

Auto Didactic

The in-house AV team in Social Sciences at Flinders University has developed an automated multi-camera vision capture system to record lectures. Finally, hardworking students get to sleep in.

Text:/ Andy Ciddor

ALTHOUGH LECTURE CAPTURE may sound like a very 21st-century idea, recording lectures for later replay by students has been the stock-in-trade of the AV departments in education institutions since shortly after the introduction of the compact audio cassette.

Generations of AV staff have been engaged in recording, editing, cataloguing, dubbing, loaning and duplicating the words (and, later, pictures) of generations of educators. For much of this time, the product has been substantially below what the community would expect to find in a commercial recording on the same medium, and even with advances in all of the technologies we use, this gap isn't going away any time soon.

LESSONS LEARNT

On the audio side of the equation, the digital signal processor has finally given us an affordable kit of processing tools to assist us with eliminating some of the more egregious problems that arise when attempting to capture intelligible audio in an acoustically inhospitable environment, from presenters with neither voice training, nor any particular interest in microphone technique or basic acoustic principles.

While in the video world, we may have been making recordings of lectures since the first U-matic machine tangled a tape around a head drum, things haven't really improved all that much. Granted that today's cameras are much more light sensitive, and that industrial video lenses and pickup arrays are now generally capable of resolving the writing on the board and sometimes even the skin colour and gender of the presenter. Some 40 years after it all began, the vast majority of lecture video recordings are still only unexciting static wide shots, even if they're now recorded by chip-based cameras on hard disk or flash RAM.

While it's widely accepted that multi-camera coverage of a teaching session captures substantially more useable information, and

consequently produces much better knowledge transfer than a static wide shot, the simple fact is that not even the best resourced of teaching institutions has the staff or the money to provide a fully switched multi-camera coverage for every lecture or presentation.

In the Faculty of Social Sciences at Flinders University in Adelaide, some lectures were being manually switched and recorded in 2003, but Matt Cooper, the Manager of the Audiovisual & Multimedia unit, was sure the processes of recording and distribution could be cost-effectively automated with available technologies. His interest in the possibility of an automated system was shared by Jonathan Wheare a Computing Systems officer in the same faculty, who believed that image switching automation could be achieved using available machine vision technology, running on standard PC hardware. To prove their point, in

“the inputs from a four-port PCI slot video input card are monitored for specific types of changes which are used to produce event triggers”

their own time, Cooper and Wheare developed a proof-of-concept prototype system, built entirely from spare parts available on-hand.

SWITCHING LOGIC

They analysed the existing manual vision switching processes to derive rule sets that could be used in an automated switching system. Using the same machine vision technologies employed for security, surveillance and process monitoring, the inputs from a four-port PCI slot video input card are monitored for specific types of changes which are used to produce event triggers.

Each input has its own set of specific content and trigger rules. On data feeds, an absence of sync information or a blank screen will produce a 'no picture' message for the switching decision logic. If there is a change in the image from a data feed or document camera, the change must meet defined threshold requirements before producing a 'changed image' message. This is required to prevent a moving mouse pointer or picture noise being mistaken for new data. Defined areas of camera images may be identified as areas to be ignored. These areas are set up to avoid false triggering from the random movements of students dozing off or throwing paper darts or such artefacts as reflections from specular objects in the frame.

The vision switching logic is provided with a list of event priorities, hold times and timeouts that produce the best possible representation of the presenter, computer data feeds, replay vision and document cameras, without random flash cuts or dead air. The wide shot camera is the default output, with other, changing sources, taking priority in a defined hierarchy, until they haven't changed for a specified hold period.

Similarly the audio inputs are gated, compressed, limited and gain controlled via a DSP system to produce the best possible mix of presenter's voice and the audio outputs of any replay devices used in the presentation.

STORE & FORWARD

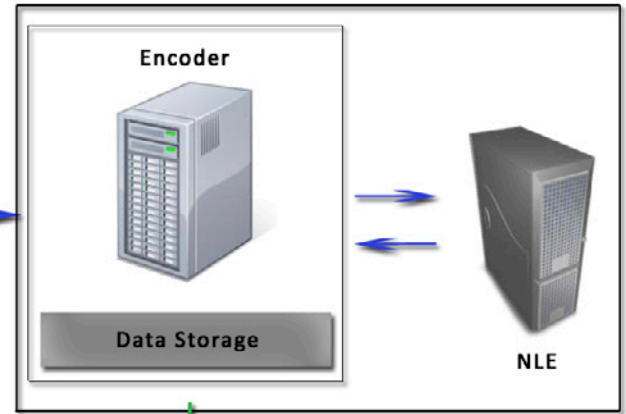
System output is stored locally in DV format on the hard disk of the capture computer and uploaded to a media processing server as bandwidth becomes available. The 'store and forward' approach enables the recording process to complete successfully independently of the network backbone, and places lower peak demands on backbone bandwidth.

In the media processing server, lecture recordings are transcoded into the requested video and audio formats, typically Flash for steaming video and MP3 for the audio content.

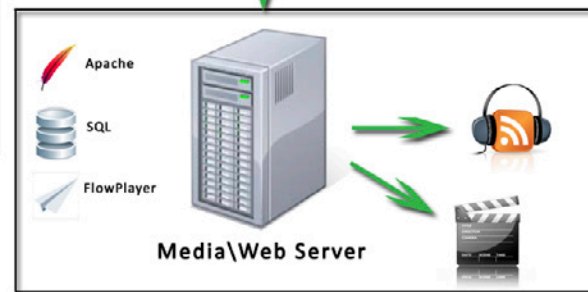
CAPTURE



PROCESSING



DELIVERY



LECTURE CAPTURE WORKFLOW

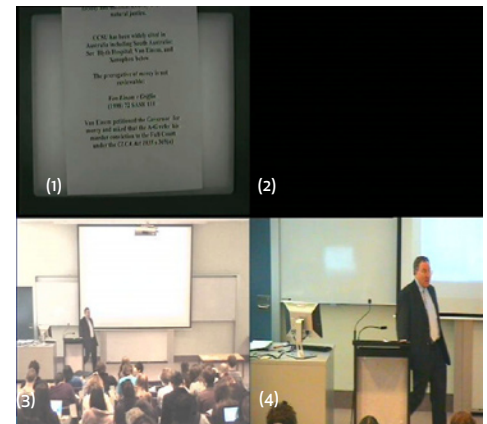
In the teaching area inputs are intelligently switched to produce a single image stream that is stored locally on the Automated Vision Capture System. The raw stream is transported across the network to a central processing system for encoding and any manual editing that may be required (usually just the removal of pre and post lecture shuffling about). The processed audio and video are then uploaded to the media server for viewing via the web interface.



MCMS index

Job ID	Added / Rec Time	SLA	Client FAN	Topic Code	Details	Elements	Actions
L1172	09:00 Mon 03/08/2009	OK	mode0144	BUSN1009	North 3 Video / Link / C1_91	[Icons]	Review
L1256	10:00 Mon 03/08/2009	OK				[Icons]	Review
L1398	11:00 Mon 03/08/2009	OK				[Icons]	Review
L1568	11:00 Mon 03/08/2009	OK				[Icons]	Review
L1007	14:00 Mon 03/08/2009	OK				[Icons]	Review
L1714	14:00 Mon 03/08/2009	OK				[Icons]	Review
L977	16:00 Mon 03/08/2009	OK				[Icons]	Review
L1200	09:00 Tue 04/08/2009	OK				[Icons]	Review
L1574	09:00 Tue 04/08/2009	OK				[Icons]	Review
L1022	10:00 Tue 04/08/2009	OK				[Icons]	Review
L1370	11:00 Tue 04/08/2009	OK				[Icons]	Review
L1462	11:00 Tue 04/08/2009	OK				[Icons]	Review

SLA	Status	Media Types
OK	Captured	Audio CD, Portable MP3, High Quality MP3, Low Quality MP3, Add to FLO MP3, Add to FLO MP3
Warning	Uploaded	Video DVD, Large powerpoint WMV, Small powerpoint WMV, Add to FLO MP4, Add to FLO MP4, Multimedia DVD
Critical	Editing	Edited, Encoding, Encoded



WHAT THE AV STAFF SEE

Above: Remote monitoring quad split of the vision sources for the session. 1) Document camera. 2) Computer output. 3) Fixed wide shot camera. 4) Moveable (PTZ) camera. Left: Content management screen that tracks and routes a lecture session from booking the recording session to publishing the finished media on the appropriate web server page.

BUSN1009 - Quantitative Methods

1. Mon, Jul 27 2009	2. Wed, Jul 29 2009	3. Mon, Aug 03 2009
Click to Play	Click to Play	Click to Play
Presenter: Prof Daniel Leonard Theatre: North 3 Video Link	Presenter: Prof Daniel Leonard Theatre: North 3 Video Link	Presenter: Prof Daniel Leonard Theatre: North 3 Video Link
Duration: 00:45:26 min	Duration: 00:46:40 min	Duration: 00:41:43 min
Watch Video (116MB)	Watch Video (119MB)	Watch Video (119MB)
Download MP3 (7MB)	Download MP3 (7MB)	Download MP3 (6MB)

NEXT



WHAT THE STUDENTS SEE

Left: The media resources selection screen for a lecture series.

Above: The familiar Flash media player interface that pops up from the selection screen.

3 Lectures available for this topic.

The files are then logged in to the content management database and transferred into the content servers for WebCT.

Automated lecture capture, however, is only one small part of the system, which provides a full cradle-to-grave automated management system. It begins when a lecture recording session is booked on a web page and ends when that lecture recording has been converted to the appropriate media formats and made available for replay on the selected WebCT pages.

BROADER APPEAL

In late 2004, while algorithms, thresholds and priorities were being refined, the prototype automated system was run in parallel to the existing manual switching system in Social Sciences Lecture Theatre 3.

With the prototype system capable of full operation, in 2005 it was examined

by a number of groups with the university, including the external commercialisation arm, which recommended the university assess the patentability of the concept while looking at introducing it across the wider campus. Unfortunately, neither of these recommendations has yet been implemented.

However The Faculty of Social Sciences picked up the system, and has since installed lecture capture installations in four lecture theatres and set up a portable capture system that can move around between a range of smaller teaching spaces that have the appropriate video and audio facilities.

The capabilities of the system continue to be extended, with capture servers now having the capacity for eight video inputs and beyond. With contributions from faculty AV staff Bean Keane and Melchior Mazzone, a content management system has been added to the back end to allow web-based online

selection and distribution of recorded material to a range of distribution and viewing pages.

A monitoring and media tracking package is the most recent addition to the system. This allows a single technician to remotely monitor the status and operation of multiple capture systems and all steps in the processing path. With its quad-split video monitoring of each system's inputs, it enables direct intervention in recording and processing operations, right down to the level of modifying machine vision settings, overriding switching decisions and tweaking the framing and focus on the PTZ cameras.

The complete automated system has now been in full production for over three years, capturing and processing around 70 lectures per teaching week and serving up some 50,000 web sessions per semester; a feat of unlikely productivity for such a small AV unit. ↵

SOON LIGHTING CONTROL WILL BE CLEAR CUT

