

RGB HD- Efficiencies in the Digital Intermediate Process.

By Rick Harding and Chris Bone

Introduction

The growing dependence on the use of Digital Intermediate (DI) processes to preserve, enhance, and transform artistic intent to the cinema screen is well underway, and while there are still issues of standardization yet to be reconciled, the popularity and demand for DI processes among the creative community continues to grow at an exponential rate.

Although the current industry goal is 4K DI workflows, they are generally considered impractical at this point in time for a number of reasons. Therefore, we see more adaptable 2K processes applied to the DI in the interim as it will likely be some time before 4K processes are considered cost effective. Be it 2K or 4K, the prospect of data looms large over the post production industry and the question of the day is: "Will it be necessary to completely retool to a data infrastructure in order to compete in this new 2K world, or can existing 4:2:2 HD infrastructures be adapted to full band RGB HD 10 bit Log video workflows to achieve comparable and predictable results?"

2K data scanning – Taking a look at the numbers

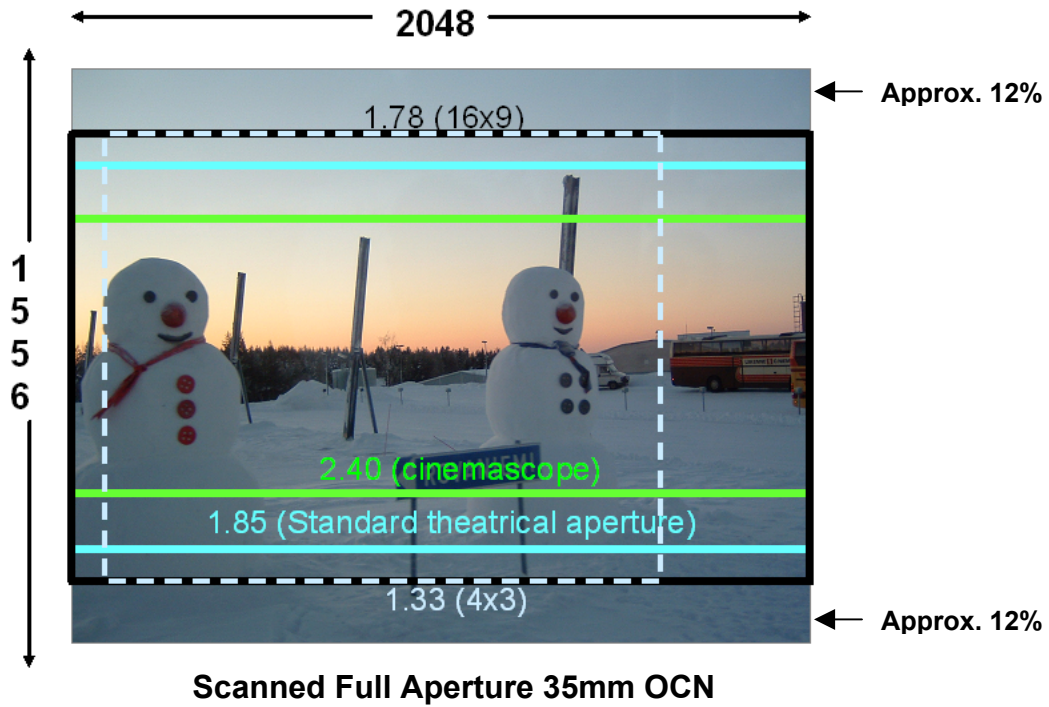
Although the term "2K" has been liberally applied of late to a number of digital acquisition and display devices, it is still generally understood that when scanning 35mm camera negative for DI applications, a pixel count of either 2048 x 1536 or 2048 x 1556 constitute a standard 2K file. These pixel counts manifest a 1.33 and 1.316 aspect ratio, respectively, and encompass the entire four perforation 35mm aperture of the Original Camera Negative (OCN).¹

Regardless of common framing reference when shooting full aperture 35mm (common center – top – bottom), digital capture of the entire OCN aperture yields flexibility in the DI process and provides a one-to-one data master of the 35mm aperture for archive. However, in actual theatrical exhibition only a percentage of this captured data will be displayed. Of course, if scanning an anamorphic film element, all of the information contained within the aperture will apply; but generally speaking when shooting spherical four perf. 35mm film an extraction will normally be required for proper exhibition aspect ratio.

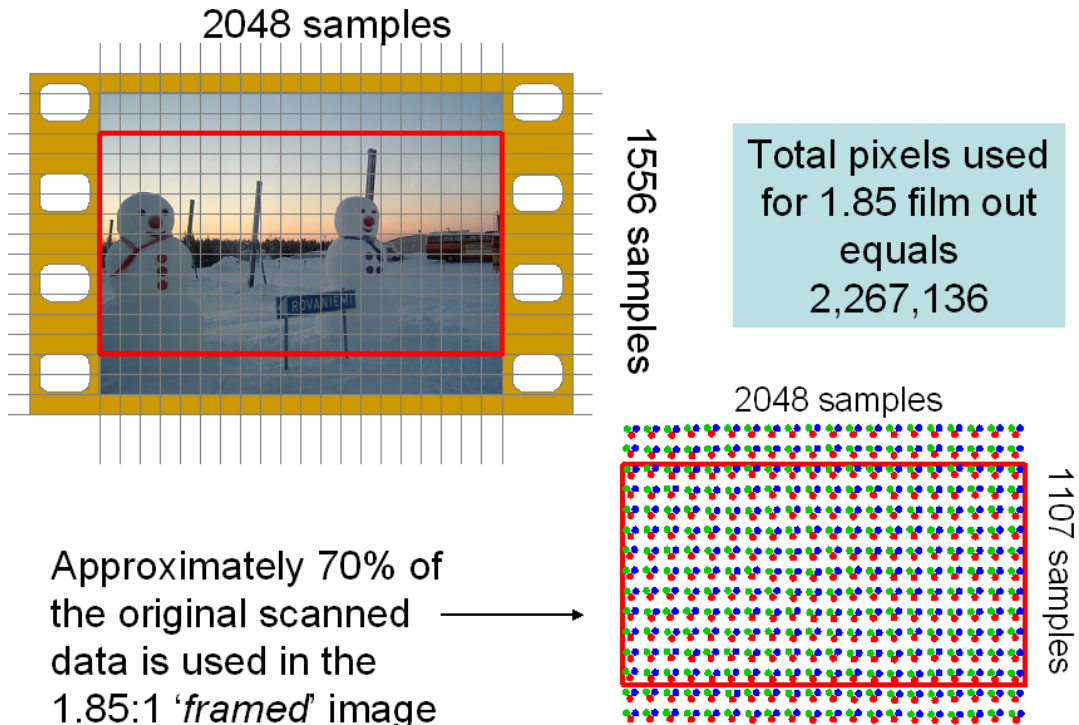
As the following full aperture 35mm (common center) film frame illustrates, none of the applied aspect ratios for both film and video exhibition will benefit from the full vertical 2K scan. This includes any 4x3 video elements, primarily due to compositional considerations and the probable presence of production equipment in the full aperture image. This being the case, it is fair to say that

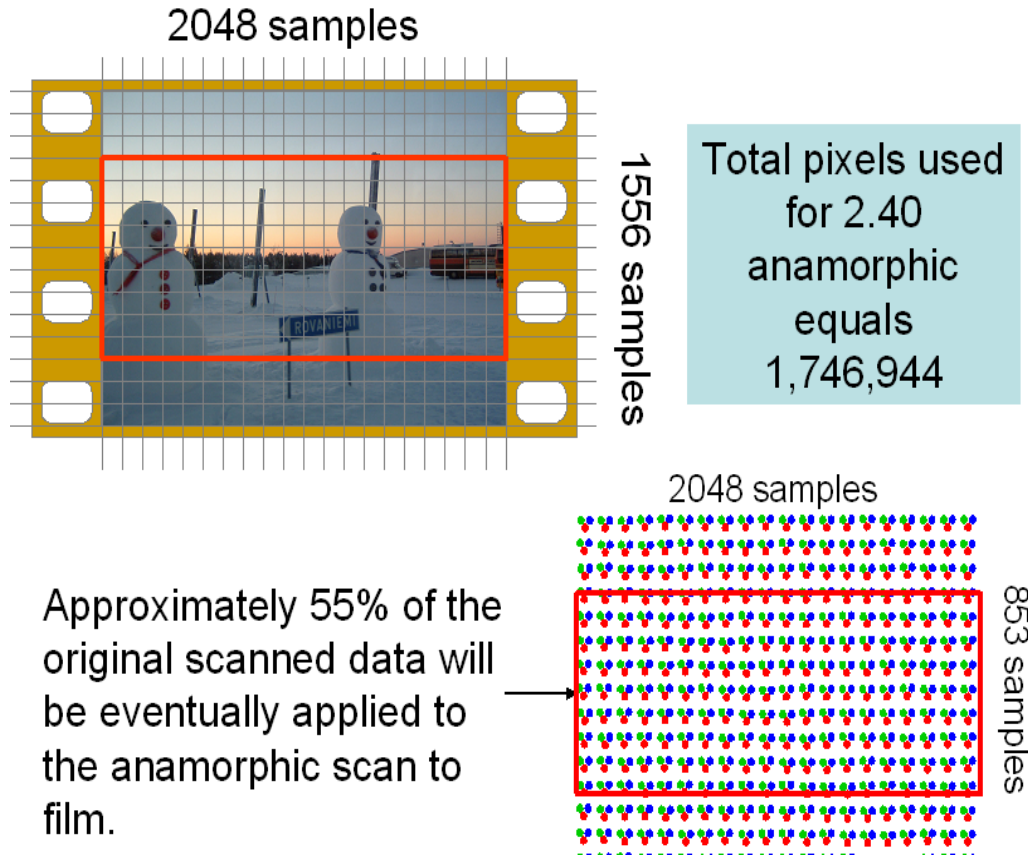
¹ Refer to SMPTE 59-1998 35mm 4-perforation Image Areas, camera Instrumentation and Full Frame Motion Pictures, Style C Specification (.981" x 735" – plus/minus .002).

that at least 24% of the originally scanned data will likely never be applied to any meaningful purpose other than archive.



To further elaborate, the following illustrations depict pixel count extractions from full aperture super 35mm for both 1.85 and 2.40 theatrical aspect ratios.





Considering these illustrations, it is fair to say that applied pixel counts are much closer to HD resolutions than that contained in the original 2K file. In fact the greatest differential between 2K and HD will be approximately 6.7% on either axis for 1.85 and 2.40 applications. This differential may seem significant, but the differential may be completely lost when recording back to film, depending on the film recorder.

When creating printing negatives from a finished data file, the file will usually be recorded to academy aperture 35mm film.² This ensures first generation quality to the printing negative and select print quality to the theater. However, academy aperture is designed to accommodate an optical track within the same physical dimensions of all 35mm film. Thus the image area of an academy aperture is significantly reduced from that of the original full aperture OCN. As a result, film recorders commonly scale the 2K image to the smaller aperture. This normally involves mapping the image to the entire academy aperture, or hard matting the image to a specific aspect ratio such as 1.78 or 1.85; often requiring an offset and extraction to align image center for both applications. This is accomplished through a variety of pixel count solutions for full aperture academy 35mm (1.37:1)

² Ref. SMPTE 111-2001 35mm printing made on continuous contact printers – exposed areas for picture and sound.

recording. Depending on facility variables, applied pixel counts to full aperture academy 35mm will range from 1806 x 1317 to 2048 x 1493. If recording hard mattes to the film, the vertical pixel sample will depend on horizontal resolution. For example, a 1.85 hard matte might be recorded from the original 2K file at any one of the following pixel counts, depending on the recorder; 1806 x 976, 1828 x 988, 1920 x 1038, 2048 x 1107.

Regardless, whether anamorphic, hard matted, or matted in projection, at the end of the day the film image viewed will not display the full pixel resolution contained in the original 2K file, and actual applied pixel counts will be less than or essentially comparable to HD resolution depending on exhibition aspect ratio. Further, considering Digital Cinema applications, the Digital Cinema Initiative (DCI) with consensus from all major MPAA studios, has recommended 2048 x 1080 as a standard for 2K theatrical delivery. Theatrical aspect ratios applied to the screen in this scenario would be 2048 x 853 for 2.40:1, 1998 x 1080 for 1.85:1, and 1920 x 1080 for 1.78:1.

Ultimately, when considering these applied pixel counts, the question arises: Is 1920 x 1080 RGB HD, 10 bit Log, a viable candidate for high-end Digital Intermediate workflows?

HD is NOT “TV”

It is not the intention of this paper to assert that RGB HD and RGB 2K data are equal. They are not. But “HD” does not equate to “HDTV” either (an assumption that some content creators make). HDTV like NTSC is what we receive in the home; a highly compressed 4:2:0 signal created specifically for the purpose of broadcast.³

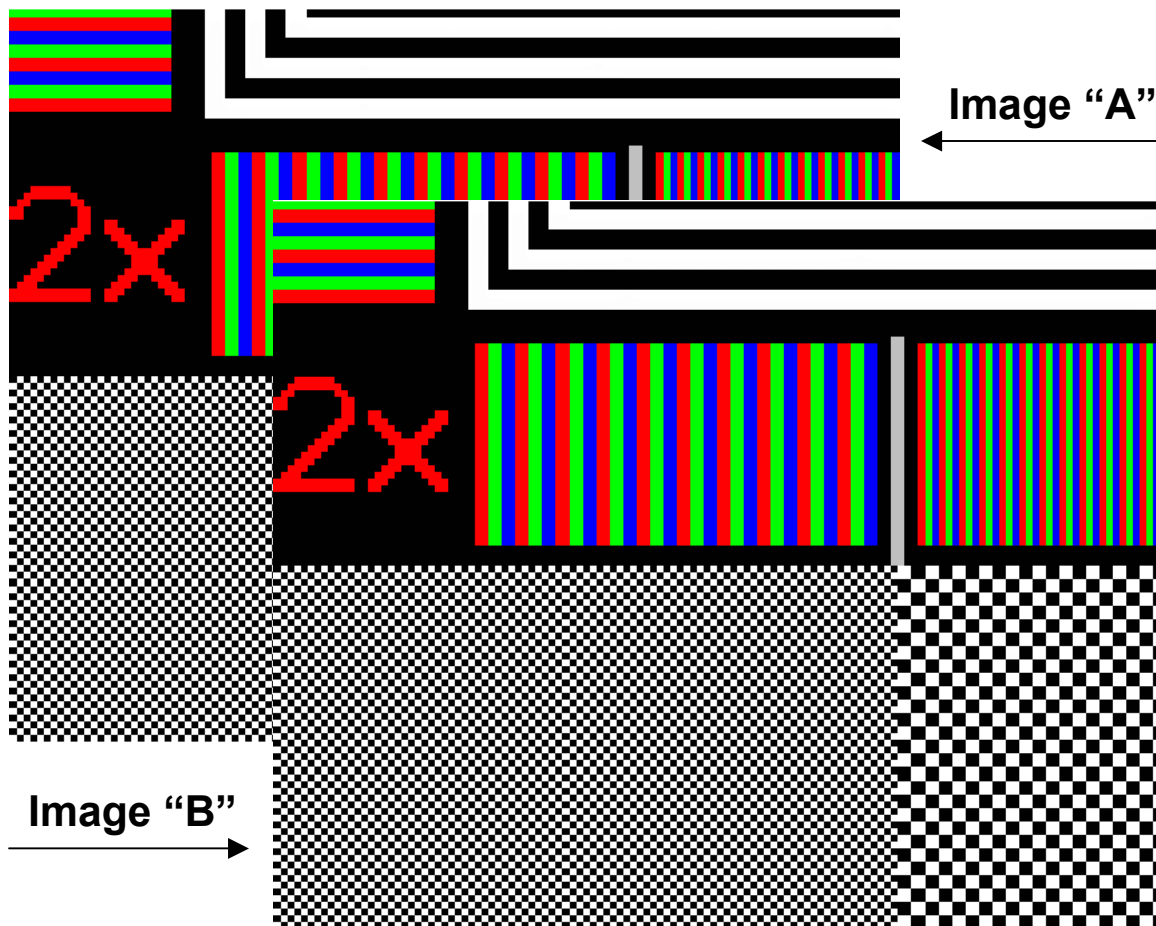
Looking at full band HD and 2K data imaging, we find much more similarity than disparity between the two. Indeed they are close cousins. As digital representations of the scanned image, both offer full band per channel RGB signals; both offer minimum bit depths of 10 bit (Log or Linear); both can accommodate metadata and other user bits; both can be applied to film or video product with a high degree of predictability and quality; both can be manipulated, monitored, and applied from servers in real time. However, with today’s technology only 10 bit Log RGB HD can be accessed and stored to tape in real time, via HDCAM SR™ technology, giving RGB HD an operational throughput advantage that promises time savings, and ultimately reduced project cost. Therefore, assuming comparable quality to the final film product (as our look at applied pixel counts suggest) throughput advantages alone make RGB HD workflows worth investigating as a viable, cost effective, alternative to working in 2K data.

³ See ATSC Doc. A/53B with current amendment

HDCAM SR – The Technology⁴

Created for the purpose of enabling full bandwidth RGB acquisition, HDCAM SR technology offers the post production community a practical bridge between conventional 4:2:2 HD video, and data operations. Capable of recording 10-Bit linear or log RGB, or 10-Bit extended range wide color gamut R'G'B' (per SMPTE Recommendation 1361) through an advanced codec, HDCAM SR recordings are virtually identical to their uncompressed source. The term “virtual” is used only because it is impossible to process a signal without affecting it in some way. Nevertheless, given the combined efficiencies of the HDCAM SR’s advanced chip set and the MPEG 4 Studio Profile (SP) compression scheme, it is “virtually” impossible to visually discern any significant difference between an HDCAM SR recording and its uncompressed source.

To illustrate this claim, consider the following images. Both images were extracted from 2Kfiles, and both images were blown up, 500%, in an attempt to simulate a theatrical sized screen comparison.



⁴ Information gleaned and extracted from SMPTE paper “Super Reality – Refining HD” by Philip C. Nottle, Hiroshi Kiriya, and Kaoru Urata, presented at SMPTE conf. July 2003, Sydney , AU

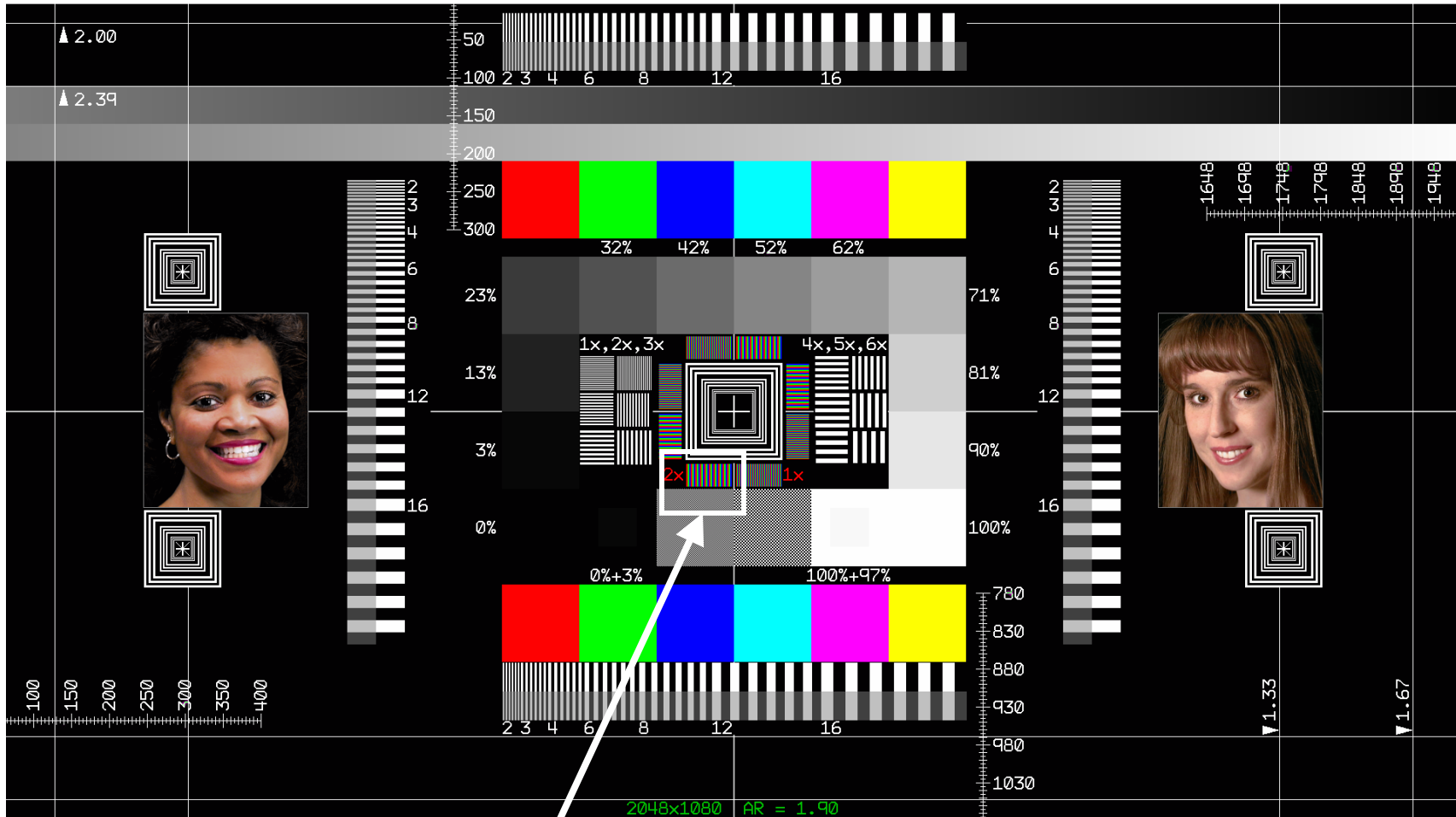
Another view of the same extractions:



Image "A"

Image "B"

1920 x 1080 cropped extract from 2048 x 1080 resolution chart⁵



Approximate area of sample extractions taken from both files.

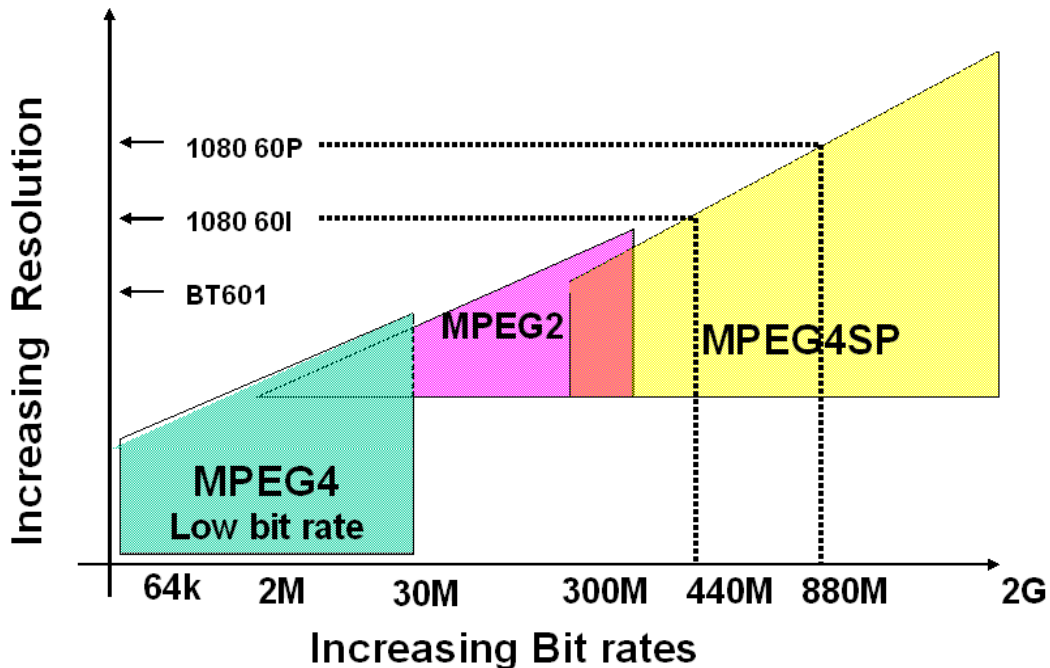
⁵ 2K file courtesy of the Digital Cinema Initiative. 2K file was derived from the original 4K Res. chart.

Starting with a 2K file, a 1920x1080 extraction of the entire resolution chart was taken, in server. This file was then converted to video and recorded to HDCAM SR tape. The RGB recording was then played back to disk, converted to a new 1920x1080 Tiff file, and recorded to CD Rom. Image "A" was then extracted from the new Tiff file created from the RGB playback. Image "B" was extracted from the source 2K file (2048x1080) that was also recorded to CD Rom. Both extractions were taken from their respective Tiff files in Photoshop (6.0) and blown up 500%.

Although this simple comparison is by no means a method for critical evaluation, and although it does not demonstrate HDCAM SR technology's remarkable proficiency in handling spatial information in the temporal domain with respect to moving images, it helps make the point that the HDCAM SR codec is (for all intents and purposes) transparent.

MPEG 4 Encoding

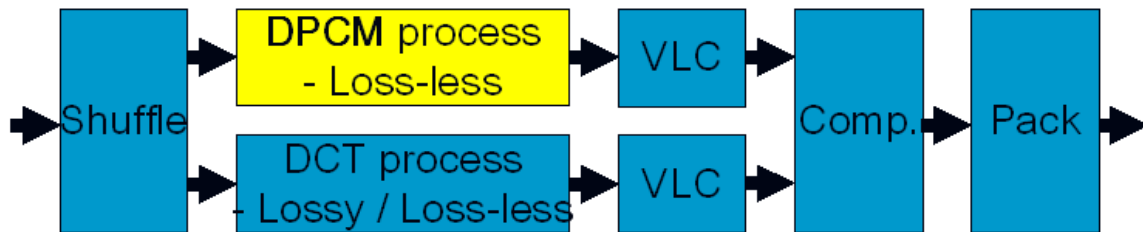
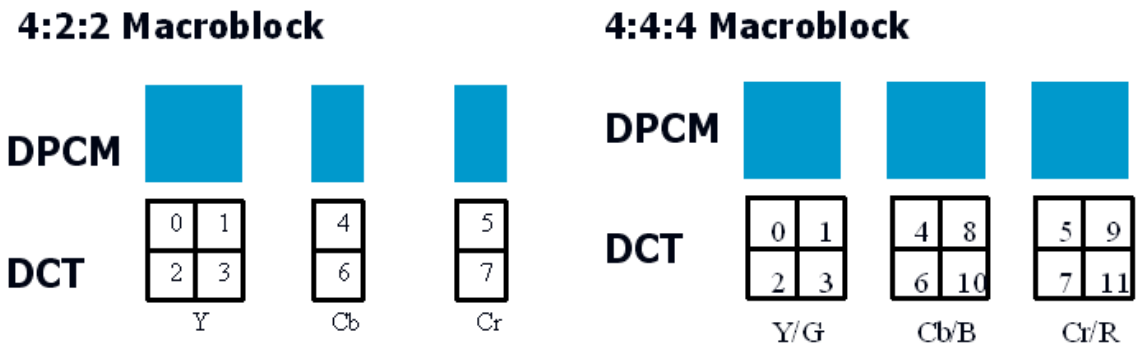
At the heart of HDCAM SR technology is the encoding/decoding chipset and the MPEG-4 Studio Profile (SP) compression strategy. MPEG-4 SP was created to meet the demands of high resolution imaging. It is not constrained by GOP structures and is scaleable in its pixel count (SD, HD, Data) at bit depths of 10 and 12 bit, with RGB color resolution, and head room potential up to 2 gigabits per second.



To fully realize MPEG-4 SP's potential, HDCAM SR technology uses both intra-frame and intra-field compression for progressive and interlaced recording (respectively) and two parallel processes; Differential Pulse Code Modulation (DPCM) and DCT, to achieve optimum results.

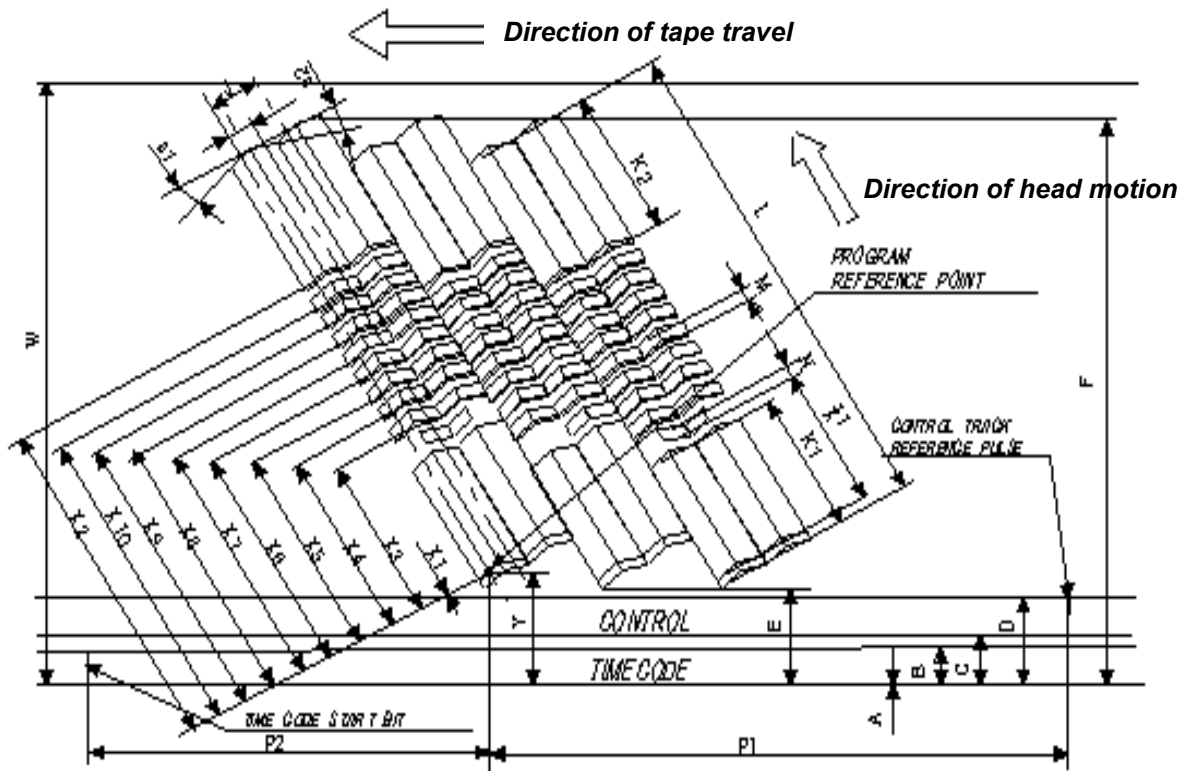
DPCM processing separates data from structure. Instead of encoding a pixel's intensity value, it predicts a value based on the value of one or more surrounding pixels. This is an ideal method of predictive coding. Data is grouped in macro-blocks and the first pixel value of the block serves as the reference level. The next pixel's value becomes the differential to the reference. The following pixel values are then the differential of the preceding pixel. There is no loss of original picture information, only a reduction in the data required to carry the information. Thus DPCM could be considered a virtually loss-less compression system.

When DPCM is used in parallel with the DCT process, both processes receive the same shuffled data. The DCT process operates at the 8 x 8 block level while the DPCM operates at the 16 x 16 macro-block level. The DPCM process is very efficient at encoding large flat areas of picture. When the outputs of the DPCM and DCT processes are compared, many times the bit rate from the DPCM process will be smaller. When this occurs, the comparator will choose the DPCM data stream for output data. This will leave the DCT with a high bit rate capacity allowing the quantizer in the DCT process to operate at lower compression, thus improving overall picture performance.



Head Drum and Tracking

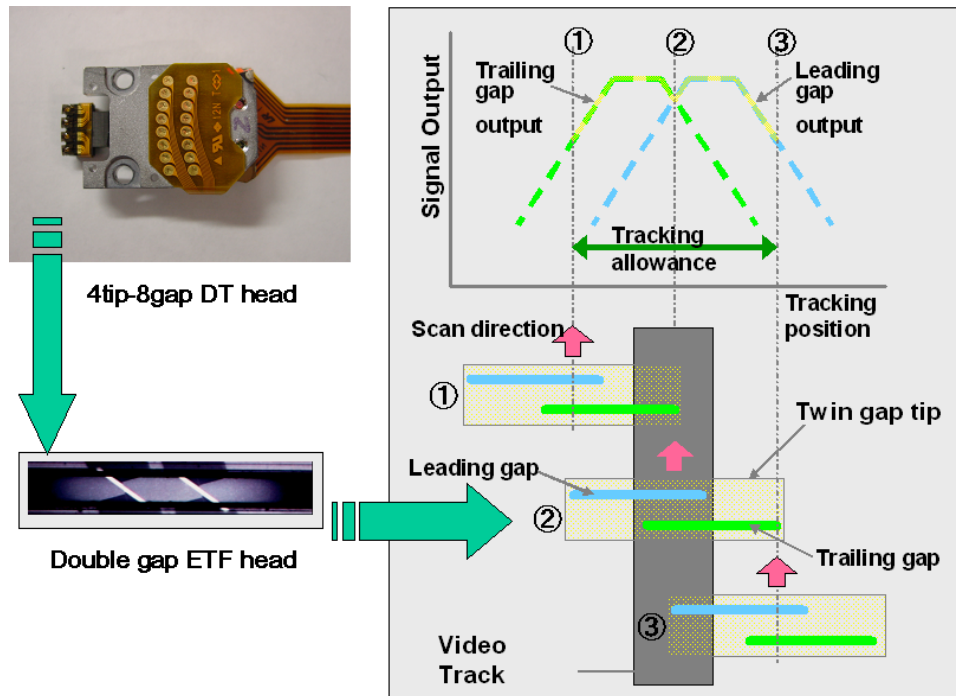
Apart from the chipset and compression scheme, HDCAM SR technology also incorporates the latest advancements in head drum design. As illustrated below, the helical scan records two video segments on either side of the audio segments, and each field (or frame) consists of 12 tracks with a pitch of $13.2\mu\text{m}$ and minimum recorded wavelength of $0.294\mu\text{m}$.



HDCAM SR™ Footprint

This very precise record capability allows a video data rate to tape of 440Mbps in standard record speed and 880Mbps in 2x record speed (SRW1 only). To provide accurate retrieval of the data from the narrow tracks, a twin gap head was developed. The use of twin gap head tips with offset gaps, give an efficient means of increasing tracking tolerance and providing an overall improvement in Bit Error Rate.

As both gaps travel across the recorded tracks in playback, each record track is effectively traced twice with each head pass. In comparison to conventional systems, this unique method of scanning allows a wider tolerance in head-to-track tracing by effectively providing an overall tracking width of over $21\mu\text{m}$. This not only provides accurate retrieval of the data but it eliminates the potential issues associated with machine interchange.



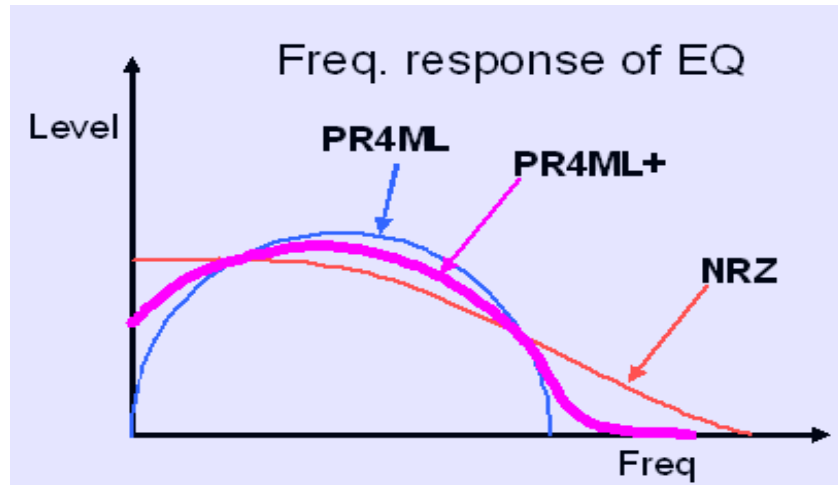
HDCAM SR™ Tracking Technology

Advanced Equalization

There are three strategies used in the advanced equalizer to maintain optimum output. Enabled by advanced chip design and supporting circuitry, Nyquist type “over-sampling” is applied to the playback RF signal. This technique reduces aliasing distortion by ensuring the highest frequency signal recovered from the tape does not come close to the sampling frequency of the Analog to Digital Converter.

Second, the Auto-EQ circuit uses an innovative method to achieve signal equalization. Part of this new method is a process called Least Mean Square (LMS) and a “Transversal Filter” with fifteen gain taps. The Least Mean Square process and the Transversal Filter work in parallel to control which of the 15 Transversal Filter gain taps the equalizer uses. This method provides more than sufficient range for the EQ process and is a big improvement over previous analog processes.

The final strategy in the advanced equalizer is the improvement to the Partial Response 4 ML circuit. This uses a combination of PR4ML and NRZ (Non Return to Zero) to change PR4ML characteristics and yield better response at both low and high frequencies. This response is similar to the characteristic of the combined tape and head response, thus optimizing data recovery and improving error rate by as much as three times.

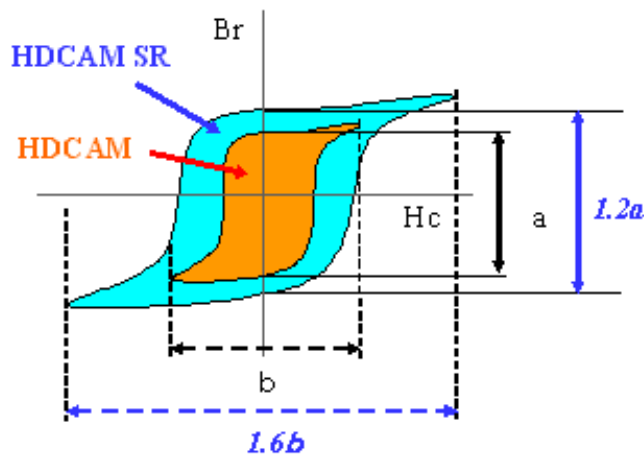


Advanced Equalization:
Error rate is improved 3 times over PR4ML alone.

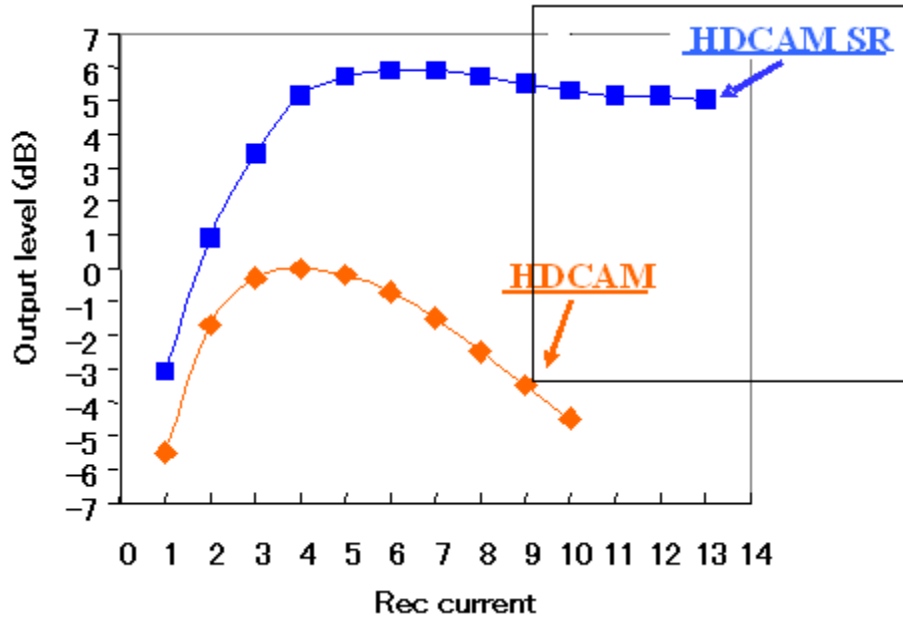
The Tape

The final element in the HDCAM SR™ advanced equation is the tape stock. Developed specifically to meet the saturation demands of this new format, state-of-the-art technology has been applied to every aspect of this new media. Metal particles were selected that are only half the length and $\frac{3}{4}$ the diameter of those used in HDCAM® format tape. This makes for a dense, smooth, pliable, recording surface that is capable of receiving high energy record signals, while maximizing tape-to-head contact.

The B-H curve of the HDCAM SR™ tape is substantially wider than HDCAM™ with better than 50% increase in coercivity. Additionally, output signal vs. record current is improved over HDCAM by 6dB.

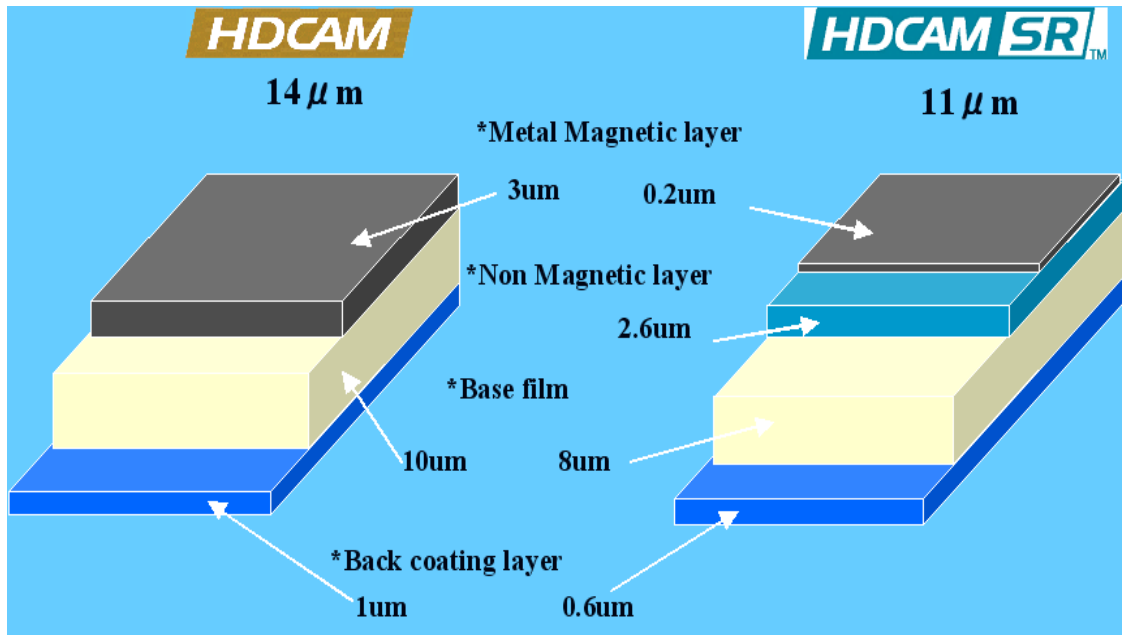


HDCAM SR B-H Curve



Six dB increase over HDCAM Tape

Finally, an inventive new insulating layer has been applied between the metal particle layer and the film base material. This not only guards against signal bleed-through, but it contains a timed release lubricant that keeps the media pliable over time, extending potential shelf life and improving long term playback performance.



HDCAM SR™ Tape Construction

Building a Bridge – Repurposing Existing 4:2:2 HD Infrastructures

The post production industry has seen many technological innovations over the past 25 years, but none, not even the advent of non linear editing, has impacted the industry to the same degree that the Digital Intermediate will. Core post production businesses such as conventional telecine, on-line editing, standards conversion, frame rate conversion, and dubbing are severally threatened as the technologies that make the DI possible continue to evolve, powered by Moore's Law.

Given that the DI produces a finished theatrical product in file form, all required distribution products can be derived from that finished product file. Software is now capable of doing what has traditionally been accomplished with hardware. Consequently, conventional HD service providers planning to compete in the new paradigm are faced with a dilemma - Invest significantly more capital into data infrastructures now, or maximize current investments to compete with emerging 2K data infrastructures, while evolving over time to a data infrastructure that will ultimately support 4K workflows. Either way, there is risk. But given facility client relationships and their potential to educate their markets, the latter option may well prove to be the most practical, cost effective, road to the future.

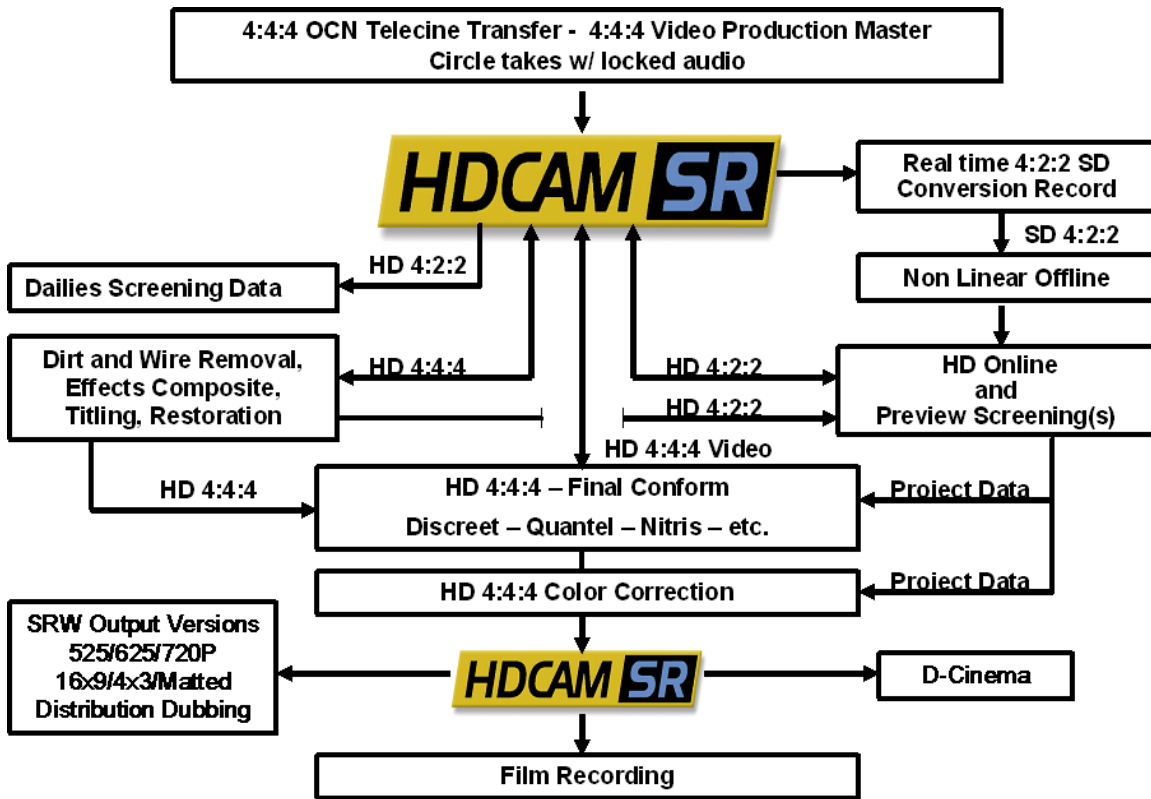
The following diagram represents a simple workflow model applying 4:4:4 10 bit Log signals to an existing 4:2:2 infrastructure for RGB DI. Generally speaking, RGB capable telecine, color correction systems, and workstations, such as SGI, Quantel, or Avid, already exist in most conventional HD facilities. Items not necessarily found in most facilities might be 2K digital projection (2048x1080), film recording (which can be outsourced), Color Management Systems, and 4:4:4 digital to analog conversion.

The RGB I/O called for in this diagram is HD-SDI Dual Link 292M Interface for 1920 x 1080 picture raster, as extended for 4:4:4 R'G'B'/10-bit application per SMPTE 372M-2002⁶. This versatile interface transmits data through link "A" in 4:2:2 sample ratio of all G' channel samples, plus even-numbered samples from the B' and R' channels, and in link "B" at an 0:2:2 ratio of odd-numbered samples from the B' and R' channels. Although link "A" is characterized as 4:2:2, it is not used in this instance as Y'C'BC'R, and as such, it cannot be monitored correctly without interposing the additional channel data carried in the 0:2:2 sample. To meet the standard, both "A" and "B" link must be within a time tolerance of 40 ns from the sender, and consequentially within the tolerance to the receiver. Fortunately, current routing systems for HD SDI and the EQ clocking and re-clocking capabilities of most current devices, allow for these two signals to be routed through existing routing networks to this tolerance without need of modification.

⁶ Refer to SMPTE 372M-2002. Other channel mapping such as Y'C'BC'R 4:4:4 MSB (dropping the 2-bit LSB) of 12-bit signals could be carried owing to the scope of the standard. See also SMPTE standards 274M and 352M.

Critical monitoring in RGB HD is best accomplished with either the recommended 2K projection system or a Sony BVMF24 RGB HD-SDI Dual-Link monitor. NAB '05 will likely feature new RGB monitoring options, however, Sony BVMD24 and BVMD32 HD monitors may also be used effectively for critical viewing using analog feeds. OmniTech makes an RGB waveform monitor that provides a full band RGB D/A converter, which lends itself well to this purpose.

Monitoring 4:2:2 signals that have been derived from the original 4:4:4 is also a viable option for all but the most critical 4:4:4 viewing requirements. The HDCAM SR internal format converter is a good source for this 4:2:2 signal. To make this approach even easier, various manufacturers are working to release new 4:4:4 A/D converters which will also output a 4:2:2 HD SDI derivative of the RGB input signal.



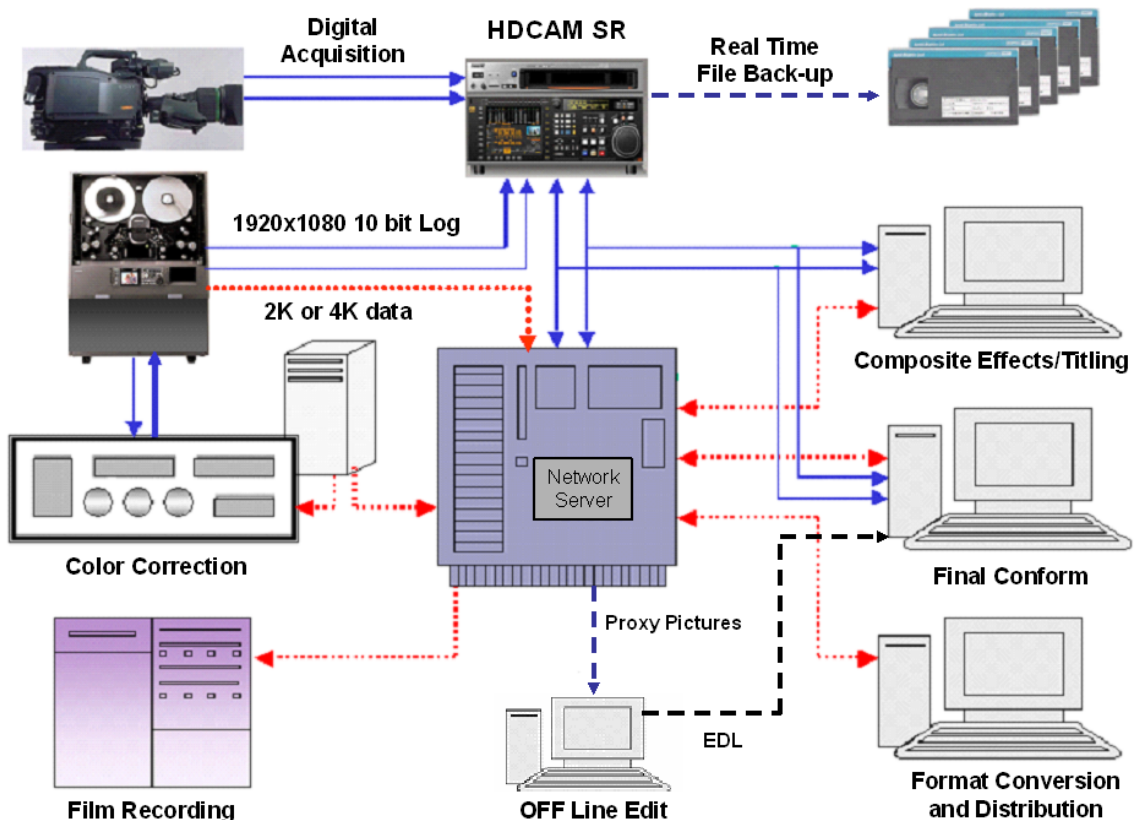
RGB Linear Workflow

Apart from the obvious implementation advantages, this workflow model is not limited by finite server limitations and does not require vast secondary RAID storage to avoid throughput bottlenecks when transferring server data to tape for back-up and archive. These advantages will directly affect overall project throughput and ultimately user costs, making project pricing a factor in competitive operations.

Non linear Systems

RGB workflows also offer throughput advantages to those already employing non-linear data systems. As previously asserted, it is highly unlikely that any aspect ratio beyond 1.78 will be required from the 2K file for any display requirement. Given this fact, and given the additional fact that the OCN is normally considered the ultimate archival element, an argument can be made for RGB applications in non linear data applications.

By reducing file size to a 1.78 extraction (2048 x 1152) the file size is effectively reduced by 26%, and server space is suddenly expanded by 26% allowing more material to server and ultimately the possibility of a real-time data dump to tape in the form of scaled RGB video to the HDCAM SR format; both conditions contributing to greater throughput potential by reducing file size to a 1.78 extraction.



**Simple Non-Linear Workflow Model
Using 1920 x 1080 RGB as Working Files**

Scaling the 2048 x 1152 to 1920 x 1080 reduces file size even further.⁷ Ultimately, working with scaled 1920 x 1080 files will net an additional 6% server capacity and allow real-time 10 bit Log data movement from server to back-up tape storage throughout the course of the project.

In 4:4:4 mode HDCAM SR technology is capable of recording up to 110 Mega Bytes per second (880Mbps mode/SRW1/compressed 2:1). Double that and you have the equivalent uncompressed volume. Moving this much uncompressed data to SAIT, DTF, LTO, or other data storage options would take more than seven times longer. Assuming 440 Mbps recording with SRW5000, HDCAM SR technology still offers data rates to tape at nearly four times that of the fastest data tape options currently available.

Once the project is complete, and ready for film recording, the 1920 x 1080 files can be seamlessly rescaled to 2048 x 1152 for film-out and archival storage. Or the 1920 files can be taken directly to film with comparable result (as illustrated). Either way, film record time will be significantly reduced over full aperture recording and may yield further project savings.

As our thinking toward conventional post flows change with technological innovation, so must our thinking with respect to film printing elements. Some may consider working with matted data elements a compromise of sorts, thinking hard matted film elements are somehow inferior to full aperture printing negative. In the new DI paradigm, printing elements should be viewed similarly to the various video masters created from archival sources specifically for video distribution. The theatrical printing negative is no different than these video masters in that it is created solely for the purpose of creating release prints for theatrical exhibition in either 1.85 or 2.40 aspect ratio (generally speaking).

As for the original full aperture data file, this along with the OCN becomes an archival element. As the negative may be ultimately cut, the original full aperture data file may also be conformed to the Edit Decision List (EDL) and Color Decision List (CDL). Any 1.78 matted effects shots would insert into the full aperture archival file to create the finished product. This has been a common practice in film for many years, and effects rendering may be limited to applicable exhibition aspect ratios regardless (likely 1.78).

If deemed necessary, an inter-positive (IP) film element can also be made from the conformed full aperture data file. This would serve as an additional archival back-up to the conformed full aperture file and provide a relatively inexpensive printing resource for the purpose of striking new inter-negatives (IN) for future

⁷ By cropping vertically to 1152 from a common center, followed by simple constrained proportion re-sampling of the horizontal and vertical dimensions, again with a common center, 2048 x 1152 then becomes 1920 x 1080

printing requirements. A cut OCN, and/or a conformed full aperture data file should be considered more than adequate to meet any future requirement. In fact, they could rightly be viewed in similar terms as conventional original cut negative and IP.

Conclusion

It was not the intention of this paper to comprehensively cover all aspects of the Digital Intermediate process. Issues such as color space and color management options, and the variety of scanning and recording variables that exist from one facility to another have little bearing on the point presented. It is, however, the purpose of this paper to point out a viable and practical business option for service providers interested in expanding their business opportunities in the new paradigm.

As the footnoted evaluation of image resolution in 35mm theatrical presentations suggest⁸, digital scaling from 2048 (h) to 1920 (h) and laser recording direct to 35mm Inter-Negative (IN), should yield at least comparable MTF results to release prints created through conventional methods of optical reduction from OCN or Inter-Positive (IP) to Academy aperture 35mm IN using industry standard Nikon, "Nikkor"⁹ lenses. This has been further corroborated in practice as the industry has produced many high profile and successful film products over the past few years by applying 1828 horizontal pixels back to film. Consequently, if it can be concluded that the application of 1920 horizontal pixel rates to film produces comparable quality results to accepted practice, then one can surmise that full band 10 bit Log RGB HD workflows have a place in the new Digital Intermediate world.

Ultimately, the Digital Intermediate will prove to be a double-edged sword. On the one hand, it provides great creative flexibilities and improved products. On the other, it consolidates workflow and threatens core businesses that have supported the film and video industry for the last quarter century. As the industry embarks on the road to 4K, it may prove a perilous trip. However, those who are capable of effectively adapting existing capital investments to maximum benefit while working toward a certain future, over time, will best survive the journey.

⁸ "The Image Resolution of 35mm Cinema Film in Theatrical Presentation", by Vittorio Baroncini, Henry Mahler, and Matthieu Sintas, SMPTE Motion Imaging Journal, Vol. 113 No. 2 & 3, pp. 60-66, February/March 2004. "Letter to the Editor" (in rebuttal), by R. Morton, A. Cosgrove, and A. Masson, Eastman Kodak Company, SMPTE Motion Imaging Journal, Vol. 113 No. 4, p. 102, April 2004.

⁹ 150mm 1:1 lenses (used with calibrated variable projection and camera plane offsets). Note: 95mm 2:1 lenses are used for printing applications such as between 35mm:16mm motion picture film. Each lens has an MTF that is superior for its designed purpose over the other. Information sources: Alan Fridge, Electronic Systems Design Consultant, Pasadena, California; Vince Roth, currently with FotoKem Film and Video in Burbank, California.