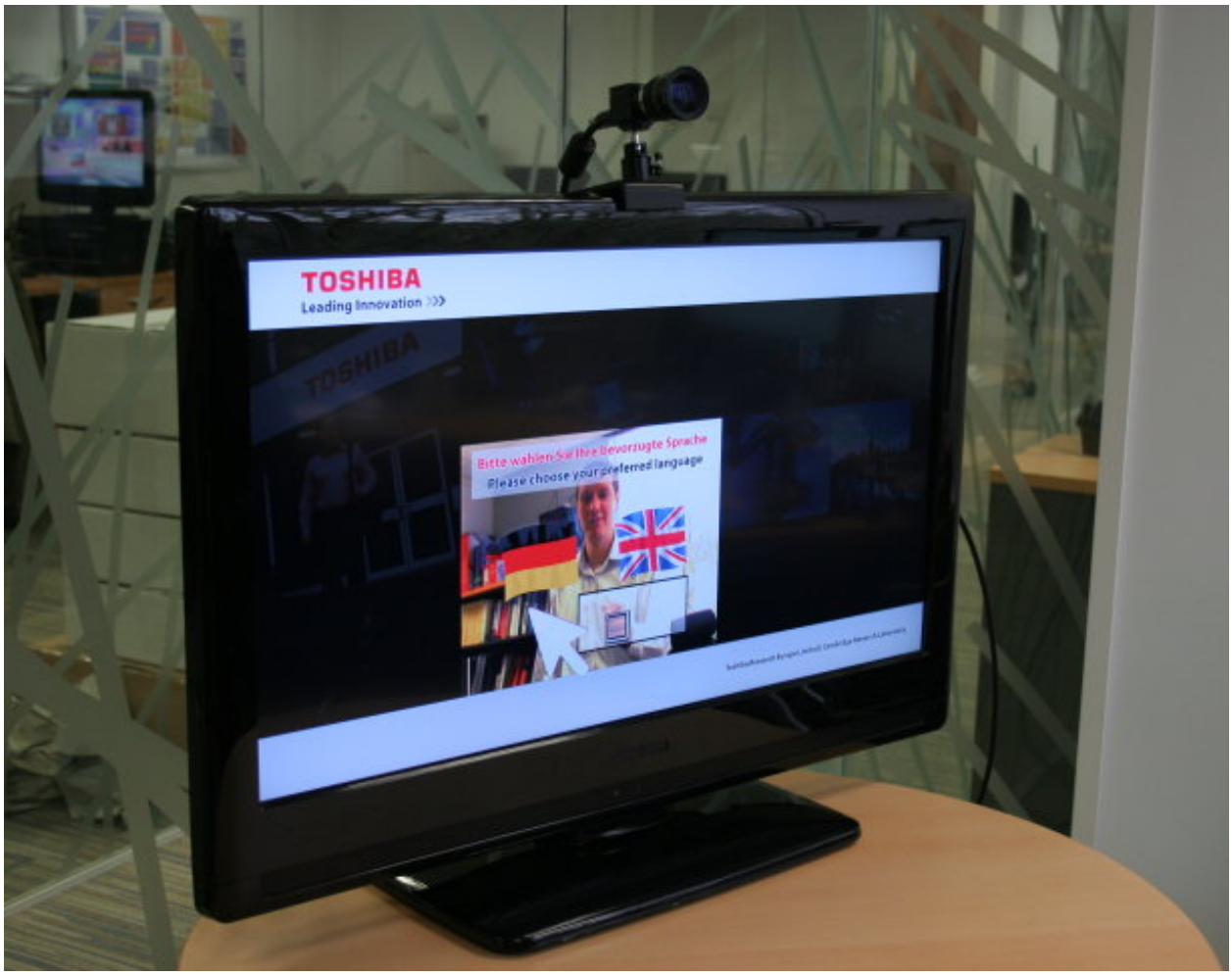
Cipolla also envisages a use-case for virtual actors and actresses - where movie stars can be rented by the hour and their captured digitised expressions used to emote a script, the next best thing to (and presumably cheaper than) the real thing. Another application could be for email messages that can be 'read out' by your avatar, complete with realistic smiles and frowns.



A real-time 3D digital model of silicon.com's Natasha Lomas

Photo credits: Natasha Lomas/silicon.com (top and middle), Roberto Cipolla (above)

Another area where computer vision technology has been applied at Toshiba's Research Lab is gesture-based interfaces - including the interface shown on the TV above.

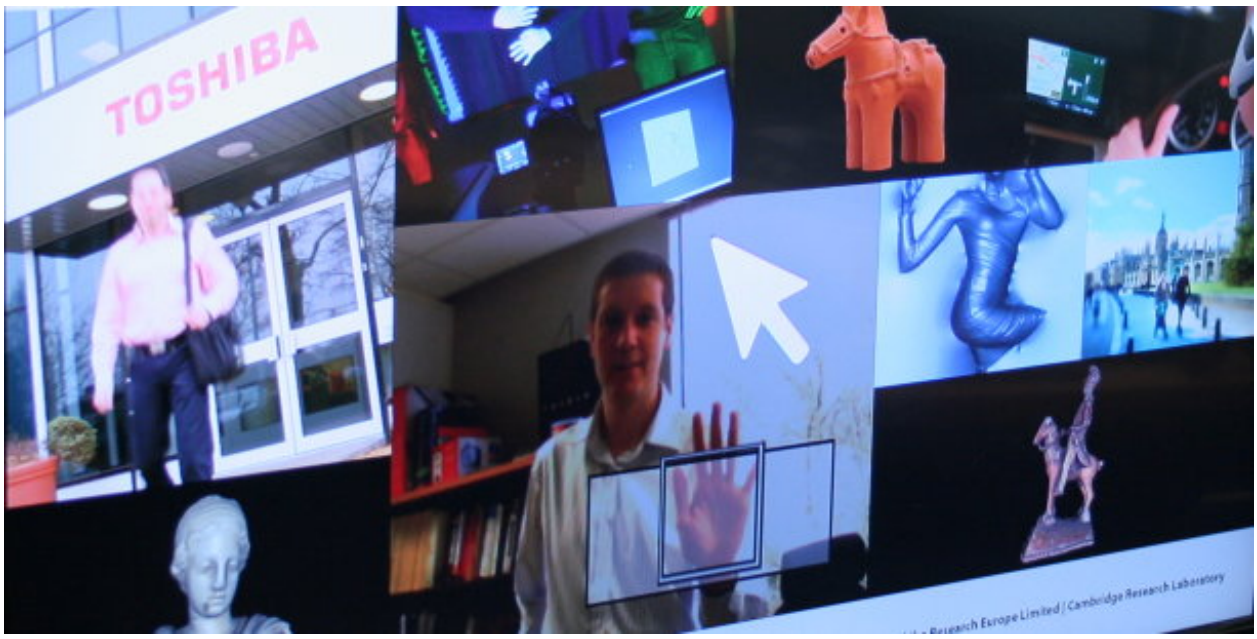


A closed fist is used to control the cursor

Stenger demonstrated how the system recognises that a person is in front of the screen by identifying their face, prompting them to take control of an on-screen cursor by placing their fist in the box (shown above).

Moving the fist moves the cursor to the desired location

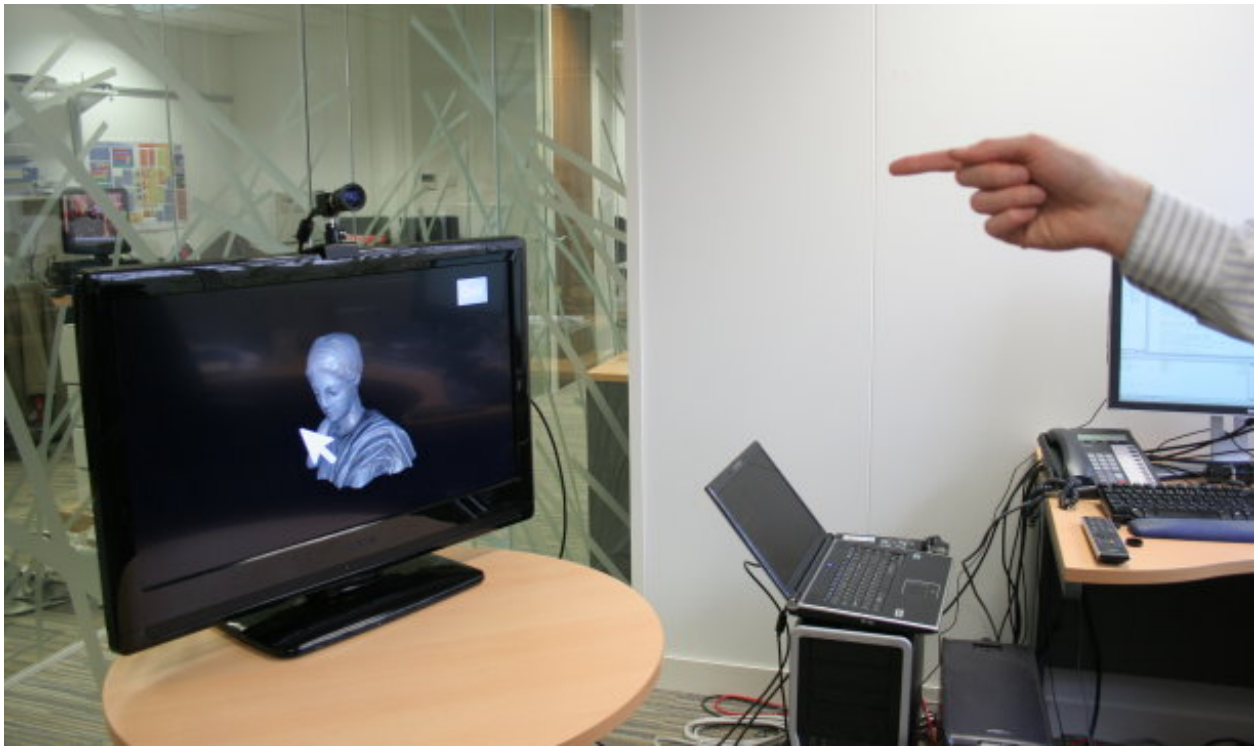
The user then moves their fist to make the cursor hover over a video to play the content. Videos can be stopped by showing the camera an open hand (shown below).



An open hand stops video playback

Photo credits: Natasha Lomas/silicon.com

The system also enables users to manipulate the viewing angle of 3D objects by pointing around the screen (shown above and below).



Pointing around the screen moves the viewing angle of the object

The gesture-based input system has also been put into Toshiba Cosmio laptops, using a built-in webcam.

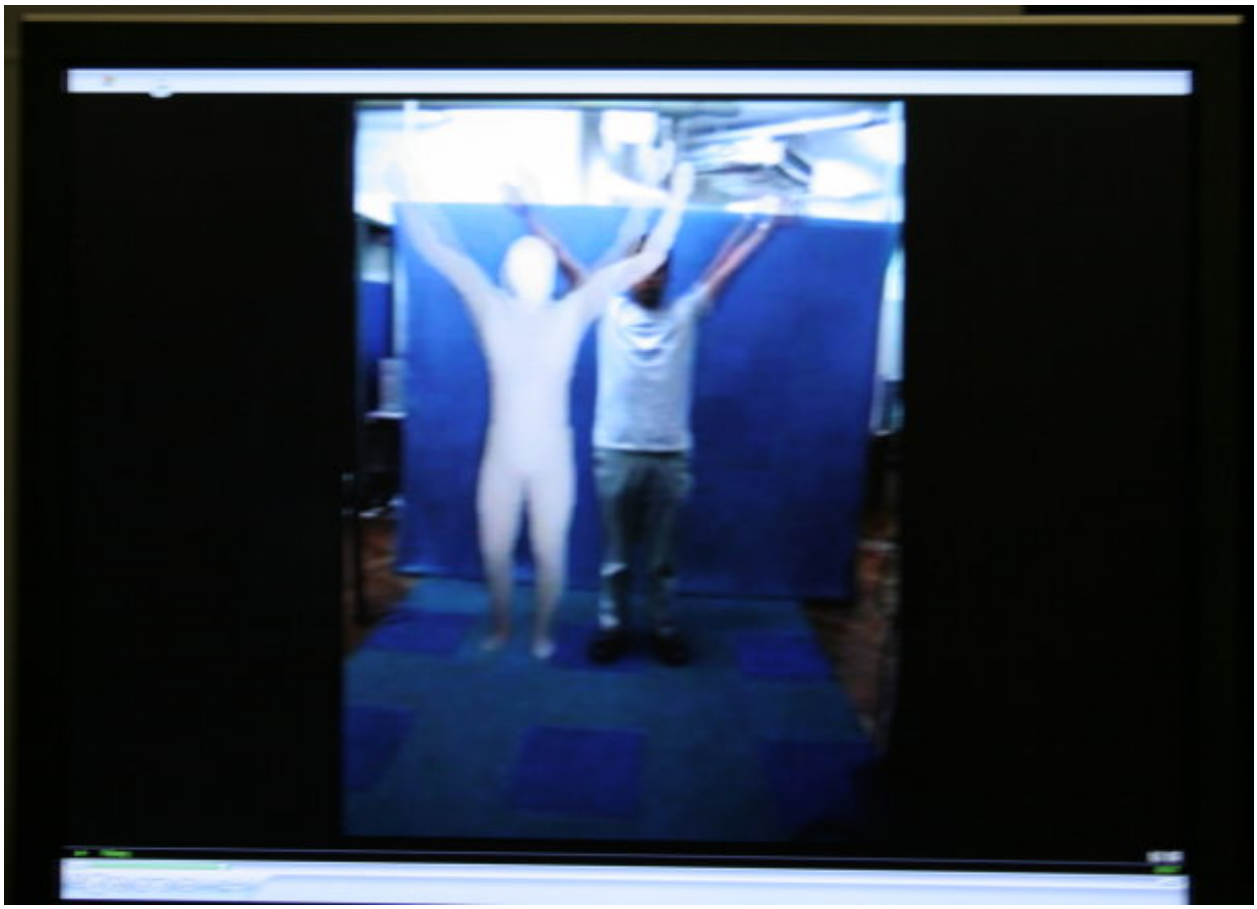
Photo credits: Natasha Lomas/silicon.com

Another gesture system developed by Toshiba research staff predates **Microsoft's as yet unreleased Natal** gaming input system by several years.

Developed in 2005, the real-time motion capture system pictured above used a standard camera but required a blue screen behind the user.

"What we're seeing now in the X-Box is this without the blue screen and using a special camera," says Cipolla, noting that **Microsoft's Natal gaming input system** uses a 3D camera to project a pattern over the gamer.

"They're using an invisible pattern, infrared pattern. So they get this nice clean data [segmenting the person from the background], once you've got this nice clean data you can estimate the pose in a number of ways."



A demo of a Toshiba gesture-based gaming system developed in 2005

Photo credits: Natasha Lomas/silicon.com

Check out some more photo galleries below:

Flying robots to driverless vehicles - the latest tech from the Ministry of Defence.

Searching images with the aid of **Google's Goggles**.

Nose scanning tech - is this the new face of biometrics?

Photo credit: MoD