Encoding for Blu-ray Disc and HD DVD

Reaping the Benefits of Integrated Mastering

SONIC USA • 101 Rowland Way • Novato • CA 94945 • Tel: 415.893.8000 • Fax: 415.893.8008
SONIC BURBANK • 3500 West Olive Avenue • Suite 730 • Burbank • CA 91505 • Tel: 818.823.4000 • Fax: 818.823.4004
SONIC EUROPE • 22 Warwick Street • London • W1B 5NF • UK • Tel: +44 20 7437 1100 • Fax: +44 20 7437 1151
SONIC PAC RIM • 4F Shirokanetakanawa Station Building • 1-27-6 Shirokane • Minato-ku Tokyo 108-0072 • Japan • Tel: +81 3 6408 2811

DIGITAL VISION SWEDEN • Solna Strandväg 98 • SE-171 54 Solna • Sweden • Tel: +46.8546.182.00
DIGITAL VISION UK • First Floor, 151 Wardour Street • London • W1F 8WE • UK • Tel: +44.20.7734.8282
DIGITAL VISION USA • 4605 Lankershim Blvd • Suite 700 • North Hollywood • CA 91602 • USA • Tel: +1.818.769.8111

CONTENTS

INTRODUCTION

New display technologies, video codecs and higher capacity optical disc formats with much greater read-speeds are all contributing to a new set of challenges to those tasked with producing high quality consumer deliverables like HD DVD and Blu-ray Disc.

Consumer investment in HD equipment and content is fuelled by a strong expectation of greatly improved video and audio quality over SD transmissions and DVDs. Those delivering this content must strive to meet this expectation with the highest quality masters possible.

This document is targeted at those involved in the preparation of HD DVD and Blu-ray Disc masters but many of the principles discussed translate to a wide range of post-production deliverables. It illuminates some of the new challenges of mastering for those next generation optical disc formats and proposes new and practical solutions in both processing and workflows.

Questions and Feedback

If you have any questions about the content of this paper, please contact your nearest Sonic or Digital Vision representative using the information provided below. We appreciate any comments or feedback you provide. If you have any questions, or have any suggestions on how any areas could be improved, please email your comments in confidence to whitepaper@sonic.com.

About Sonic Solutions

Sonic Solutions is the leader in digital media software, providing a broad range of interoperable, platform-independent software tools and applications for creative professionals, consumers, and technology partners. Sonic's products include professional encoding and authoring systems used by Hollywood studios to create the most-advanced and highly interactive Blu-ray Disc, HD DVD, and SD DVD titles.

Sonic has been the leading provider of digital media creation technology since the inception of digital media itself. Following on from SonicStudio - Sonic's industry leading CD editing and premastering solution - Sonic was the first company to deliver a professional DVD-Video production workstation in 1996. Sonic's Scenarist system has now been used to release over 4 billion DVD titles. A founding member of the HD Authoring Alliance, an association of top authoring facilities worldwide, Sonic is dedicated to accelerating the development of the authoring skills, capacity, and technological infrastructure required to support the rapid adoption of HD DVD. Working together, the HDAA have helped design and refine the optimum HD DVD workflow resulting in industry leading products like Sonic Scenarist and Sonic CineVision.

About Digital Vision

Digital Vision provides innovative image restoration, enhancement, color correction and data conforming systems that movie studios, television networks, and post-production facilities use to produce and enhance feature films, TV programs and commercials. The company's Nucoda product line provides a strong suite of products for the burgeoning digital intermediate 2k/4k market. The company's award-winning products are a standard of the media & entertainment industry and are deployed at top facilities around the world.

More Information

Please contact your regional Sonic or Digital Vision representative for more information.

• Sonic U.S.A. (West)	Terry Marshall - terry_marshall@sonic.com. Tel: +1.818.357.7825
• Sonic U.S.A. (East), Canada, & Latin America	Brian Murphy - brian_murphy@sonic.com. Tel: +1.415.893.7832
• Sonic Europe, Middle East, & Africa	Richard Linecar - richard_linecar@sonic.com. Tel: +44 20 7437 1100
• Sonic Pacific Rim	Terry Marshall - terry_marshall@sonic.com. Tel: +1.818.357.7825
• Digital Vision Sweden	info@digitalvision.se. Tel: +46.8546.182.00
• Digital Vision UK	info@digitalvision.se. Tel: +44.20.7734.8282
• Digital Vision U.S.A.	info@digitalvision.se. Tel: +1.818.769.8111

You can also contact your nearest Sonic reseller using Sonic's online reseller locator: www.sonic.com/reseller or visit the Sonic or Digital Vision Web sites www.sonic.com and www.digitalvision.se.

NEXT GENERATION VIDEO CODECS

HD DVD and Blu-ray Disc introduce two new video codecs, in addition to MPEG-2, to the world of video delivery in the home. The two new formats include great advances in optical media technology providing both greater data capacities and read/write bandwidth.

Format	Capacity per layer (GB)	Maximum Video Bit Rate (Mbps)
SD DVD-Video	4.7	9.8
HD DVD	15	29.4
Blu-ray Disc	25	40

The most obvious and important difference between MPEG-2 and the new AVC (H.264) and VC-1 codecs is a significant improvement in efficiency. The advances in processing technology have made practical a variety of new, sophisticated, and more computationally expensive tools which achieve a greater compression ratio for an equivalent objective quality.

Typical Bit Rates

The vastly increased capacities of HD DVD and Blu-ray Disc (30GB and 50GB for dual-layer) mean that the average bit rates that can be used for a feature length title are very high. Despite the amount of picture data for HD (1920x1080) being about 5 times that of SD (720x576) the much larger capacities coupled with the efficiency gains of the new video codecs means that in real-terms the average data rates have increased significantly. Whilst a somewhat nominal approximation, it has been frequently suggested that the newer codecs like AVC or VC-1 are capable of reproducing video at the same objective quality at half the bit rate of MPEG-2. The table below illustrates this point:

Format	Disc Capacity (GB)	Average bit rate for 2 hours 30 minutes of content (Mbps)	Equivalent bit rates if resolutions were equal (Mbps)	Equivalent bit rates taking into account codec gains (Mbps)
SD DVD-Video	4.7	7.5	7.5	7.5
HD DVD	15	26.7	5.3	10.6
Blu-ray Disc	25	44.4	8.9	17.8

Note: Blu-ray Disc actually has a maximum video bit rate of 40Mbps so the 44.4Mbps quoted is for the sake of comparison only.

So at first sight everything looks like it should be a great deal easier than it was before and in some cases this is true. For certain material it is certainly the case that, because the overall capabilities of the encoder are greater, the emphasis moves back to how good the source actually is. However, there are now some significant new challenges that perhaps eat into the head start the new codecs and increased bandwidths afford us. We will discuss these challenges, and the shift in which stages of the workflow are key to the eventual quality of the disc, have moved between standard definition DVD-Video and the new HD formats Blu-ray Disc and HD DVD later in this paper.

AVC and VC-1

Both of the new codecs available for optical disc delivery (AVC and VC-1) share a lot of the same technological advances so we will give a brief and general introduction to some of the key ones here. It's worth noting straight away that the key downside of all of these advances is an increase in processing overhead either at encode time, at decode time, or in some cases at both stages.

In-loop De-blocking

Both AVC and VC-1 streams convey to their decoders whether or not they should apply a filter which is designed to detect the edges of visible macroblocks and filter their edges to disguise them. VC-1 only allows for a simple on or off control for this filter but AVC actually contains two further parameters to control the strength of the filter. Since it is processed by the decoder itself, de-blocking adds a processing overhead into decoders but in the vast majority of cases it improves the perceived quality of the compressed video a great deal.

The de-blocking filter is described as in-loop since, as well as being displayed, each de-blocked frame can be used as a possible reference frame for encoding/decoding other neighboring frames.

Smaller Block Sizes

All video codecs work on the basis of splitting an image up into square groups of pixels called macroblocks and indeed all three of the codecs here use a macroblock size of 16x16. However, in the case of AVC and VC-1, each macroblock can optionally be subdivided into smaller blocks as small as 4x4 with each sub-block gaining its own motion vector and color information. This has particular benefit for encoding camera effects such as zooms where spatial motion estimation can break down somewhat if the “size” of content within a frame is changing rather than just moving.

Improved Entropy Encoding

Once again all three video codecs employ entropy encoding (a lossless compression method that works by detecting patterns and storing how frequently they occur) after the picture information has been transformed into the frequency domain. AVC has a choice of two entropy modes CABAC and CAVLC. Content adaptive binary arithmetic coding (CABAC) is the most computationally expensive and complex entropy coding mode available in all the codecs but it does achieve up to a 20% efficiency gain over the VLC entropy coding scheme used in MPEG-2. By way of illustration, if it was possible to use CABAC in MPEG-2, theoretically an 8Mbps encode could look the same as a 10Mbps encode. VC-1 uses a slightly less complex high-order entropy coding model which also offers significant gains over MPEG-2.

MPEG-2

Whilst not dwelling on the technical details of the MPEG-2 codec (since they are well documented and widely understood) it is important to understand the advantages MPEG-2 can still have over AVC and VC-1.

Its principle strengths lie in its maturity and simplicity of both encoding and decoding. Both AVC and VC-1 present a tough challenge for decoders to achieve real-time and, in particular, AVC requires a very high performance multi-core processor in order to decode in software in real-time. This is, of course, less of a consideration when the target playback device is a set-top player.

NEW CHALLENGES

In standard definition DVD-Video production the lower resolution and basic MPEG-2 video codecs employed mean that the ultimate complete tool-set for achieving a very high quality deliverable is actually quite small. However, the move to high definition along with the use of new advanced codecs like H.264 and VC-1 means that there is now a much larger set of tools available. In DVD-Video production the area of the workflow key to delivering a high quality encode was the quality of the MPEG-2 encoding used and the only ways to improve it were via adjustment of encoding parameters and the use of simple noise removal filters.

As production has begun for new formats like HD DVD and Blu-ray Disc a new set of challenges has become clearly apparent. The notable thing about these new challenges is that they primarily centre around the interaction between mastering and encoding rather than the quality of the MPEG-2 encoder.

With an increased range of tools, which interact with each other, key decisions affecting the final quality of the encoded master take place during that part of the workflow previously handled during mastering.

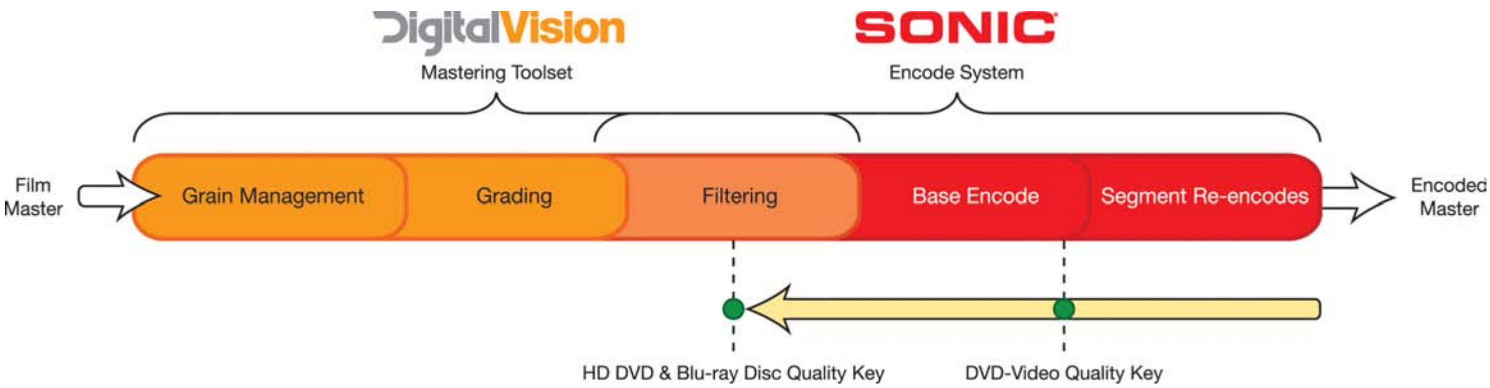

Fig. 1 - The point at which encode quality is determined for the HD formats is further 'up' the workflow than for SD DVD.

The remainder of this chapter discusses the two most fundamental challenges in the HD encoding world: the inherent quality limitations of 8 bit 4:2:0 video and the impact film grain has on successful encoding.

The 8-bit 4:2:0 Barrier

The HD DVD and Blu-ray Disc specifications list a large number of constraints upon the video codecs that they employ. Amongst that list one fundamental item (the profile adopted) has a massive impact on the overall quality that's attainable by inferring (amongst other things) that the bit depth and sampling format is 8-bit 4:2:0.

What does 4:2:0 mean?

'4:2:0' refers to the sampling rate of the luminance and chrominance signals. It means that the chroma components Cr and Cb are sampled half as often as the luminance component in both the horizontal and vertical domains. In simple terms, each square of four pixels has a luminance value for every pixel but only color information for the block of four as a whole.

Since the human eye is significantly less sensitive to chroma information this is rarely perceptible when viewing 4:2:0 video. However, it is important to be aware of this fact when working with an encoder. Encoders share their attention between the luma and chroma components and if their attention is too weighted in favor of luminance then noticeable and unwanted artifacts can be present only in the chrominance information. A good encoding system will permit the user to manually adjust the weighting between the two in order to correct problems of this nature. It's also important to remember that an inevitable by-product of chroma sub-sampling is that there is some "bleed" between luma and chroma. Consequently, poorly reproduced chroma information can affect the luminance to which we are significantly more sensitive.

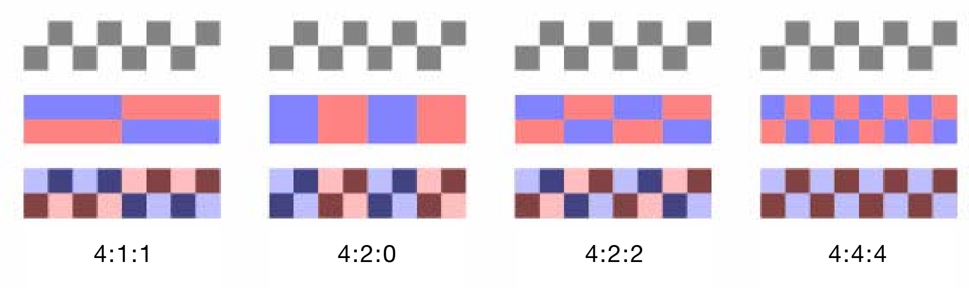

Fig. 2 - Illustrations of the Y, U, and V video signal components at different sampling formats.

Is 8-bit sampling a problem?

8-bits can convey 256 discrete values. Consequently, in this conversation about 8-bit 4:2:0 video for each block of four pixels, each individual pixel has a 0-255 level for luma and the block as a whole has a 0-255 level for Cr and a 0-255 level for Cb. In reality however, the 0-255 range is actually reduced to 16-235 by the specification of the use of ITU-R BT.601 which limits the range to avoid overshoots in transmission. 16 therefore represents black and 235 represents full white and consequently there are actually only 219 possible levels.

Since we are now all very used to viewing 8-bit video (DVD-Video has been considered a high quality video format for over 10 years now) it is clear that this doesn't necessarily present much of a problem. However, there are certain situations where this limitation of 219 levels produces noticeable and almost always undesirable visual artifacts.

The most common problem is when large areas of an image have a block of the same color or, even worse, a very gradual and linear gradient from one color to another very similar color. For example a large area of blue sky might only traverse two or three levels in the 16-235 ranges and consequently if large enough areas that quantize to exactly the same levels are present it can be quite easy for visible steps from one level to the next to emerge from the image. This effect is often called "banding".

This problem is also more prevalent in animation and other CGI material where gradients and colors tend to be a lot more "perfect" than in real life. To some extent banding in animation has become an accepted part of the viewing experience but when encoding this sort of material it is important to ensure that the encoding process does not introduce further banding or exaggerate it.

An example of some 8-bit 4:2:0 banding from a CGI feature is shown in figure 3.

Fig. 3 - An example of 8-bit 4:2:0 banding

Finally, it's also important to be aware that 8-bit banding is most obvious and irritating to the viewer in darker areas. The step between 16 and 17 is almost always very much more noticeable than the step from 234 to 235. Most encoder algorithms treat all luma levels equally and produce what internally to them is an equal quantization irrespective of the actual luma level. For this reason it can often be necessary to look at methods both in mastering and encoding to combat this.

Encoder techniques for combatting 'banding'

The approach to improve banding at the encode stage is to apply a dithering algorithm to the levels within the image. Ideally this should only be applied to areas of the image that actually contain bands, so some detection of the bands should also be included. Dithering areas of an image that do not contain any banding will only serve to distort the content slightly.

The images in figure 4 show an extreme example of dithering in action. The first image is the original full resolution frame followed by 16 color version first without dither and then including dither.

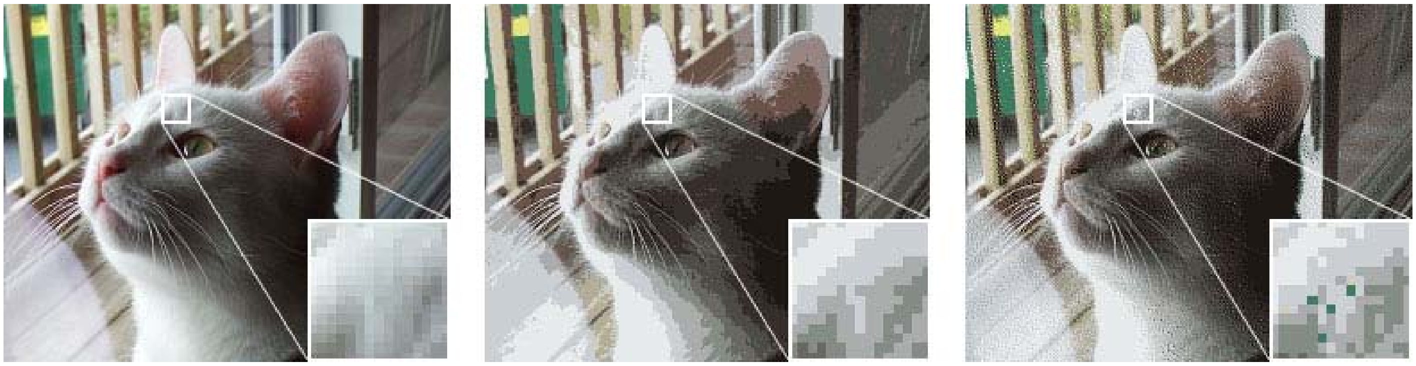

Fig. 4 - Examples of dithering

The primary problem with attempting to fix banding within the encoder component itself is that it will always be working in the 8-bit 4:2:0 space. Dithering is only effective when working with the higher resolution material (a 10-bit source in the examples here). Therefore, for an encoder compressing 8-bit 4:2:0 material it is difficult for it to also include an effective dithering stage since the 10-bit information has already been lost. It is clear that the dither should be applied to the frames before the actual compression process begins.

One other simple tool also warrants a mention here. Because banding is most objectionable in dark areas, a quick improvement can sometimes be achieved by normalizing levels that are near black to black itself (for example adjusting everything in the 16-20 range to 16). Unlike dither, this is a simple tool that can be applied by the encoder itself within the 8-bit space.

Pre-processing techniques for combatting 'banding'

Modern film digital post-production is often performed at either 2k or 4k pixels per line and at 10-bits or more per RGB component. Reducing this to HD resolution and 4:2:0, 8-bits per YCbCr component, involves making compromises. Ideally the impact on the image of these compromises needs to be assessed whilst making the final creative decisions in the color grading suite. Here there is an opportunity for the director to preview what the final disc will look like and make adjustments to the "look" of the film based on that preview. Many things contribute to this "look"; overall color balance, creative color control of individual elements, softness or sharpness in the image, and the amount of grain or noise all have an impact.

By using an integrated suite of tools the director has the opportunity to take creative decisions to improve the final look of the disc, decisions which it may not be appropriate for an operator encoding the film to take.

At first it may seem appropriate to reduce the grain as much as possible in an image in order to allow the encoding algorithm to compress the image paying attention to image detail rather than noise. Unfortunately this gives a look which is quite unlike film and is often undesirable to the director. Therefore a balance is needed between film grain and encoder in order to achieve the best possible result on a domestic display.

Computer generated (CG) imagery can suffer from the opposite problem. Very smooth, graduated areas of pure color can make banding more visible. The addition of a small amount of artificial grain to these images may improve the final result. Proper integration between the re-grain algorithm and the encoder is required if the grain is not to be discarded in the conversion from 10-bit RGB to 4:2:0.

Typically a noise algorithm will be used to help reduce banding. There are various different methods available which use dithering, rounding or some other method of disturbing the chroma or luma values slightly on a pixel basis. However, film images will contain an amount of natural grain which has an organic nature, spreading over more than one pixel (see below). DVO Re-grain offers the opportunity of introducing grain which visually matches that within the existing material, whilst allowing the operator to control the amount, size, sharpness and effect in the dark, mid or bright areas. This allows the operator to address particular issues such as banding within the shadows. By modeling the normal characteristics of film grain we can avoid the possibility of introducing visually disturbing results.

Again, these are creative decisions which the director may want to have control over and that is only possible in a well integrated multi-format mastering environment.

The Grain and Noise Problem

The nature of grain

In film the image is recorded by grains of silver halide in the emulsion of the film. It is the size of these grains which causes the granular appearance of film images. Larger grains are more sensitive to light, so filming in dark conditions, or at high speed requires a fast film stock which will exhibit more grain. The example in figure 5 is taken from a small section of a 4k film scan.

Fig. 5 - An example of film grain taken from a section of a 4k film scan

Grain is more noticeable in darker regions of the image and in the blue channel as illusrated in figure 6.

Red Green Blue

Fig. 6 - Film grain is more noticeable in darker regions and in the blue channel

Grain size is dependent upon a number of factors; the film stock, the exposure, and how it is developed. Film grain varies in size a great deal but may be between 10 and 30 microns. On a piece of 35mm film in Academy format the image area is 22mm wide. At standard definition each pixel of a line of the final image would be approximately 30 microns across. This means that in an SD image you may be aware of the some variability in the image due to grain, but would be unlikely to see the grain itself. In a high definition image, by comparison, each pixel would be approximately 10 microns across making the grain far more likely to be directly visible.

Our eyes are very sensitive to grain in a moving picture. For example when computer generated (CG) visual effects are added to film footage a great deal of effort is expended in post production to duplicate the grain of the original film and add it to the CG. Without it the CG stands out in a most artificial way.

What of digitally captured media? If the media is not originally shot on film then one would expect all of this talk of grain to be irrelevant. Unfortunately it turns out that some cameras produce CCD noise which is different from, but analogous to film grain. If the lighting is not good, or the camera setup is not perfect, then CCD noise can be even more visible than film grain and is sometimes considered by Directors and Directors of Photography to be less visibly appealing. The management of this noise therefore becomes a critical creative and technical issue in the finishing and mastering of the HD DVD or Blu-ray Disc.

Grain reduction

Straightforward attempts to reduce grain by spatial filtering lead to a significant loss in sharpness and detail. In applying spatial filtering information is removed from the image. What is needed is a method to generate the information that should replace the grain. This is achieved by using a method of temporal and spatial analysis called motion estimation.

The Digital Vision motion estimation algorithms have been under continuous development for 20 years and have won numerous awards. They have been used on a vast number of films and television programs in order to manage grain, conceal dust and scratches, and to convert television programs from one standard to another (when entirely new frames need to be created in going from 25 frames per second to 30 frames per second).

The motion estimation allows the grain reduction algorithm to use information from adjacent frames to create new detail for the current frame. This gives the highest possible quality.

The DVO Grain algorithm allows the user fine control over how the grain is reduced. A simple "Mode" control gives a quick setting to control the amount of grain reduction. When more control is required the user can control how much the grain reduction effects luma or chroma channels, and its effect in the dark, mid or bright regions of the image. In order to improve creative flexibility there is also the ability to blend back into the image an amount of the original grain.

The creative need for grain

From an encoding point of view one might have thought that the best route would be to completely remove the grain or noise. Unfortunately material which has been processed like this loses it's "film look" and this may be contrary to the creative intentions of the Director or Director of Photography. Often a certain level of grain may be associated with a particular period, mood, or time of day. So a balance must be sought between the creative intentions of the Director and the technical needs of the encoder in order to achieve the desired result in the domestic viewing of the disc.

The relationship between grain intensity, encoded bit rate, and perceived quality

In the research phases of this white paper Digital Vision and Sonic undertook a variety of tests to analyze the actual effects of different levels of film grain upon the encoding process. One hypothesis we explored was whether a clear relationship could be defined between grain intensity and the bit rate used during encoding. For example, if a scene is proving problematic for the compressionist, is increasing the bit rate by 1Mbps comparable with reducing the grain intensity by 10% perhaps?

In the table below, encoding tests were performed using the AVC codec in CineVision at several different bit rates on three versions of the same piece of 1080p24 source material. The three sources had been processed to contain different levels of grain. The PSNR values were then recorded for each as one measure of quality. See *Appendix A* for a definition of PSNR and its limitations.

Encoded Target bit rate (Mbps)	PSNR from original source (dB)	PSNR from source with medium grain reduction (dB)	PSNR from source with coarse grain reduction (dB)
5	41.8	44.3	47.4
15	43.2	45.7	48.8
25	44.3	46.8	49.7
30	44.8	47.3	50.2

At first sight these results are staggering. In audio terms 1dB is generally regarded as the smallest discernible difference in level. However, in encoding terms, if we encode the same source twice and achieve a 1dB difference in PSNR this normally means a quite significant difference in perceived quality. The encode of the original source at 30Mbps does not even achieve a PSNR level close to a 5Mbps encode of the same source when the film grain has been largely removed. As we have previously intimated, PSNR does not necessarily equate to perceived quality. However, it is a good reflection of how well the encoder is able to reproduce exactly what has been fed to it and these values at least reflect just how significant an effect film grain has on an encoder.

Unfortunately it isn't possible to illustrate in full the effects of significant amounts of grain and noise on encoders within the confines of this paper since the most important effects are on the temporal consistency from frame to frame rather than simply how good the grain looks on an individual frame. When encoders are presented with more grain and noise than they are able to reproduce it is the consistency in look of that grain from one frame to the next that tends to worsen first of all.

The important fact to consider is that encoding video with a high noise or film grain content will inevitably filter that grain and noise to some degree. The amount of filtering is somewhat proportional to the bit rate available for the encoding and the quality of the encoder in general. Consequently, there is a trade-off between applying more high quality noise and grain reduction during pre-processing, and keeping enough grain in the video to maintain a natural look that is faithful to the original. Too much grain in the source will bring unwanted encoding artifacts whilst too little will make the video look unnecessarily lacking in texture.

Clearly what is required is an encoding system that quickly gives the compressionist access to high quality grain and noise reduction on a scene-by-scene basis. Adjustments to that pre-processing can then be quickly compared with adjustments to encoder parameters to determine which is producing the best results for any particular segment of video. This is precisely the approach adopted by Sonic CineVision 2.0 and its integrated Digital Vision DVO toolset.

OPTIMUM WORKFLOWS

Figure 7 shows the typical workflow employed by the majority of DVD-Video production houses for video post-production and encoding.

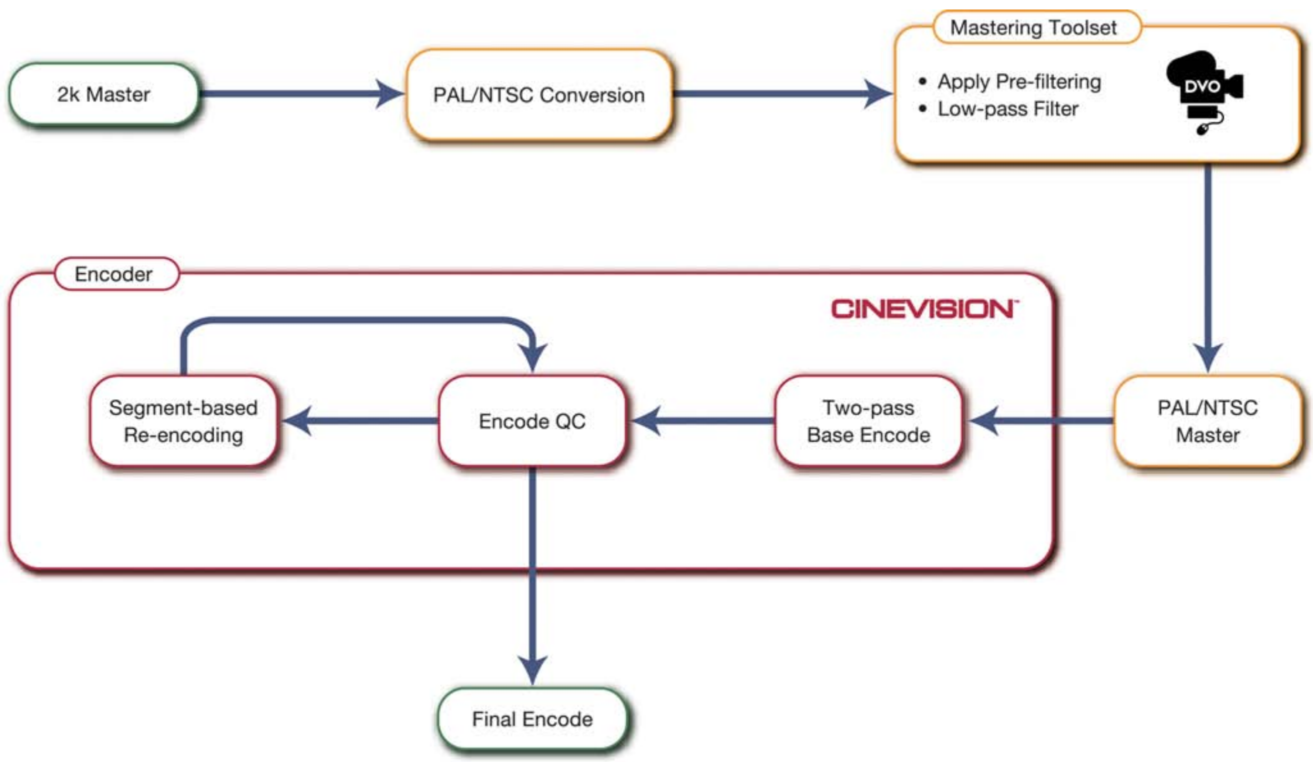

Fig. 7 - The typical mastering and encoding workflow for SD DVD-Video

In the majority of cases the post-production phases (creation of the graded 2K master and its conversion to a PAL or NTSC tape) are completed in a separate facility to the encoding processes shown. For that reason the mastering toolset of grain management, aperture correction and so on are often not available to the compressionist working on the DVD-Video MPEG-2 master. They're normally forced to rely upon simple low-pass filters provided by the manufacturer of the encoding solution they're using.

However, with the new challenges presented by encoding for HD formats as discussed in the previous chapter this represents a significant disadvantage. The following sections define new workflows that integrate both mastering tools and encoding tools to deliver significant benefits in encoding results.

Optimum Workflow for the Next-generation Optical Disc Formats

Figure 8 illustrates a new workflow for post-production and encoding that addresses the new challenges of encoding for HD optical disc formats.

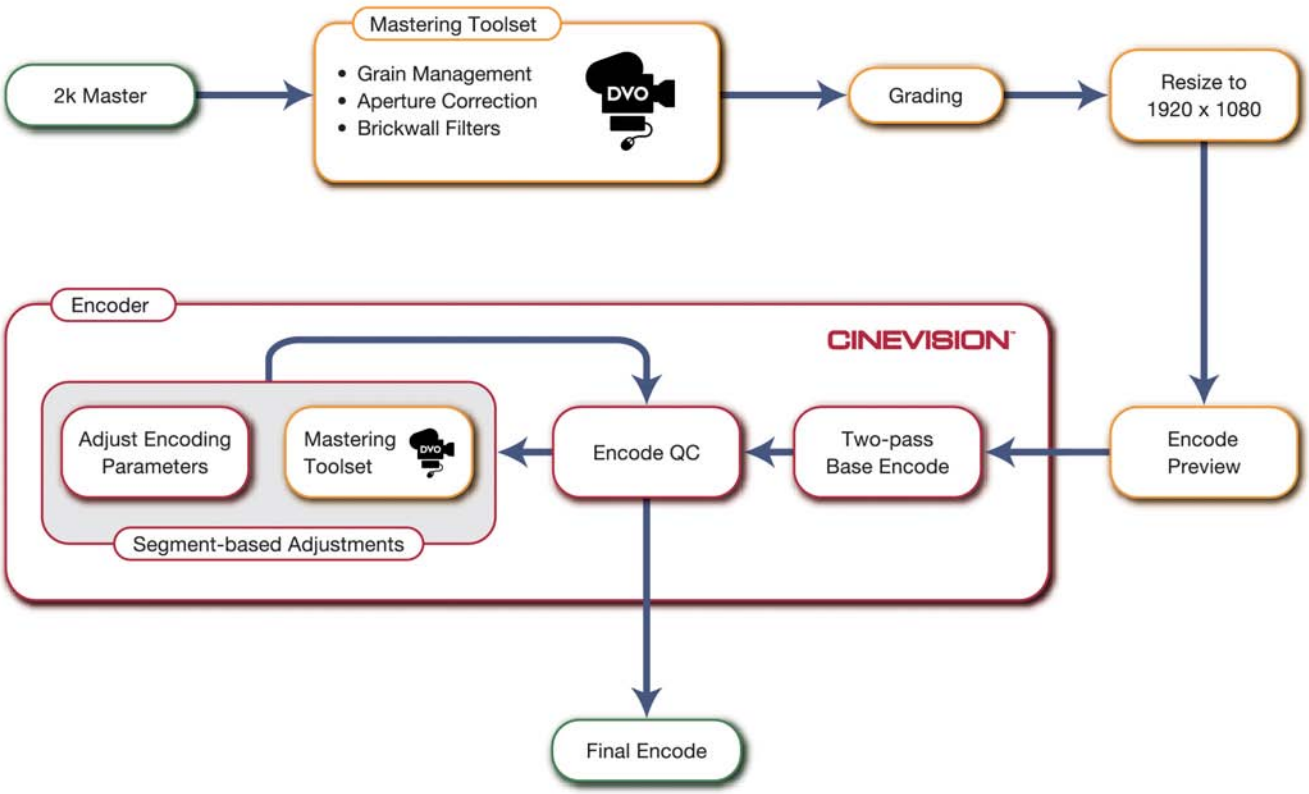

Fig. 8 - The integrated mastering and encoding workflow for the next-generation optical disc formats

The key characteristic of this new workflow is the integration of both mastering toolset and encoding tools and the added feedback loops between the two.

The initial stages of the chart show a typical post-production process with the addition of one important new stage – an inline preview of what the encoded version looks like. For the sake of practicality this inline preview may not necessarily adopt precisely the same parameters that are ultimately used in the encoding system (for example it may actually be single pass instead of two pass) but a preview using the correct codec and bit rate can be invaluable at this stage. This preview should allow the Mastering QC to identify, at a much earlier stage, fundamental problems with the grain management and correct them before passing the material to the compression team.

As the workflow progresses into the encoding system we see the other important change. Namely that it includes some access to the same tools that are used in the previous phase. Because on a scene-by-scene basis there is a good chance that a fine adjustment to the grain management or filtering may provide a better solution than an adjustment to the encoding parameters themselves, it's without question that providing these tools on the encoding station itself is invaluable.

Focus on Mastering

The notion of a universal master supposes that the onward operations are linear. In the past that may have been true but it was always a compromised one-size-fits-all concept. Now there are some requirements that challenge that notion, and some others that dictate an altogether different approach. Good technical quality for all viewing platforms, creative aims, commercial and censorship requirements as well as the ever-present desire to make changes to anything right up to the last moment, all mean that a new methodology for making deliverables and versioning is needed that uses a far more flexible form of master.

The look of images on the screen is a part of the creative aim and includes the control of color and grain. The process of grading tends to stretch the contrast and so increase the appearance of grain. Therefore the ideal workflow order here is linear in that grain management should be completed before grading begins.

The universal digital master is only a suitable solution if the client is prepared to accept one grade on all deliverables. In modern times they rarely are. If any of the deliverables are to be encoded or compressed then the image noise or grain can adversely affect the efficiency of the coding. For example, DVD, HD DVD and Blu-ray encoding all involve 4:2:0 sampling of YCrCb color space at 8-bit resolution. This “pre-compression” also has the effect of magnifying the noise and so increasing signal-to-noise ratio of the encoder input. Movies are coded for theatrical exhibition using a totally different process with 4:4:4 sampling of XYZ color space and wavelet-based JPEG2000 compression that is only intra-frame and here some grain is often desirable to give a filmic look. The upshot is that the grain, or noise, management of a ‘film’ master will need re-visiting to further reduce its appearance and to give a clean crisp look on video discs and in transmission. That process is best done before grading – effectively pushing the grain management back up the workflow to before the grade.

Different audiences may require, or be delivered, different versions of the production. This may involve straight cutting, in which case it may be possible to work directly from the universal digital master, but in many cases it will not. The increasing need for geographic specific product placement and censorship can dramatically increase the number of editorially different versions.

Clearly there are ample grounds for seeking a new workflow for post production and producing a new type of master. This should give access to any part of the program and allow changes to any aspect – even to work that may have been completed early on in the workflow and is now ‘buried’ under several further layers of further work in making the universal digital master. The traditional workflow does not allow this as the universal digital master is flattened, merging all layers and not carrying any of the data describing the original sources and all the processes that have been applied to it since.

The requirement is to store all the metadata associated with making a master, this is not just a shot log or an EDL, but all information about the footage used and all the detail of the processes applied to the media from grain management through all post. This should then allow re-visiting any stage of the operation at any time. At the end of the process the finished project can exist as the essence, the original footage including VFX shots, and metadata, all the information of the processes applied to the material to make a master. This requires a database to store and recall all the metadata and a storage infrastructure that allows fast access to all the media, as well as all the equipment and applications needed to process the media. While organizations with big budgets can build and support their own systems, most post houses do not have the resources, specifically the programmers and engineers, needed to create such bespoke solutions. The requirement is for a more off-the-shelf solution to meet the needs of modern post that includes a streamlined method of producing the versions and deliverables in an efficient way.

Multiple Format Masters

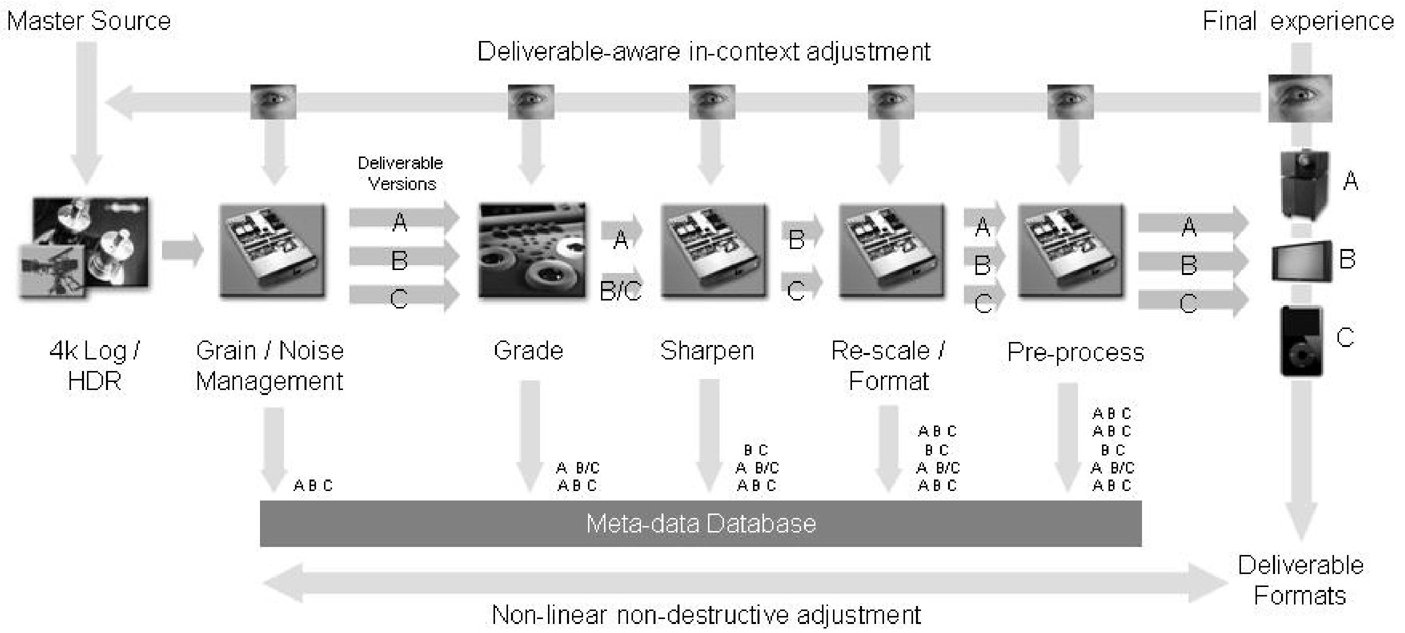

Fig. 9 - Multiple format masters

Optimal order for mastering/pre-processing steps

Typically most image processing works best on an unprocessed signal however this is not possible when multiple processes are to be applied. In addition, some artefacts in the image may disturb the algorithms of some processes and must be dealt with first. There will always be compromises, but this proposed order will provide the best result.

The DVO Dust algorithm works best when the dust and scratch artefacts have clean, sharp, edges. For this reason it is best to perform this process first.

Grain/noise reduction on the other hand works best when other artefacts like dust and dirt have been removed.

Aperture correction is better applied further down the chain as one side effect of this process is that it can enhance noise slightly. If the picture has already been grain/noise reduced then this is less of an issue. While sharpening of the picture may contradict a lower bit rate/higher quality encoding, the combination of overall sharpening plus the sharp cut-off at higher frequencies of a brickwall filter can result in a subjectively sharper image without affecting the encoding result.

A brickwall filter limits the spectral content of the picture and should therefore be applied at the end of the chain. Reducing the sharpness of the picture at the beginning of the chain would reduce the effectiveness of the other processes. While the other processes not only reduce the bit rate required in encoding, they also affect the picture quality positively. A brickwall filter on the other hand can almost be considered to “degrade” the picture, although also in a positive way for the encoding result.

The ability to preview the encoded output during the pre-processing and grading of the master provides the opportunity for the final effect of these different image processing steps to be clearly seen. Collaboration between Sonic Solutions and Digital Vision enables this possibility within the Nucoda Film Master and for decisions about pre-processing to be taken in confidence during the final grading whilst maintaining the important flexibility of a non-linear workflow.

Focus on Encoding

Figure 10 shows in further detail an optimized workflow for the encoding process itself which is more suited to the challenges of HD.

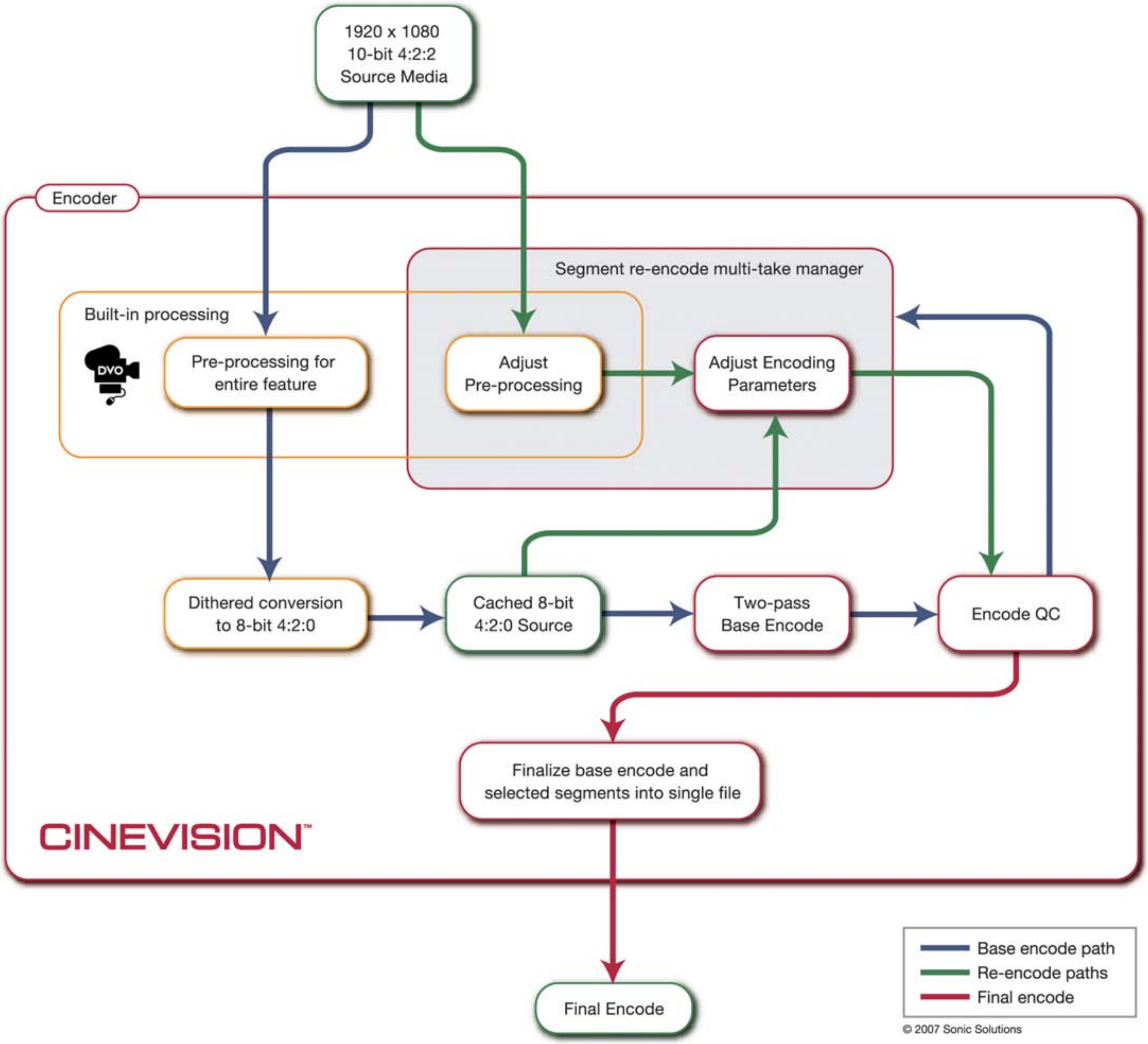

Fig. 10 - The re-encode flow: The source used for re-encodes (10-bit or 8-bit) depends on the nature of the re-encode. Re-encodes that require adjusted pre-processing will use the 10-bit source whereas re-encodes involving encoding parameter changes only will use the cached 8-bit source.

There are two particularly notable points to take from this workflow:

Caching of pre-processed source is important

Uncompressed 1920x1080 progressive material can have a data rate of up to about 160MB/second at 10-bit 4:2:2 (which means over 1TB for a 2 hour movie) but unfortunately for an efficient production workflow it can actually be necessary to have two versions of the source material on storage central to the mastering/pre-processing system and the encoding station.

As we've covered, effective pre-processing requires a sample depth above that of the encode itself (we've used 10-bit 4:2:2 as a common recommendation in this example) whereas the encode process can be most efficient in its destination

8-bit sample depth. By caching an 8-bit 4:2:0 version of the source after its pre-processing and dithering phases, segment-based adjustments to the encode parameters alone can be performed without redoing the same pre-processing every time. On the other hand, when the segment-based change incorporates a pre-processing adjustment the segment manager should pull the source from the original higher sample depth version as shown above.

With so many tools you need a multiple-take approach

Another important workflow enhancement that Sonic has developed for CineVision is the ability to create multiple takes or versions of the same segment and quickly switch between them to determine which has produced the best results. Once mastering tools like Digital Vision’s DVO technology are tightly integrated with an advanced encoding system like CineVision it quickly becomes apparent that an effective management interface is very important in order to fully realize their mutual potential.

This multiple-take approach also means that the operator is able to queue up a number of subtly different parameter setups for the same segment and create all of the encodes for them without needing to be with the system.

By making it so easy for compressionists to compare the effect of so many tools for problem sections results should be improved. CineVision also allows segment parameters to be saved as presets allowing successful setups to be recalled immediately.

SEGMENT RE-ENCODING

Segment based re-encoding has always played a very key role in professional encoding systems for DVD-Video. After a QC pass on the encoded video it allows short sections (segments) to be re-encoded to address problems without the need to re-encode the entire feature.

With the compressionist’s toolset now expanding a great deal more, segment re-encoding looks set to be even more powerful. Whilst a good encoder should be able to dynamically adapt to different material using all the tools at is disposal there is never any replacement for making sure that that toolset is also available on a more manual basis to an experienced compressionist.

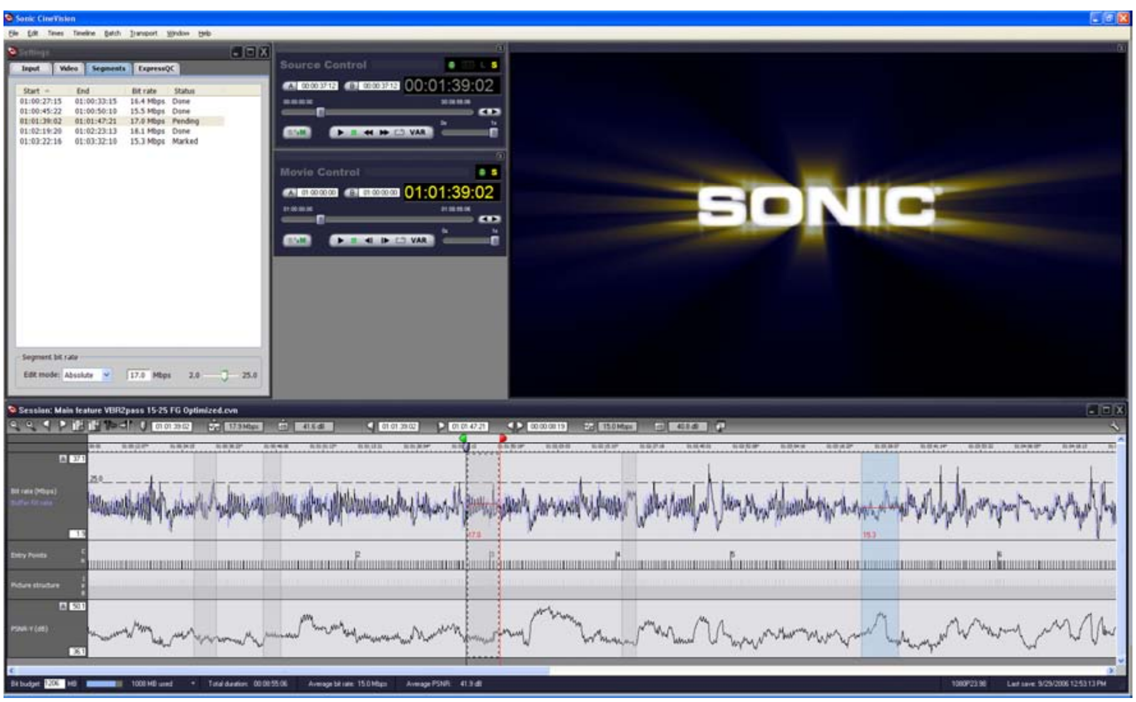

Fig. 11 - As well as best of class video encoding, Sonic CineVision delivers all of the professional tools a compresionist needs: extensive DVO filtering (from Digital Vision), exact parameter control for all three next-generation video codecs, and segment-based encoding.

Maintaining a Compliant VBV/HRD Buffer Model

Like all video codecs, the three under discussion here all have very stringent requirements for the manner in which data enters and leaves the buffer from which they are decoded. Indeed they are all defined by their decoder's requirements for a compliant bit stream rather than being specified at the encode side. Although somewhat similar in reality, MPEG-2 and VC-1 are defined by a video buffer verifier (VBV) model and AVC by a hypothetical reference decoder (HRD).

These models have a significant impact on the process of segment re-encoding. When replacing a section of video from a longer encode it is crucial to avoid buffer model problems like underflow and (in some cases) overflow. Particularly for AVC, the chances of doing a re-encode without any reference to the status of the buffer at the in and out points and obtaining a legal bit stream are quite low. Sonic CineVision features a unique approach to this challenge which allows high quality segment re-encodes to be performed at different bit rates without any requirement to re-encode portions longer than the user actually needs.

Beyond Bit Rate-only Re-encodes

For SD DVD-Video and the limited tools available in MPEG-2, an adjustment of the bit rate was to a large extent the only feasible solution available for improving quality. Now there are various tools that can often be more effective. We will discuss a few of them here.

Adaptive quantization

Adaptive quantization is not a new thing (and indeed some MPEG-2 encoders for DVD-Video featured some adaptive quantization modes). However, with the expanded toolsets available within codecs like AVC and VC-1 it is now more effective than ever.

Adaptive quantization is a heuristic approach to solving encoding problems by changing from a linear quantization approach to one that reduces the quantization when particular characteristics are detected (for example darkness). In general, adaptive quantization is not effective when applied to long pieces of video with a wide variety of characteristics (the encoder's default quantization approach should be most effective for this) but can be extremely effective for segments containing a specific issue. In addition to an adjustment to the quantization level applied to an area of a frame, adaptive quantization can also employ other tools in the codec's armory to optimize results for particular challenges.

De-blocking filter

As discussed back in *Next Generation Video Codecs* these filters are extremely effective at improving the perceived quality of many encodes but, in the case of AVC, can cause problems if the filter is set to too high a strength where high frequency content in the frames themselves can be softened. Similarly, if the filter is not strong enough, visible macroblock edges may become more prominent (particularly important at lower bit rates where quantization is higher anyway). Consequently, being able to control the parameters of these filters on a segment basis can help to optimize results significantly.

Black level normalization

We have already mentioned this tool in the section on the challenges of 8-bit 4:2:0 banding. Because we are very sensitive to variations between the levels near black, it can sometimes help to simply force all levels near to black to full black. Whilst clearly a crude approach, this can sometimes be the best way to fix problems of grain and noise or banding in near blacks.

Since the effect of this tool is clearly to remove information from the picture it would not be advisable to use it across an entire feature but it can be extremely valuable on individual segments.

Maximum quantization

A maximum quantization control allows the user to force a ceiling on the quantization level (QP) that the encoder adopts. This forces the encoder to use a lower (better quality) QP value if the user believes that the encoder is not assigning enough data to a particular segment of video.

Maximum quantization is a similar control to the minimum bit rate controls which are common amongst standard definition MPEG-2 encoders for DVD and is consequently not a new thing. Sonic CineVision displays quantization average, maximum and minimum levels for each frame so that irrespective of the bit rates involved the user can clearly see how "compressed" a frame really is.

Once again, in the majority of cases it would be inadvisable to use a maximum quantization across a long encode as it will constrain the encoder's ability to control the bit rate in general and negatively affect some sections. However, for specific segments of concern, it is a tool worth having access to.

Re-encoding with Pre-processing Adjustments

A key basis for the writing of this white paper has been the realization (through various testing by Sonic, Digital Vision and our customers) that access to high quality pre-processing tools during encoding and on a segment basis is now crucial in order to achieve the best possible results. In the sections below we discuss why some of these tools are of such significant benefit.

Grain

In *Next Generation Video Codecs* we discussed the principle new challenges for encoding for HD formats like HD DVD and Blu-ray Disc: Combating banding and encoding film grain successfully. It has also been shown how grain management can be a useful tool in both of these scenarios. Film grain can be filtered if the bit rate being used for encoding isn't sufficient to achieve a reasonable reproduction of it and can also be added as one approach to improving upon banding problems.

On a segment basis these tools are of great benefit for fine adjustments. Particularly for scenarios where bit rates for a project are lower than would perhaps be desirable there may be cases where for segments small reductions in grain intensity may be the best tool available to remove some objectionable compression artifacts.

It is also common for some banding artifacts to only reveal themselves in the final encode so the ability to correct them quickly with grain insertion during the segment re-encoding process is extremely useful.

Filtering

Low-pass filters have always been an important tool for compression and that is still the case for HD. The DVO tools within CineVision offer a variety of different filters but the principle behind them all is similar. We have discussed in detail how problematic grain can be to encoding algorithms since, for the most part, it can be classified as uncorrelated noise.

In a similar way, a lot of video material contains a lot of high frequency noise content which can be removed to give encoders an easier time. In more extreme cases, removing more and more high frequency content from a signal can avoid visible artifacting in the encoded frames since by its nature, encoding high frequencies consumes the most data. The obvious disadvantage of these filters is that the more they are used the "softer" the material becomes. However, access to high quality motion-compensated filters is once again highly beneficial in a segment re-encoding phase in conjunction with the other tools to correct objectionable encoding artifacts.

Aperture

Aperture (sharpness) correction becomes more important as soon as a greater number of pre-processing filters are available. The unwanted side-effect of more extreme grain management or low-pass filtering is an increase in softness (a reduction in sharpness). For that reason, use of aperture correction in conjunction with grain management and other filters can deliver the improvements without losing too much of the real high frequency content which a viewer expects from high definition video.

CONCLUSIONS

The primary goal of this paper and the research behind it was to outline the principle challenges that face those tasked with mastering, grading, and encoding video for high quality optical disc formats like HD DVD and Blu-ray Disc, to look at how they contrast with the production challenges of the past, and to provide practical solutions to them. In summary the conclusions we have reached are:

The 8-bit 4:2:0 barrier is now significant

The inherent limitations of 8-bit 4:2:0 video are now much more challenging, when trying to achieve good quality encodes, than they were for standard definition encoding. The most successful tools to address the problem are found in the pre-processing stage and include applying rounding, dithering and introducing new grain.

Integrated control of film grain is crucial

The largely uncorrelated nature of film grain or CCD noise presents a massive challenge to encoders. It was determined that it's important to apply the correct amount of high quality grain reduction in pre-processing so that the encoder reproduces it as closely as possible. Too much grain in the source will bring unwanted encoding artifacts whilst too little may make the video look unnatural and unfaithful to the original. An encoding system that allows segment adjustment of grain preprocessing is key to successful HD encoding.

Mastering workflows need to change

The single universal master tape delivery for SD DVD-Video production houses does not serve HD needs well. Modern mastering and grading needs to become 'deliverable aware'. We have outlined how effective management of metadata can allow multiple masters to share previous work avoiding repeating an intensive linear process.

Segment re-encoding is more important than ever

With the advent of more advanced video codecs like AVC and VC-1, the concept of segment re-encoding is now significantly more powerful. These codecs provide many new tools in addition to bit rate which may yield the best "fixes" for problem sections. Add to that the ability to make pre-processing adjustments on a segment basis and it is clear that an effective segment management system which allows the creation and auditioning of multiple versions is very beneficial indeed.

Bringing high quality mastering and encoding tools together reaps great rewards for HD production

It has been shown that providing compressionists access to advanced film grain management, filtering and other mastering tools within an encoding system, and conversely, those tasked with mastering and grading, a preview of what encoding would look like, will facilitate significant improvements to the final product.

APPENDIX

An Introduction to PSNR

Peak Signal-To-Noise Ratio (PSNR) is a method for calculating the quality of an encode expressed in decibels (dB). Luminance intensities across the encoded image are compared with those from the original source material and a signal-to-noise ratio is calculated from the difference. It's worth noting that most PSNR calculations do not take chrominance intensities into account and indeed the PSNR graph displayed in Sonic CineVision uses the luminance component (Y) only.

Most people are more aware of dB values with reference to audio rather than video. In audio a difference of 1dB is generally considered to be the smallest discernible difference that the human senses can detect. However, in video a 1dB difference in PSNR generally means a much more significant and discernible difference in video quality.

It is also worth noting that, even when using exactly the same encoding parameters, the PSNR for any given encoded frame can vary widely based on the material in the image. This limitation is averaged out over the course of an encode so if the encode is sufficiently long the average PSNR value becomes useful.

Higher PSNR values generally represent a better quality encode:

PSNR = 40dB

PSNR = 30dB

PSNR = 20dB

PSNR = 10dB

PSNR = 0dB

Fig. A1 - Example PSNR values with images

The only really meaningful application for PSNR is when comparing the encoding results of the same original source material. It should be noted that PSNR is just one of many quality metrics used for analyzing compression systems and that an increase in PSNR does not always infer an improvement in the "quality" a viewer perceives from watching the material.

ACKNOWLEDGEMENTS

This white paper was authored by Sam Orton-Jay and Chris Steele.

Sam Orton-Jay - Senior Product Manager, Sonic Professional Products Group

Sam Orton-Jay is Senior Product Manager for Sonic CineVision, Sonic's professional HD DVD and BD encoding solution. With over 5 years experience at Sonic, Sam has spent the past two years working closely with members of Sonic's High Definition Authoring Alliance and, with their help, has been instrumental in defining new workflows and functionality for encoding for HD DVD and Blu-ray Disc. Prior to working for Sonic, Sam spent two years as a software engineer for a professional audio editing system manufacturer. He graduated from the renowned Tonmeister (Music and Sound Recording) degree at the University of Surrey in the UK.

Chris Steele - VP of Product Development, Digital Vision

Chris Steele is a 20 year veteran of the film and television industries, working to develop tools in computer graphics, paint, compositing and animation. He joined Digital Vision in 2006 after holding senior positions at Parallax Software (Matador and Illusion), Avid Technology and Eastman Kodak (Cineon). He was most recently Director of UK Operations for OMG plc, having started 2d3 Ltd for OMG and brought the award winning camera tracker; boujou to market. He is currently VP of Product Development at Digital Vision.