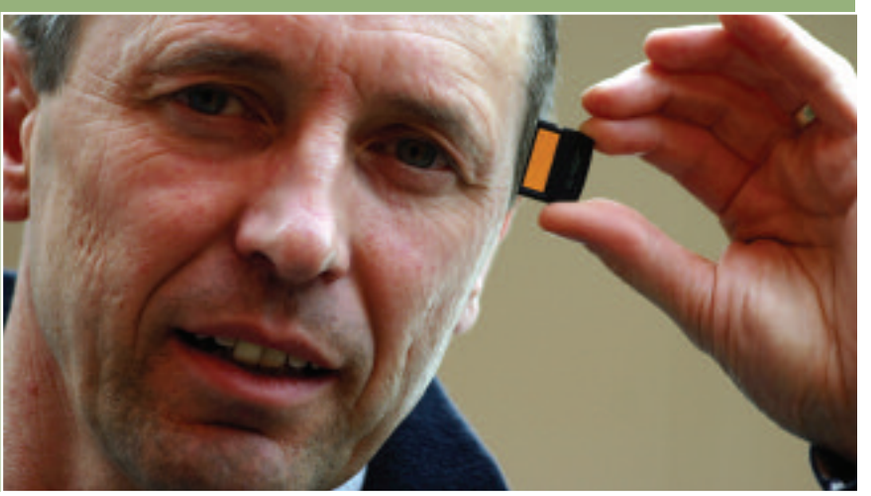
Navigating through the content of digital videos – the viewer's cut!

A unique cross-disciplinary collaboration is yielding exciting developments in the field of automated analysis of digital video information. **Cormac Sheridan** talks to Prof Alan Smeaton of DCU's Centre for Digital Video Processing about how his cutting-edge research is attracting multinational investment.



When one world-leading information technology company funds a university research lab, it's a reasonable indication that it is engaged in good-quality work. When two do so – and do so within the same month – it's more of a confirmation that a research effort of a significant order is underway.

Alan Smeaton, Professor of Computing and founding director of the Centre for Digital Video Processing (CDVP) at Dublin City University, received such a validation last February, when Google and Microsoft each funded different research projects within his 45-person research group.

"They're engaging with us because they think we're on a par with the best universities in the US," he says. "Irish research activity is now good enough to attract the beneficial interest of the large multinationals."

Established in 1999, the CDVP is a cross-disciplinary collaboration between the schools of computing and electronic engineering at DCU with a mission to develop techniques and tools for automated analysis of digital video information. It integrates a software-focused, computer science culture with a more hardware-oriented, engineering ethos.

"When you put those two together you create a very unique cocktail of people, who can both take new media and pull it apart to understand what's in it, and then also do useful things with that content," says Smeaton.

Neither of the current projects is a simple industry contract to perform a set of research and development tasks determined by the sponsor. Neither Google nor Microsoft is taking any proprietary position, in fact. Indeed, Microsoft has stipulated that it wants the outcome of the project it is supporting to be placed in the public domain. For its part, Google will maintain a watching brief on the project it is funding and may in-license the resulting technology at a future point.

Although the two projects involve radically different technical challenges, each is motivated by the massive expansion of digital information that has been enabled by the internet and by the widespread adoption of mobile electronic devices. Both address a common requirement – the need to navigate through very large quantities of material in order to obtain meaningful information.

In the case of the Google project, the material comprises handwritten text. Smeaton's group is adapting a technique it has developed for analysing digital video images to making handwritten documents amenable to word searches.

His group has developed a technology for segmenting individual elements in a shot, such as a person, and separating them from the background. This has potential application in the area of video compression. Only the dynamic elements in a shot need to be transmitted, while the background can be captured as a static image and transmitted just once.

"Integrated right into one of the base layers of this is an algorithm, a very clever algorithm,

which we've patented, to match the shape of something against the shape of something else," Smeaton says.

It ensures that deformations in shape that occur – for example, when a person moves through a space or when the camera undergoes a tilt, a pan or a zoom – can be tolerated. In the current project, it is being put to work in order to capture the subtle variations in shape and size that can occur in an individual's handwriting.

Currently, Smeaton's group is collaborating with teams at the University at Buffalo in New York State and at the University of Massachusetts at Amherst, both of which had taken a different approach to the problem.

"We don't try and identify the individual letters in a word. We just take the word as a blob, whereas they were trying to identify the individual constituents of a word." By combining the two techniques, the three groups hope to achieve a higher level of precision than would be feasible using either one on its own.

The CDVP secured funding from Microsoft Research through a competitive bidding process, following a request for proposals from the software vendor under a research programme called 'Digital Memories', which is intended to explore new applications for Microsoft's SenseCam, a sensor-activated, wearable digital camera.

SenseCam, which is still a prototype rather than a commercial product, generates a visual record of the wearer's day by taking several thousand pictures, each of which is triggered automatically by some kind of change, such as movement or alterations in light or temperature.

Smeaton's group is devising methods for identifying landmark or significant events over short or long timescales. The approach will integrate what Smeaton calls 'an ensemble of inputs'. These include elements such as colour, texture and shapes, which can be analysed using existing methods. His group is developing a technique for automatically determining the 'salience' or principal content of an image. Additional features, such as location-based information, and bluetooth signatures from mobile phones, which will provide personal profiles of people in the wearer's proximity, will be processed. His group has also built 'semantic feature detectors', which can differentiate between indoor and outdoor environments and between urban and rural or marine environments. By integrating all of these analytical methods, Smeaton's group aims to come up with a technique that will identify patterns and variations or exceptions to particular patterns that would constitute noteworthy highlights in an image log.

"When you put those two together [computing and electronic engineering], you create a very unique cocktail of people, who can both take new media and pull it apart to understand what's in it, and then also do useful things with that content."

DCU is the only university outside the USA to receive one of the seven Microsoft Research grants on offer. It's in good company. Other recipients include Massachusetts Institute of Technology, Princeton University and Columbia University.

Alan Smeaton attributes the current strength in IT research in Ireland to the Technology Foresight initiative, which gave rise to Science Foundation Ireland (SFI). Before the latter agency was established, there was around IEP 5 million available annually for university-based IT research, he says. The CDVP is now part of an SFI-funded group, the Adaptive Information Cluster, which also involves the National Centre for Sensor Research (NCSR) at DCU and several research teams based at University College Dublin. This has enabled Smeaton to enter collaborations with researchers in other fields, including, for example, Prof Dermot Diamond of the NCSR. It's a direction he is keen to follow further.

"My belief – and I've always maintained it – is that the best science and the most interesting science is done when you collaborate with people that you don't normally collaborate with, because it creates unusual applications that you wouldn't have thought of, if you're just focusing within. I see us collaborating more and more with these groups with whom we don't have a completely homogenous intersection."

Cormac Sheridan graduated from DCU in 1993 with an MA in Journalism. He currently works as a freelance science and technology journalist, contributing primarily to international publications.