Effective Management of Digital Video Assets:

The CatDV Pro Advantage

Rolf Howarth

Square Box Systems Ltd.

March 2003

Over the last decade or so we have witnessed a revolution in digital media management and content production. First came non-linear edit applications, replacing slow and inflexible tape-based edit suites with disk-based editing and digital video effects. Then satellite, cable and digital revolutionised television – with vastly more channels, there came a huge demand from broadcasters for low cost program material. Cheap digital camcorders and personal computer-based editing software can meet some of this need at the low end of the market, achieving a quality that professional broadcasters would have been more than happy with a few years ago with equipment that is now available to consumers.

Bandwidth, disk space, and computer processing power are getting cheaper all the time, but new applications to make use of this capability have kept up proportionately, or even made us greedy for ever more as our eyes are opened up to the possibilities. The huge popularity of MP3 audio files and digital cameras have given consumers something to do with their increased disk capacity. Video has even greater demands, but whereas previously a broadcaster might have been content with a basement containing dusty shelves stacked with reels of film and videotape, perhaps with a few handwritten labels to give a clue as to their contents, now we expect such content to be instantly accessible.

Cataloging

Just like managing a library of books, archiving a large collection of video assets has two main aspects. First, the media itself needs to be available and stored somewhere safe, most likely on magnetic tape and in a digital format (DigiBeta, DV, etc), though this obviously depends on the original source of the material. And secondly, we need some sort of index or catalog of the collection so we can retrieve the library's contents.

The purpose of building up such a catalog is so we can find clips without having to review the entire collection, which could contain thousands of hours of material. Just like a book has a catalog number and pages, each tape will need a unique name or number identifier, and individual frames on the tape will be identified by a linear increasing timecode value. We can then store information *about* the collection in our catalog without having to store the media itself – we will still know unambiguously what we're referring to, e.g. "shot xyz starts at $35^{1}/_{2}$ minutes into tape 93".

The general term for data about other data is 'metadata'. An obvious approach to cataloging our library is to store various attributes describing each tape or video clip (our metadata) in a database and use these for browsing and searching the catalog. We can use textual attributes to describe the clip, such as a description, the project name, tape name, camera operator, the shot location, various keywords describing the subject, and so on, and perhaps even a verbatim transcript of any spoken dialog. We could store dates, such as the date of recording, date of capture, or the date the final program was edited. We can store technical attributes of the clip, including video format, audio sample rate, capture filename and size, timecode values, and so on. Finally, recognising that best way to get a quick overview of a large amount of information is visually, we might be tempted to store a still thumbnail image for each scene, or even a low-resolution video proxy to remind us about the clip's contents in the absence of the full-size original.

All of these metadata attributes are potentially very useful, but the effort required to manually build up a detailed catalog containing all this information may be prohibitive. To be effective, therefore, it's

vital that as much as possible of this information is collected automatically from all the available sources during the cataloging process.

Video logging

Video production has several distinct steps. The steps up to and including the shoot itself (writing a script, hiring the talent, planning a shoot, lighting, directing, sound recording, etc.) occur before the media is created and stored in digital form and do not directly concern us here. Once the original footage has been recorded, however, and before it can be edited into a program, it normally needs to be 'logged'.

Logging involves watching the material and making notes of which shots to use. Traditionally, this might have been done by a producer watching a VHS dub with burnt in timecode and making pencil and paper notes of timecode values of scenes of interest. This list would be passed to an editor, who would digitise the requested segments, trim them roughly, and lay them out in a sequence, then once the timing is correct any required transitions and titles or effects can be applied. Another common approach is to log straight from the master tape, cueing and reviewing to mark in and out points, then batch capturing all the selected clips.

In either case, because digital video at full quality requires so much hard disk space, users are accustomed to capturing or digitising just as much as they'll need in the final program, and then deleting the material again at the end of a project. Batch capture can then be used to automatically recapture the required material. Given a unique tape name and timecode that can be read, non-linear editing programs will prompt the user to insert the necessary tape, then remotely control the tape deck or camcorder to play the requested portions of tape so they can be recaptured. Because each frame has a unique timecode value it's possible to recapture the same material with single frame accuracy, giving precise repeatability.

The process of logging, especially accurately finding the start and end of each shot, is time consuming and mundane. Logging by shuttling a tape backwards and forwards also has the disadvantage of wearing out both the heads of the deck and the precious master tape, so any alternatives to simplify and automate the logging process would be very welcome to most editors.

The DV format

In 1995 a consortium of manufacturers led by Sony introduced the first digital camcorders using the DV format. DV was primarily intended as a consumer format but professional variants of the format exist too, and because of the low cost and compact size of cameras it is also gaining a lot of popularity in areas such as news gathering and documentary production.

DV is a digital format, giving a near broadcast quality, but unlike MPEG there is no interframe compression, making it highly suitable for editing. Many personal computers include an IEEE1394 FireWire interface, allowing DV video to be captured and edited with no generational loss in quality using readily available hardware.

As well as its low cost and ubiquity, the DV format has several features that make it interesting from the standpoint of logging and cataloging video. Apart from the video and audio sample data itself, each frame can include additional metadata embedded within it, including the timecode value of the frame, the date and time of recording, the camera exposure and shutter settings, the aspect ratio, and a marker that identifies when the operator started or stopped recording.

All of this additional metadata stored with the video data is very helpful when it comes to automating the process of logging clips and building up a catalog, though not all these features will be supported by all video cameras or in all circumstances. For example, some professional cameras record an additional auxiliary timecode value that can be set from an external source, and the date will only be valid if it was correctly set on the camera at the time of recording.

CatDV Pro

CatDV Pro is a media cataloging application that is designed specifically to deal with video and take the above points into account. There are of course many superficially similar applications that provide a catalog of images or media files, but these tend to operate at the level of *files*, identified by file name, whereas CatDV Pro deals with video *clips*, identified by a tape name and timecode value that can be used to relate back to the original master recording.

A stack of master tapes in a box or on a shelf is quite inscrutable, so CatDV Pro uses the following information to describe and represent a segment of video in a catalog:

- clip name, description, and an unlimited number of other user-defined textual attributes (eg. location, videographer, project, scene/take);
- one or more still thumbnail images, typically from the first frame or midpoint of each clip;
- tape identifier and timecode In and Out values;
- technical attributes such as duration, the format the video and audio is stored in, the image aspect ratio, audio sample rate, file size, data rate, and so on;
- the original media itself, assuming it is currently online (i.e. available in digitised form on disk). Alternatively, a low resolution preview version of the media may be created and stored locally and used as a proxy in place of the full size original media.

CatDV Pro deals with different types of clip object that can come from different sources. For example, clips can come from a batch log or project file exported by another application, be entered by the user, created automatically by scanning a DV tape, or come from importing and analysing a media file stored on disk (this could be a captured movie, or it might be any other kind of media file such as a still image, audio clip, or computer generated animation). Not all the attributes listed above will be present for all types of clip. For example, if the media has never been captured to disk then media file name and file size have no meaning, and an audio clip has no thumbnail.

Automatic logging

When a movie file is imported into CatDV Pro it is automatically analysed to see if it contains any scene boundaries, and if so a separate clip record is created for each scene. Automatic scene change detection is a useful first step in the logging process as it removes one of the main time-consuming activities.

Two algorithms are available for detecting scene changes:

- If the media was originally recorded in DV format the video camera will have written a marker whenever the camera started or stopped recording. If CatDV Pro detects either this marker or a discontinuity in the date/time stamp recorded with each frame, then it knows a scene change has occurred with total frame accuracy.
- Alternatively, it is possible to apply a visual scene detection algorithm that compares the brightness and contrast of successive frames looking for significant changes that would indicate the start of a new shot. The sensitivity threshold can be adjusted but by its nature this algorithm won't always be 100% accurate. It can be fooled into thinking there is a scene change if a shot contains rapid motion, or might miss a change if the camera is on a tripod and there is very little change between shots. This method works with any kind of video material, however, not just material shot in DV format.

The scene boundaries detected by CatDV Pro can be reviewed and adjusted manually, either merging together two clips that were split in error, splitting a clip in two if a scene change was missed, or creating new secondary clips from a selection within a clip.

Each clip can be annotated with a name, description and other attributes which can be used for searching and categorising the clips. Clips can be grouped into 'bins', flagged as 'good' for

subsequent capture, and assembled into rough cut sequences. If the source material includes a lot of spoken dialog then the verbatim logger can be used to type in free format text as the video plays, with hot keys to pause or back up the movie slightly or to insert a marker for the current timecode value.

Once clips have been logged in CatDV Pro they can be used in various ways. They can be printed out, either with or without thumbnails, copied to another application, or exported as a batch capture log or edit decision list (EDL) for subsequent editing in an online editing application.

This is the traditional functionality provided by a logging utility, and pure logging functionality might be all that is required if a program needs to be edited together as quickly as possible after the shoot. If the results of the logging process are saved in a database, however, then it becomes a useful resource in the long term should it ever become necessary to refer back to existing material. Keeping the logging results in a database is also useful on large projects as a way to organise and manage the material during production.

Logging and archiving are normally considered to be two completely separate processes but they are integrated within CatDV Pro because it is logging that typically provides the raw information that will be used in a catalog.

A media asset database

By storing all the information collected during the logging process in a database we build up a permanent index of the contents of our tapes. The material on each tape is described by one or more clip records, each with a timecode start value and a duration. Each clip has a name, notes, and other user-defined text fields, as well as technical details such as record date and format, and a thumbnail. The point to note is that this information is available in our catalog even if the original media is no longer available online, for example if the digitised file has been deleted from hard disk and the master tape is locked away somewhere in a vault.

If a new project is commenced that requires material from the archive then it is a simple matter to search across all the existing catalog(s) for matching clips, based on type, date, containing a particular keyword in the notes, and so on. Complex queries can be created and saved for future use, for example “find all clips in 16:9 format that contain the keyword ‘New York’ and were shot on or after 1 Jan 2002, plus any other shots on the same tape taken on the same day (even if they don’t include that keyword)”. All the matching clips are displayed, complete with their thumbnails and other data, and can even be previewed if low resolution proxies were created when the tape was originally scanned. Those clips which are suitable for inclusion in the new project can be marked, and sub-selections made within these clips by marking in and out points if necessary, then exported as a batch log or EDL for use in the editing station.

As well as providing a search capability, CatDV Pro has a flexible user interface that allows users to browse the catalog in different ways. Clips may be shown in tabular form or as a grid, with thumbnails of different size and showing a user-definable list of attributes. Clips can be sorted, filtered and categorised in various ways, and any media can be presented in windowed or full screen display. If preview versions of all the tapes are available, for example, you can import an arbitrary EDL and instantly display a cuts-only version of the program.

In addition to cataloging media files, other types of file that describe the contents of a tape, or indicate which projects a clip was used in, can be imported into a CatDV Pro catalog too. These include edit decision lists, project files, and arbitrary tab-separated text files. Combining information from different sources in this way might result in duplicated or overlapping clip details, so CatDV Pro provides a special summary view that automatically merges details from multiple clips together to provide a concise overview of the contents of a tape. This can be printed out as a single page index sheet, much like the contact sheet returned when a 35mm or APS film is processed, and filed with the tape itself.

CatDV Pro catalog files themselves are very small, tiny in comparison to the original media, as all they contain are textual information about each clip and the thumbnails. They can thus be kept as a permanent record of a tape or project that was worked on.

Supported media formats

CatDV Pro uses QuickTime to provide its core media capability. This means that media files in any format supported by QuickTime can be imported and indexed in a CatDV Pro catalog. This includes not just QuickTime and DV movies (and DVCPro/DVCAM) but many other common video formats such as AVI, MPEG1, MPEG2, and more recently MPEG4. In addition, formats other than movies can be catalogued, including still images (TIFF, PSD, JPEG, GIF, BMP etc.), audio clips (AIFF, AU, WAV, MP3, etc.) and various other more specialised file formats (QTVR, Flash, PDF, sprite movies, etc.)

When a media file is imported into CatDV Pro, technical metadata describing the format of the file is extracted and added to the catalog, as well as thumbnails (if the file has any visual content), and any 'user' metadata. For example, a QuickTime movie may contain title, author and a copyright notice; JPEG, GIF and Photoshop images may contain comments; and an MP3 file may contain artist and song title. Modern digital cameras and image editing applications store information about the image (date, aperture, exposure, focal length, etc.) in Exif metadata tags and these are also extracted and catalogued. As a project might involve other types of media file than just video it makes sense to be able to catalog them all in the same database.

As well as the file formats supported by QuickTime, CatDV Pro also has built-in support for reading a number of proprietary file formats, including Avid OMF Interchange media files, Pinnacle MIF files, Cinestream project files, Final Cut Pro batch files, and so on. This support is provided via an extensible framework, so additional specialised formats can be added as required.

Finally, in the case of DV files, CatDV Pro has special support to read the metadata embedded in each frame: timecode, date/time of recording, aspect ratio, audio format, auxiliary timecode, and camera aperture, gain and white balance settings. These are recorded in the catalog and can be used for searching, sorting and filtering. The record date in particular is very useful as an automatically defined attribute to help us search for and group video clips in a media database.

Client-server architecture

CatDV Pro can be used as a standalone desktop application, with clip catalogs held in self-contained documents on the local file system. When searching for clips matching a query the search can be performed within the catalog that is currently open in a window, or a limited search can be performed across all the catalog files in a given directory.

To permit much more sophisticated searching and effective collaborative working within a team, however, CatDV Pro can be used with an optional workgroup server. In this scenario, clips are stored in a central database (rather than a flat file) and the CatDV Pro client application communicates with an application server program that runs on a separate server machine. This server performs queries and updates against the database and returns the results over the local area network back to the client.

Using a relational database means that complex queries across a large database can be performed quickly and efficiently. Result sets consisting of clips from different individual catalogs can be returned and used exactly as if they were stored in a catalog file locally. Once clips have been returned to the client they can be edited, perhaps adding further keywords, creating new secondary clips, or marking clips of particular interest for other team members to look at. The changes can then be published back to the server. If two users attempt to update the same clips concurrently then the second user is given the opportunity to resolve any conflicts that may occur.

Users can work locally (disconnected from the server, perhaps using a laptop to log in the field), saving their catalogs to the local file system. They can publish these catalogs to the server once they

are connected again, so the clips then become available to other users. Conversely, they can download catalogs or the results of performing a query from the server and save these locally.

Just as when saving catalog files locally, the database only holds the textual and numeric attributes that make up a clip record and are used for searching, together with the thumbnail poster images. It is relatively small in size therefore. The preview movies and any original source media files are not stored in the database. Instead, they are stored as normal files and can be accessed via the file system. They can be stored on a file server (either the same machine as the workgroup server or a dedicated media file server) and accessed as a remote volume mounted on each client machine, or they can be accessed via a web server. The CatDV Pro application simply deals with path names or URLs to the media, providing great flexibility in how the system is configured, depending on the requirements.

Rather than using the CatDV Pro client application to access the database, it is also possible to use an additional web publisher server option to publish the contents of the CatDV Pro database via http and HTML. This allows anyone with a web browser to browse or search the database and display or print reports of matching clips, though not to update the database or perform logging (or other media intensive operations, such as analysing movies or displaying slide shows). It is possible, however, to customise the web interface and integrate it into other systems as part of a broader workflow.

Open standards

CatDV Pro is based on non-proprietary open standards from the outset. In an effort to keep costs down, and avoid being locked in to one particular supplier, several states are now mandating that public bodies actively consider open standards-based proposals in the tendering process. We are a firm supporter of this principle.

Although QuickTime is a proprietary Apple technology, it is available for both Macintoshes and Windows-based PCs, and is widely supported by different application vendors. The format itself is published and was used as the basis of the MPEG4 open standard. QuickTime is a framework that allows third party vendors to develop plug-in codecs to support particular file formats or hardware. Any such third party features are then fully available to CatDV Pro through the QuickTime API.

CatDV Pro is developed in the Java programming language, a platform-neutral language and run-time environment. This means the client application is available both for Apple Macintosh and Microsoft Windows, and means that the server can run on a very wide variety of Unix and Windows platforms, including Mac OS X.

The workgroup server is based on Java J2EE technologies. By default it uses the open source MySQL database and Apache Tomcat servlet engine. By adhering to open standards, however, it is possible to reconfigure CatDV Pro to use other databases or application servers, providing great flexibility and scalability, and compatibility with many typical enterprise infrastructures.

Conclusion

This is an exciting time, as more and more media assets are processed in digital form and become accessible to the same kind of productivity and workflow improvements that other more mundane forms of data have been subject to for some time.

The problem when dealing with any large collection of media assets is how to condense the sheer mass of data involved down to a manageable form. To that end, it is important to extract any existing metadata that may help us, and leverage what is already there (for example, in the DV format) to the maximum extent possible.

CatDV Pro does this, and as a flexible, low cost, standards-based tool is finding increasing acceptance as the media cataloging solution of choice, for users at all levels from ‘prosumers’ through to independent video producers, advertising agencies, and broadcasters.