A Universal Mastering Format for Post-Production Finishing: 1080p/24

By Patrick Byrne

This paper addresses the need for a universal mastering format as a result of international acceptance of DTV broadcasting and the ensuing demand for content in various media. With the FCC adoption of the ATSC standard and growing acceptance of DTV internationally, dramatic changes will be seen in the post-production of programming content. Nonlinear disk-based finishing systems offer a flexible solution. Market needs, system requirements, and several possible post-production workflows are examined.

The Federal Communications Commission adoption of the ATSC Digital Television Standard (A/53) in 1996, which defined 18 DTV formats, set the groundwork for dramatic change in the broadcast and post-production industry. Three major U.S. networks announced their plans to support different DTV formats two years ago, and momentum continues to build for digital television broadcast. As of October 1999, there were 75 DTV stations on the air, and by the year 2002, that number is expected to rise to 39% of all stations in the U.S.

There is growing acceptance of DTV internationally. Various countries throughout the world are planning the transition to digital, and announcing objective dates for ceasing analog transmission. In 1997, Canada and Japan adopted digital terrestrial broadcasting using the ATSC DTV standard. In 1998, Taiwan and Argentina did so.

The Argentine decision marks the first time that a PAL-based country has adopted ATSC for its digital transmissions. Argentina chose the ATSC standard over the European DVB-T standard, which it also evaluated. Recent reports indicate that Australian post-production professionals, who will begin digital broadcasting in 2001, support the acceptance of 1080p/24 as a universal mastering format.

The availability of a new 6-MHz digital channel offers options to broadcasters: transmit a single high-resolution signal or use the bandwidth to broadcast up to six standard-resolution programs simultaneously. In February 1999, 43% of stations surveyed said that they plan to multicast different programs with the new digital signal. This decision increases pressure on program producers to create new content or reformat existing content, to fill this additional airtime.

The Universal Mastering Format

The need to produce new programming content in a variety of distinct DTV formats, as well as NTSC and PAL, during the transition period to digital, has furthered the call for a universal mastering format. Universal mastering enables programming content to be produced and post-produced in one format from which all of the necessary delivery formats may be derived at the end of the production pipeline. They include DTV, film, NTSC, PAL, DVD, or web-based. The universal mastering format, one of 18 standards from ATSC Table 3, is 1080p/24. That is, a resolution of 1920 active pixels, by 1080 active lines, progressively scanned at 24 frames per sec.

Combining the best attributes of the ATSC standard, 1080p/24 offers the highest possible spatial resolution, a common frame rate to standard film production, and a wide aspect ratio. It maintains the same properties of film acquisition and distribution—since film scanning and recording devices are also progressive in nature. The high-resolution nature of 1080p/24 also enables output to 35mm film for theatrical release.

Benefits

By using the highest possible spatial resolution from Table 3 (1920 x 1080), more image detail is preserved from the original source. These large frame sizes, progressively scanned, provide the viewer with an unsurpassed visual experience, while at the same time enabling easy, high-quality conversions to smaller frame size formats. The inverse is not always true: up-scaling ("up-rezing") standard resolution imagery to high resolution often results in loss of detail, and may potentially create image artifacts. The 1920 x 1080 frame size is approximately six times the size of today's ITU-R 601 image, often requiring advanced hardware processing to manage increased amounts of data.

The 16 x 9 aspect ratio of 1080p/24 offers a widescreen cinematic experience to viewers who may be accustomed to the standard 4 x 3 aspect ratio. The wider frame creates a greater sense of space from a creative standpoint, but converting between formats is challenging. As cinematographers, directors, producers, and artists educate themselves about this new process, attention must be paid to the compositional elements and their position within the frame, knowing that the images must be converted from 16 x 9 to 4 x 3. One solution is to provide pan/scan capability within the universal mastering system, although this does not necessarily alleviate the creative burden caused

Presented at the 141st SMPTE Technical Conference (paper no. 141-S8), in New York City, November 19-22, 1999. Patrick Byrne is with Discreet, Montreal, Que., H3C 2L7, Canada. An unedited version of this paper appears in *Sprockets, Samples, and Satellites: Moving Imaging into the Third Millennium*, SMPTE, 1999.

by supporting two separate DTV aspect ratios.

1080p/24 workflow is somewhat different from standard resolution post-production techniques, but is similar to the familiar workflow for film. One telecine transfer may be performed at 1920 x 1080, with all the distribution formats generated from one pass. This may be contrasted to previous workflow, where delivery in additional formats required an external standards converter that could reduce the quality of results, or where the entire project was re-scanned from the original film. This offers exciting new possibilities.

Since approximately 85% of today's episodic television is shot on film, working at the 24-frame rate in post-production simplifies matchback to the original source material. Also, disk-based editing and visual effects post-production systems offer a 20% savings in disk storage, as only the original 24 frames of material is saved to disk versus 30 frames for NTSC. Furthermore, fewer frames are processed during image compositing and image manipulation, resulting in significant time savings.

The capture, storage, and processing of progressive images eliminate the need for filtering and sampling common to interlaced systems. Maintaining a progressive frame throughout the finishing pipeline results in fewer interlace artifacts and creates higher quality composited images. Progressive systems alleviate having to manually maintain a consistent 3:2 pulldown sequence, a basic requirement for MPEG-2 encoding. Rather, a consistent sequence can be added automatically during output from the system.

Requirements for a 1080P/24 Universal Mastering System

Several requirements must be met in order for a system to deliver a complete working model at 1080p/24. First, it must be format independent, that is, capable of supporting multiple DTV standards with various I/O resolutions, as well as NTSC and PAL. In order to be cost-effective in today's economic environment, the system must support HDTV resolutions while maintaining all of the same functionality currently available at ITU-R 601 resolution. This will not only ensure a secure migration path into the future, but will produce the best quality distribution copies from a high-resolution 1080p/24 universal master.

Second, it must be frame-rate independent. This implies that the system is capable of capturing, displaying, and processing the various frame rates that compose the ATSC digital television standard.

Third, the system must be compatible with current facility infrastructure and be capable of integrating with current I/O devices—whether 601 or HDTV resolution. It must also be able to interface with today's post environment from previsualization, offline, finishing, and audio post.

Fourth, the high spatial resolution of 1080p images requires that artists interact with these pictures in their purest form—thus the system must support uncompressed images at full resolution during display and processing. The hardware requirements for CPU power, graphics processing, memory allocation, and storage access are greater when working with large frame sizes, so hardware technology advances have offered high bandwidth bus and memory architecture to deal with the additional data efficiently.

Additional requirements include video processing tools for progressive to interlace conversion; audio tools for sample rate changes and pitch correction; and I/O file utilities, such as edit decision list and telecine log file support. The system must be flexible and upgradable, to deliver primary features as the market requires and secondary features in future software release cycles.

Workflow

Following is a description of the requirements for capture, processing, display, and output of 1080p/24 images, and the creation of a true universal master. For the purposes of this paper, the referenced solutions are the Discreet fire* and inferno* systems, operating on the Silicon Graphics Onyx 2 platform. It is important to note that the current software version also supports all of the same universal mastering functionality at 601 resolution on Discreet fire*, inferno*, smoke*, and flame*. Provided that several workflow criteria are met, universal mastering can be performed by fire* and inferno* at full resolution 1080p/24.

Several leading equipment manufacturers are in the process of adding 24p support to their devices, including cameras, DDRs, and VTRs. However, until this equipment is readily available and the entire post-production process can occur at 24p, Discreet has developed a workflow solution that facilities can adopt, which utilizes existing equipment and infrastructure.

First, the source material is shot on film at 24 frames/sec, or on 24p HDTV video. Second, the film-originated source material is transferred to a HDTV VTR as 1080i/60. The video transfer reels have continuous ascending time code, continuous 3:2 cadence, and a known SMPTE A-frame start.

The Discreet system captures uncompressed 1080i/60 images in realtime using the SGI XT-HD hardware option. The XT-HD board, capable of input or output of uncompressed high-resolution images in realtime, is based on the XIO bus interface available on the SGI Onyx 2 platform. The XT-HD board has a parallel input and requires the use of third-party serial/parallel converters prior to capture. The captured images are stored as RGB 4:4:4 to ensure high quality during image processing. Therefore, a realtime color space conversion is applied at the board input and output stage.

Since the current universal mastering workflow is based on 1080i/60 as input, some processing is required to convert to the original 24-frame source material. This is accomplished by automatically removing the 3:2 pulldown sequence from the 1080i/60 source. 3:2 pulldown is the industry standard formula for deriving 60 interlaced video fields from 24 film frames. For the purposes of this paper, 3:2 pulldown is analogous to 2:3 pulldown. It is a repeating 10-field sequence, where an image is repeated every 5 fields (Fig. 1).

Discreet finishing systems automatically remove 3:2 pulldown while capturing HDTV images in realtime.

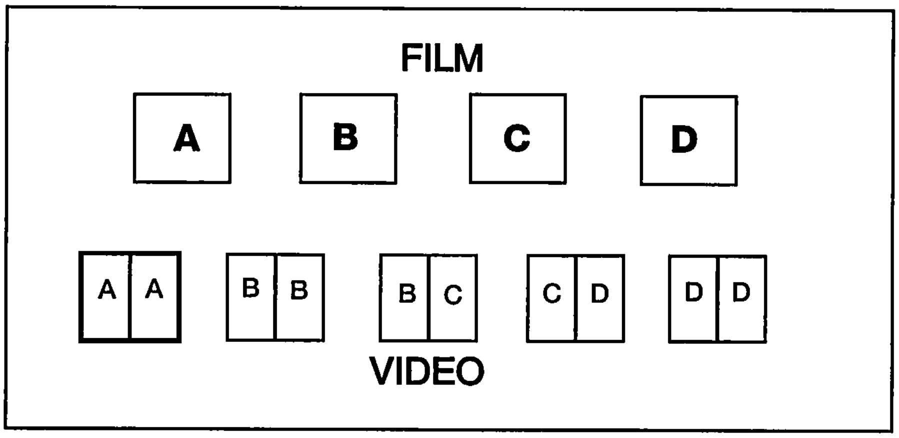

Figure 1. Illustration of 3:2 pulldown sequence.

Currently, user definition of the SMPTE A-frame start is necessary. Future software releases will provide full support of telecine log files, which contain a database relationship between film machine readable key numbers, video source time code, and 3:2 pulldown sequence.

The system reads in the first two video frames of the sequence and writes them to disk. It reads frame 3 and frame 4, holds them in memory, and copies the first field of frame 4 to the first field of frame 3. This new frame is written to disk. It then captures frame 5 and writes it to disk. Each time a frame is captured, it is also displayed at full resolution for monitoring. The four frames written to disk comprise the original A, B, C, D film sequence. In Fig. 2, the gray areas are the source frames written to disk.

Once the original 24-frame source material is written to disk it is progressive in nature. Each image is from the same moment in time as the film original, and there are no field artifacts. The user can define proxy images that are created and written to disk at the same time as the full resolution images, or as a post process. Proxy frame size is user definable and allows low-resolution display of the original HDTV images if necessary.

The inverse process can be applied at the output stage. A sequence of 24 frame/sec HDTV images can be output to a VTR as 1080i/60, 720p/60, or NTSC. The sequence can have 3:2 pulldown automatically inserted at output, with a user definable A-frame start time. The sequence is synchronized to tape time code, which allows the manual stop-down and pickup of output, or re-insertion of material into an existing clip on tape, maintaining the 3:2 cadence on the master tape.

Discreet finishing systems fire*, inferno*, flame*, and smoke*, support the loading and capturing of most industry-standard edit decision lists, or EDLs. The universal mastering functions added in the most recent software release allow the user to display the EDL in 30 frames/sec or 24 frames/sec. The system allows for capture with the automatic removal of 3:2 sequence either as a single clip input or an automated capture based on the EDL. The captured material can then be conformed in 24 frames/sec.

The high bandwidth capability of SGI's Onyx 2 enables full resolution monitoring of 1920 x 1080 images on a 24-in. widescreen graphics display and simultaneous display on an HDTV broadcast monitor at 30 frames/sec. This is accomplished by adding a "virtual" 3:2 pulldown sequence in the video player of the Discreet fire* application. This video player will play back the full resolution image with a simultaneous output to a HDTV broadcast monitor, allowing for display of 60 interlaced fields per sec. The support for automated 3:2 insertion is also an important pre-process for monitoring field-based motion artifacts of 24p content that is to be broadcast via existing interlaced transmission architecture.

Since the display is a virtual 30 frame/sec display of 24 frame/sec source material, the system tracks dual frame code modes. The application automatically tracks both 24 frame/sec time code and 30 frame/sec drop-frame or nondrop-frame time code for any clip that was captured with the automatic 3:2 removal process. The system also displays a graphical representation of the 3:2 sequence that will be added on output. This allows the user to edit and manipulate 24 frames/sec within the system, while always knowing the duration of the material after it is output as 60 interlaced fields. It also allows the user to quickly know if online editorial changes will be occurring on field 1 or field 2, once the 3:2 pulldown sequence is added at the output stage. The occasional occurrence of hybrid splices, or splices on B or C frames in the pulldown sequence, can be flagged and logged by the system.

Discreet editorial finishing systems, fire* and smoke*, support the export of industry-standard EDL formats. The output EDL function can add notations on an event-by-event

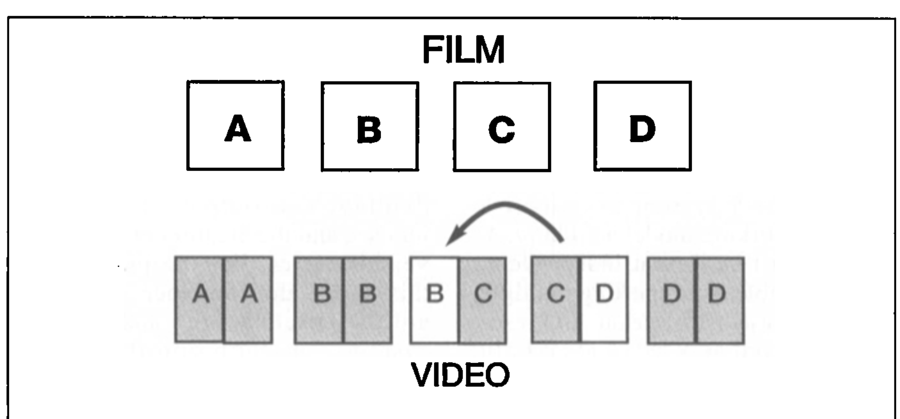

Figure 2. 3:2 removal; gray areas represent video fields stored to disk.

basis that describes the original A-frame sequence. The exported EDL also contains special marks recognized by the DaVinci color corrector that allow the color corrector to identify hybrid in and out points. This support allows the color corrector to automatically know where shot changes on field 2 will be, a vital step for tape-to-tape color correction.

Infrastructure, Storage, and Tools

The above workflow describes the input, monitoring, and output of 24 frame/sec material in the finishing system. The following describes some of the storage capabilities and advanced tools of the system.

Discreet finishing systems are based on a proprietary image file format stored on a proprietary file system, which has been optimized to deliver the performance required for high-resolution multimedia applications. The file system is dynamic and multiresolution capable. This allows writing and reading of multiple frame sizes to and from the file system without user intervention. A software process called "dynamic striping," which splits large frame sizes across multiple disks in the array, allowing fast access to and from data on disk is used. Furthermore, the user can store multiple resolutions on the same array without having to manually allocate disk space. Images of one frame size can therefore be copied to a framestore partition of another size using a dynamic, picture-based utility. This completes the universal mastering workflow, allowing one to manipulate images at 1080p/24 and then copy the entire project to a partition of a different frame size and frame rate.

For optimal performance, user-definable image quality levels can be applied to any image copy between framestore partitions. Other workflows would involve outputting the final image as 1080i/60 or 1080p/24 when recording devices are available, and performing a standards conversion external to the system.

There are also advanced tools within the system that allow for manual extraction and insertion of 3:2 pulldown sequence. This is a necessary tool for programming content that may mix original film footage with a known pulldown sequence and stock footage with an unknown pulldown sequence. The user can manually choose the A-frame sequence start and manually remove the pulldown sequence. Other advanced tools include interlacing, de-interlacing, and field merging of video images. These processing tools allow the user to easily mix multiformat images that may be progressive or interlaced.

Any clip in the system can be speeded up or slowed down—a process called timewarping. There are many attributes that can be added to a timewarped clip, including time changes based on fields or frames, user-definable mixes of frames, motion blurs, and keyframable animations. Timewarps can be automatically applied to a series of clips, which allow 60-field interlaced video to be added to a 24 frame/sec sequence. Any effects processing or compositing technique can be field or frame based, and each Discreet finishing system has motion blur capabilities. While these tools allow the mixing of interlaced and progressive images, the universal mastering workflow suggests frame-based processes for best results.

Future software releases will implement other tools necessary for universal mastering. Specialized audio tools will be included for sample rate conversion and pitch correction. This will make it possible to create PAL delivery masters from the original 1080p/24 source material without external audio processing.

However, framestore transfers of 24 frame/sec material to 25 frame/sec material can currently be performed while maintaining synchronized audio. As mentioned above, future support of telecine log files and film machine readable key numbers will provide tighter integration with film source material—a necessary function for film output applications. Current software releases support keyframable image reposition. While this allows minimal pan/scan functionality, additional functionality will be added with future releases.

Conclusion

The advancement of digital television has resulted in the need for a universal mastering format. Increasing demands for more content to fill the new digital broadcast channels, as well as multiformat distribution, require post-production finishing systems to be flexible, open, and upgradable, as well as format and resolution independent.

At first, it may appear that delivering in more than one standard or DTV format is complex, but, the 1080p/24 ATSC format has simplified the process and spawned a new workflow: shoot film or 24-frame progressive HDTV video, perform high-resolution image manipulation and compositing, and create various output formats in the highest possible quality, from a single HDTV master. Format-independent finishing systems and infrastructure from Discreet, along with high bandwidth computer platforms from Silicon Graphics, have delivered a facility solution and workflow for universal mastering, while providing a migration path for the future.

THE AUTHOR

Patrick Byrne graduated from Ithaca College, Ithaca, NY, with a B.S. degree in television and radio production. He worked in commercial post-production in New York, Montreal, and Los Angeles before joining Discreet. Most recently, Byrne was a senior editor and lead compositor, in Los Angeles, on various HDTV projects and special effects for theatrical release and broadcast. In 1997, he joined Discreet's western U.S. regional office as a HDTV specialist. In 1990, Byrne transferred to Discreet's world headquarters in Montreal where he works in product management.