DV.com - Inspiring and Empowering Creativity

Ben Waggoner

Video Format Conversion

Production and post used to be a parochial affair. Film, PAL, and NTSC projects stayed in their own worlds. Any conversion to another format was only done to a final, edited package. That's changing, though, and fast. More and more shops, and smaller and smaller shops, now need to deliver projects in multiple formats.

In the past, format conversion has been the province of expensive, dedicated standards-conversion hardware from companies such as Snell & Wilcox. These systems work quite well in general and are fast, but they are expensive and some have limited ability to change the speed of the output.

Today you can perform high-quality conversions with desktop software tools. This article explores lower cost software options and methods for converting between standard-definition formats. We will explore high-definition conversions another time.

To make a 24fps source such as film fit into a 3D fps medium such as NTSC video, the source frames get mapped to the output frames in a 3:2 pulldown, a repeating pattern of three progressive frames and two interlaced frames.

The formats

Film is by far the oldest format we talk about. The 35 mm, 24 fps sync sound version of film was introduced in 1926, and-from the format perspective-is the same world over.

The next big thing, or things, is **24 fps progressive** (24p). A variety of 24p formats are being deployed right now, each promising a single, universal master format for production and post. If there was a truly universal 24p format, this article would be much shorter. But, alas, no such format exists.

From a postproduction perspective, 24p is very much like film but doesn't require a telecine process. The 24p variants include 720 x 486 and 720 x 480 standard-definition images, and 1280 x 720 and 1920 x 1080 high-definition images. Aspect ratios are 4:3 and 16:9.

Some of the new DV-based 24p cameras actually store the image data on the tape in 60-field mode like normal NTSC; but if your NLE doesn't transparently convert the footage to 24p, you'll need to do that.

NTSC, used in North America and Japan, runs at 30,000/1001 frames per second (conventionally, but not precisely, rounded to 29.97 fps). This unusual frame rate is the source of a number of subtle conversion problems, especially if the 29.97 approximation is used instead of the actual 30,000/1001 value. NTSC formats have either 480 or 486 active lines of resolution. Note that the 480 lines, used in DV25 formats, is a subset of the 486 lines used in other 601-compliant formats. The pixel aspect ratio is the same. These are really the same format, only the DV frame is placed in the 486-line frame with four blank lines at the top and two at the bottom

PAL, the video format used in Europe and Asia, is in many ways superior to NTSC, especially for conversion from film. It has the nice, even frame rate of precisely 25 fps, and it has only one number of active lines: 576.

Adobe After Effect's inverse telecine can guess at a source's cadence pattern, but does so by looking at the first few frames of source video, which isn't helpful if a tape starts with black. A little trial and error will get things rolling correctly.

The parameters

Video formats use different resolutions. For production in NTSC, digital resolutions of 720 x 480 and 720 x 486 pixels are standard, and 720 x 486 is standard for broadcast delivery formats. PAL production and delivery both use 720 x 576 in most cases.

Note that resolution and aspect ratio are two different issues. NTSC and PAL can be produced and authored at either a 4:3 aspect ratio (by default) or 16:9 (increasingly common, especially with DVD and PAL broadcast).

Time is the most vexing aspect of standards conversion. Resolution is malleable: 486 pixels can convert into 576 and back again very smoothly. But it's much more difficult to interpolate time values and establish good intermediate frames.

Frame rate is the important factor for progressive formats, but field rate is the key factor for interlaced formats. Fifty fields go into 60 fields smoother than 25 frames go into 30.

The critical issue with temporal interpolation is the time gap between the input frame and the output frame.

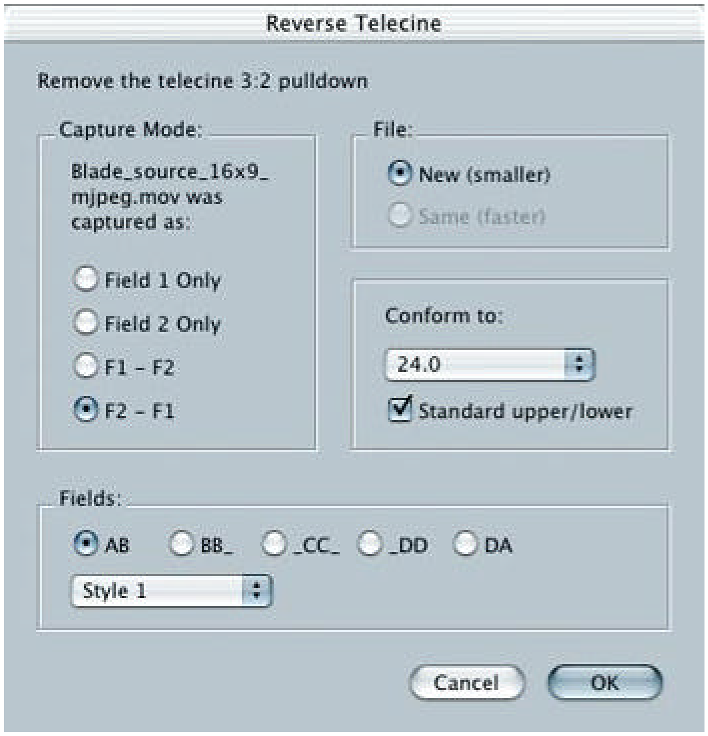
Many engineers are looking for better ways to handle temporal interpolation, so I hope we'll see substantial improvements over the next few years. For example, Microsoft's demonstrations of a realtime frame rate resampler in Windows Media 9 are promising.

In some cases, the best way to handle the transfer is to change the speed and duration of either the source or output video to get frame rates that align better. Thus the output can be significantly longer or shorter than the source.

Field order is one of the subtler and frustrating aspects of working with mixed sources, largely because poor documentation makes it hard to determine the field order used by different formats. The first line of the frame may be displayed before the second or the second may be displayed before the first. Guessing the wrong field order results in video that stutters when there is any motion because the content is being displayed out of order.

Luminance and color vary from PAL, NTSC, and film, with each displaying different ranges. For PAL and NTSC conversions, the differences are small and you generally need to adjust gamma to compensate. As long as the same internal values for luma and chroma are maintained throughout the conversion, the actual display should conform to the appropriate gamma.

Video to film conversion processes often can add film grain to match a particular stock. Although this is an essential technique for projects trying to match film and video-produced assets, I don't recommend it for projects that are entirely video source. Adding film grain is simply adding noise, which hurts quality. I'm a firm believer in playing to each medium's strengths, and one of the wonderful things about video is that it can have less grain than film. Instead of trying to fake a project so that it looks like it was shot on film, which never works, I think it's best to use the strengths of the video source and leave it clean.



Apple Cinema Tools works great with carefully telecined material, but doesn't let you preview without rendering.

The conversions

Film to 24p video transfers should be extremely straightforward-a frame of source becomes a frame of output. The only complication is if you use 23.976 fps (24,000/1001) source instead of a straight 24. In that case, apply a simple 0.1 percent time expansion to the source for accurate 24 fps output.

Film is transferred to video so often through telecine, we generally forget that the process is an intrinsic pain. That is, 24 simply doesn't go evenly into NTSC's 59.94 (60,000/1001) fields per second, so a 3:2 pulldown is applied. First the film is slowed down by 0.1 percent to 23.976 fps (24,000/1001) so 24 frames of film and 60 fields of video will have the same precise time base. Then the film is mapped to video with the first film frame becoming two fields of video, the following frame becoming three fields, the next becoming two fields, repeating ad infinitum.

This process results in a video with the characteristic 3:2 cadence. The video stream will consist of a series of three progressive frames, followed by two interlaced frames (where a given video frame's fields have images from two different film frames).

Because of the 0.1 percent slowdown as part of the telecine, the 3:2 pattern should be rock solid for video no matter its length. Unless something went wrong, an entire feature-length film will maintain the cadence throughout.

Film to PAL video transfer is a lot simpler in theory and practice, but it has some interesting ramifications. Because there's no good way to divide 24 into 25, the telecine machine is sped up to 25 fps, and the film is turned into 25 fps progressive PAL. This is a heck of a lot easier to work with than NTSC's 3:2 cadence. However, it shortens the duration of the media by a significant 4.2 percent, cutting the duration of a two-hour movie by nearly five minutes (for plenty of European films, this can be quite welcome). The speedup also raises audio nearly a full step in pitch, so the audio will need to be resampled using an algorithm that keeps pitch constant.

NTSC to film transfer, especially a proj-ect shot as standard 59.94 fps interlaced NTSC, can be extremely daunting. However, all of those folks who want to shoot a feature on DV and transfer it to film (note that I'm writing "want to") have caused a lot of smart engineers to try to make the NTSC to film process as seamless as possible.

In essence, the programs perform an inverse telecine on a source with really bad cadence breaks. Most of the high-end tools use proprietary algorithms, but here's a 10,000-foot view of some simpler ways to tackle this transfer.

The first technique is to figure out which video field is closest in time to the prospective film frame output. The closest field can then be deinterlaced and used as that frame. The resulting motion won't be perfectly smooth because 36 out of the 60 original temporal samples in the video are discarded and the remaining 24 aren't distributed evenly. But the end result isn't bad.

With the next level of complexity, the source frame of video can be adaptively deinterlaced so the parts of the image that aren't moving, aren't deinterlaced, preserving more resolution. Many software tools can provide this level of conversion.

More sophisticated tools use motion tracking techniques to gain the advantage of adaptive deinterlacing to parts of images with motion. Expect a fair amount of development in this area over the next few years, at least until 24p cameras become universally used for film-out projects.

One interesting subset of NTSC to film transfers is projects that were originally shot on film, telecined to NTSC video, and then converted back. In those cases, an inverse telecine can effectively make the NTSC into 24p. However, a strong inverse telecine algorithm will be required if the source was edited without preserving the cadence of the source.

NTSC to 24p transfer is pretty much the same as going to film, with the same issues to solve and the same tools to solve them.

But 29.97 doesn't go into 25 smoothly, so a reversible technique such as 3:2 pulldown isn't an option for converting NTSC to PAL. However, you do have the option to sync with fields instead of frames, and hence can convert from 59.94 to 50, which reduces the amount of temporal mismatch by half instead of just comparing frames. As a side note, this is why interlaced NTSC works better than 30p NTSC for standards conversion (and for film out or for 24p).

NTSC to PAL conversion takes the NTSC field closest temporally to each PAL output field, and disregards the leftover fields.

For resolution, PAL has 576 lines and NTSC has 486 lines. When you're working with 486-line NTSC video and an editing system that supports it, converting NTSC to PAL is easy. But when you're coming from a 480-line format such as DV or MPEG-2, you should add four lines to the top of each frame and two to the bottom before scaling to 576 lines. If you need to reverse field order, add three lines to top and bottom for a quick fix.

Dealing with luma differences isn't as difficult as it may first appear. PAL has a minimum black IRE of 0, instead of NTSC's 7.5. However, in digital, both formats define black as Y = 16, so this shouldn't need to be explicitly addressed during conversion.

PAL to film or 24p conversion usually involves slowing down PAL from 25 fps to 24 fps, and making one film frame out of each PAL frame. This, of course, makes the content about 4.2 percent longer, so a two-hour video presentation would be about five minutes longer than on film.

Because a proper PAL inverse telecine keeps a 1:1 relationship between source and output frames, a simple adaptive de-interlace can provide excellent results.

A motion estimated deinterlace can do even a little bit better yet. And, of course, progressive PAL wouldn't require any de-interlacing at all.

PAL progressive 16:9 converts much better to film and 24p than does anything sourced in NTSC, which is why it was used for many DV-to-film productions in the past. However, 24p will probably replace it soon.

PAL to NTSC requires converting from 576-line PAL to a 486-line image. If you're going to a 480-line NTSC format such as DV and MPEG-2, the image should first be scaled to 486 lines, and then the top four and bottom two lines cropped out of the image. To reverse field order, add three lines to the top and bottom of the 480-line image.

Progressive PAL, including images transferred from film to PAL, should be slowed down to 24p and then 3:2 pulldown can be applied. But this isn't appropriate for the much more common interlaced PAL. Instead, time can be kept constant and field interpolation can be applied in rendering the video.

Discreet Cleaner 6 offers a strong inverse telecine and provides pre-processing previews, but it can't make temporal adjustments.

The software

Hardware transcoding tools offer excellent quality and realtime performance, but they are also expensive and difficult to transport in this era of finishing on a laptop. Software tools hold the promise of much less expensive standards conversion.

My ideal tool would have automatic support for the different resolutions, precise time contraction and expansion, and a robust inverse telecine that can deal with cadence breaks. I have yet to find one tool that does all of these. Instead I use a hodgepodge of different solutions.

After Effects 5.5 ($699 Standard, $1699 Production Bundle) is a champ at dealing with different video formats. For transcoding, the Standard version is all you need. After Effects provides appropriate defaults for pixel shape and resolution for all of the major formats, and will automatically handle 480-/486-line differences. However, you must jump through a few hoops to get the temporal processing correct.

To transfer from PAL video sourced from film or progressive images to NTSC or 24p, it is best to interpret the footage as 24 fps to a 24 fps project. In the case of NTSC, apply the 3:2 pulldown and conversion to 59.94 fields per second in Render Settings.

With the other sources, things are a little more straightforward. Interlaced PAL, NTSC, film, and 24p can be imported without any time-shifting. Film to PAL should be sped up to 25p on output, of course.

After Effects 5.5 has weak inverse telecine. The source must not have any cadence breaks at all, although After Effects can try to guess what the cadence pattern is. However, it does this by analyzing the first few frames of video, so the feature doesn't work if the source has a black leader. In the end, you'll generally need to use trial and error to find the pattern.

The Preview window gets updated as soon as the Interpret Footage dialog is closed, so it typically only takes me a minute or two to find the right settings.

You'll also need to make sure any differences between source and output resolutions are corrected. The simplest way to do this is to drag a corner handle of the video frame to align with the corner of the project's active area. After Effects keeps the centers locked, so getting one corner correct should get all of them correct. However, remember to convert to or from 486 lines instead of 480 lines when dealing with PAL.

Final Cut Pro 3.0 ($999) for Mac OS does a fine job of working with different video modes, but by itself doesn't offer much facility to convert between them. Apple's $999 **Cinema Tools 1.0** for Final Cut Pro provides the ability to convert between a variety of telecined formats with excellent integration. Together, these two applications make a good solution for dealing with telecined formats that have rock-solid cadence or a window dub of the frame rate. If you don't know the cadence of the source, you'll need to find it by trial and error.

However, as with After Effects, Cinema Tools doesn't even try to deal with source that has cadence breaks, and you'll need another tool to deal with that kind of footage. Cinema Tools also doesn't provide any way to preview without rendering out a whole file, so if cadence isn't shown in a window dub, determining it can be painful. Understandably, Cinema Tools is best when used for the purpose it was designed for: working with telecined source that has window dub frame numbers or when a telecine log is provided.

I have yet to find a good way to convert between interlaced NTSC and PAL with Final Cut Pro. Although the application can easily change the speed of a track within a movie, it doesn't do the appropriate field rendering when there is a mismatch between source and output frame rate. However, in a pinch, Final Cut Pro can do a tolerable job.

Discreet Cleaner ($599) has historically had the strongest inverse telecine algorithm available at its price point. However, a serious bug in Cleaner 5.0.2 can result in duplicated and even reversed order frames. The new Cleaner 6 fixes this bug. Cleaner 6 is out now for Mac OS, with Windows coming in '03. The older Media Cleaner Pro 4 has a functional inverse telecine without the bug present in 5.0.2.

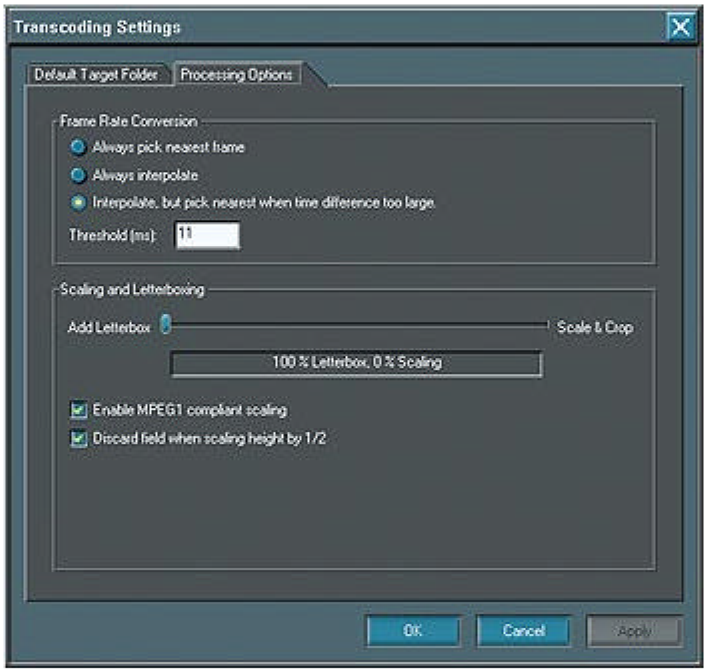
But no version of Cleaner can make temporal adjustments to footage. For standards conversion from telecined NTSC with irregular cadence, I use Cleaner 6 to create a progressive 23.976 fps out-put file, and then apply temporal adjustments as needed in my NLE.

Cleaner performs a decent adaptive deinterlace. However, it doesn't do field-savvy rendering for standards conversion, so it isn't an appropriate tool for conversions between PAL and NTSC. (Note that I worked for the previous developer of Cleaner.)

Canopus ProCoder 1.2.1 ($699) for Windows has strong features for transcoding. Beyond mere adaptive deinterlacing, ProCoder can blend between source fields when making PAL-NTSC conversions. The blending results in smoother apparent motion with most outputs.

ProCoder currently lacks an inverse telecine filter. Canopus says it plans to release an inverse telecine filter in the first quarter of 2003. Also, ProCoder currently doesn't handle temporal adjustments, leaving it less than ideal going to 24p from PAL.

Other than that problem, ProCoder's interface makes it easy to get good transcoding results. For example, if you set an NTSC source and a PAL MPEG-2 output, ProCoder automatically does the right things internally without requiring manual user tweaking. It can even automatically letterbox when going between 16:9 and 4:3.



Canopus ProCoder has a simple clean interface that makes it easy to get good transcoding results. However, it doesn't currently handle temporal changes.

Transformed

In today's multiformat world, it is possible and in some cases even desirable to do standards conversion in software instead of with hardware. The ideal software transcoding product would combine good inverse telecine, adaptive deinterlacing, and temporal expansion and contraction. A single such product doesn't exist yet, but until it does, you only need to combine a couple low-cost applications that you may already own to make a good transcoding system.

Resolution for Graphics

When they're creating graphics for use in video projects with applications such as Photoshop, many operators create files in square-pixel resolution and let the NLE or DVD authoring application handle the conversion to nonsquare video formats. Good modern applications will maintain the correct aspect ratio during conversion to NTSC or PAL, so circles won't become ovals.

It's always preferable to scale down images rather than scale them up to avoid introducing scaling artifacts into the processed videos. To that end, square-pixel graphics need high-enough resolutions so no axis is scaled up when converted to NTSC or PAL formats. The good resolution numbers for good square-pixel graphic sources are 768 x 576 pixels for 4:3 projects and 1024 x 576 pixels for 16:9 projects.

Of course, these numbers work for 486-line NTSC as well as PAL projects, but strictly speaking can lead to a 1 percent vertical squash for NTSC DV with its 480-line frame, if the image is brought into a 720 x 480 comp. For 720 x 480 comps and timelines, a 720 x 533 pixel source is more correct (Adobe uses 720 x 534 in its Photoshop preset, which is close enough).

For More Information

You can learn more about the products and companies mentioned in this article at the following Web sites:

Adobe After Effects
www.adobe.com/aftereffects

Apple Final Cut Pro and Cinema Tools
www.apple.com/finalcutpro

Canopus ProCoder
www.canopus.com

Discreet Cleaner
www.discreet.com

Snell & Wilcox
www.snellwilcox.com

Windows Media 9
www.microsoft.com/windowsmedia

Ben Waggoner has been compressing and transcoding video for over 10 years. His new book, Compression Techniques for Great Digital Video (CMP Books, 2002), is out now.