HDTV and Film — Digitization and Extended Dynamic Range

By Laurence J. Thorpe

High-definition television (HDTV) is rapidly evolving and film continues to improve. The ability to freely employ either imaging medium while ensuring its high-quality integration in HD post-production promises a new creative flexibility. Of the many attributes contributing to a final high picture quality, dynamic range remains critical to achieving a satisfactory matching of images (by intercut or blue-screen compositing). HDTV charge-coupled device (CCD) cameras have arrived, already endowed with a total dynamic range latitude equal to that of the best motion-picture film. However, the rapid move to HD digital recording and processing raises questions regarding the optimum quantization of the CCD analog output video waveform. This article discusses the complexities of this issue, the tentative decisions already adopted by HDTV equipment manufacturers, and the importance of the studies on this topic currently ongoing within the SMPTE.

An earlier paper presented at SMPTE[1] dealt with the issue of dynamic range for both film and HDTV origination. It attempted to separate issues relating to the image electronically "captured" (in terms of video dynamic range) by the HDTV camera from that which might ultimately be portrayed in terms of visual contrast range on a given HDTV display. There can be significant differences between the two.

In particular, the displayed visual contrast range can be greatly variable, being a complex function of:

- Display technology used
- Technical alignment of the display (brightness/contrast)
- Ambient lighting conditions, within which the display is viewed

The same can be said to apply to film. The wide dynamic range that might be captured by the original negative film (9 to 10 *f*-stops of "latitude" not being untypical) may ultimately result in a poor reproduction of visual contrast range in a given cinema, as shown in Fig. 1.

The degree of that degradation is a function of:

- Subsequent photochemical processing of the original negative-to-positive print (particularly dependent upon the number of stages of development and quality control at each, for a final release print)
- Quality of the film projector (notably flare)
- Ambient light of the viewing environment

There is little control that might be exercised over viewing conditions for the consumer at large. For professional viewing, however, a great deal has been examined, written, and recommended.[2]

In the case of HDTV, it is to be hoped that technological developments and an improved discipline in setting up the display environment will capitalize on the full potential of the system, particularly when attempting to exploit the striking capabilities of the HD CCD camera. As has been pointed out,[3] the psychophysical sensation of image depth, which can impart an almost three-dimensional experience to the large-screen HDTV image, is very much a function of the displayed contrast range. Contrast range is one of the vital "multidimensions" of the displayed HDTV image and, thus, an important contribution to the new sense of reality. However, as will be shown, the issue of digitization of

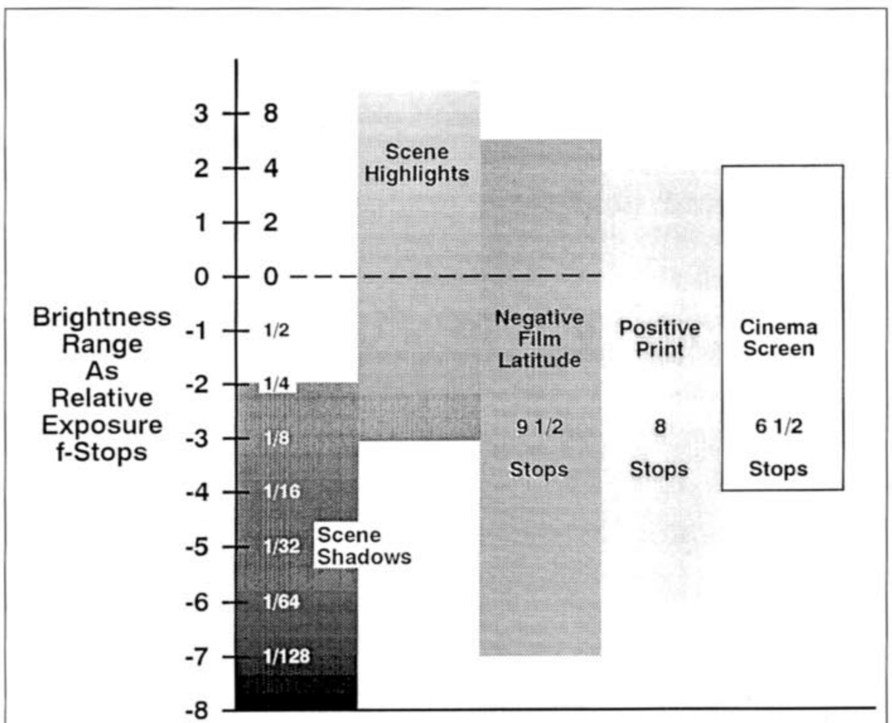

Figure 1. Typical capture of scene brightness range by negative film followed by progressive loss of contrast through subsequent system.

Presented at the 133rd SMPTE Technical Conference in Los Angeles (paper no. 133-100) on October 29, 1991. Laurence J. Thorpe is with Sony Advanced Systems, Montvale, NJ 07645.

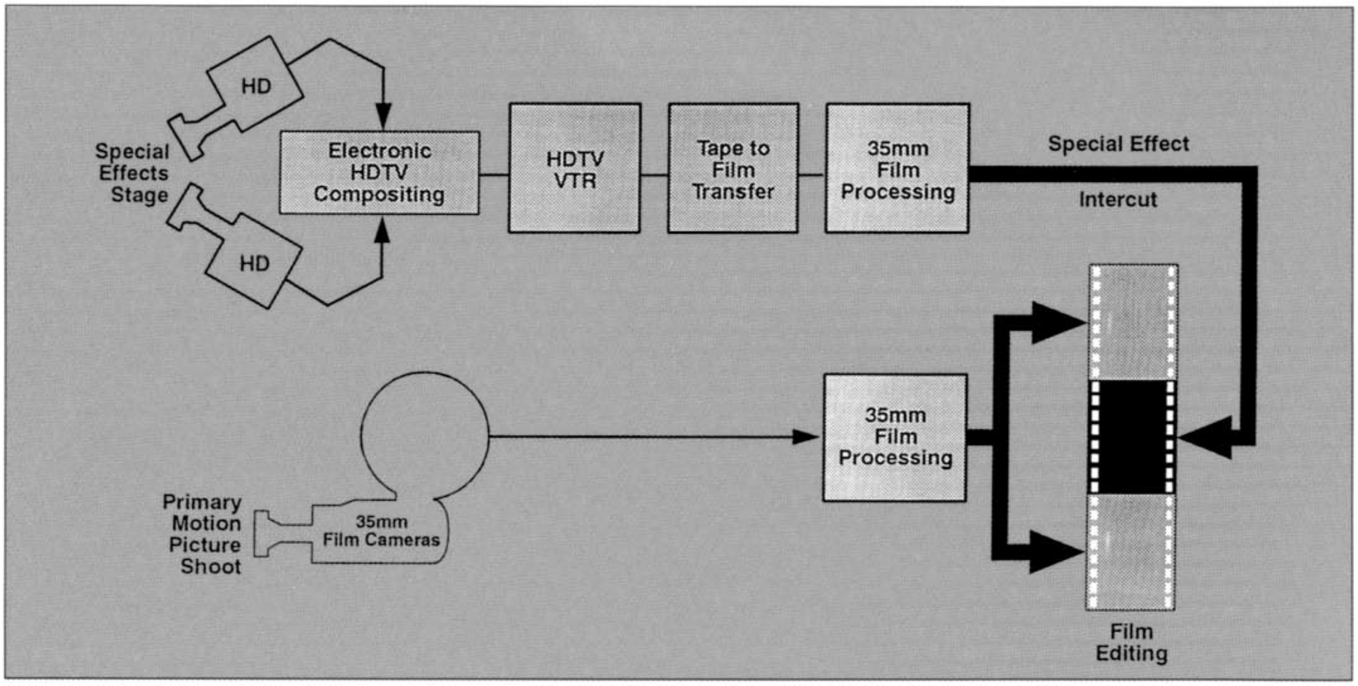

Figure 2. HDTV electronic blue-screen composite special effect is transferred to film and intercut with primary direct film origination.

the HD video signal will become a significant determinant in what can ultimately be achieved.

The HD Electronic Intermediate

Quite apart from the issues surrounding the final portrayal of HDTV (or film), there are the more important implications of when images from the two media might be brought together in the production/post-production process.[4] In this context we are much more concerned about "matching" the two images to achieve a seamless intercut, or possibly a blue-screen composite. This integration of images, originated from the two separate media, can take place in the film domain (following transfer of HDTV to film by electronic beam recorder or laser recorder), as in Fig. 2 or, it can occur in the electronic domain, following transfer by telecine of the original film imagery across to HDTV (Fig. 3). The latter electronic integration is perhaps the one likely to be more common in practice because of the advantages of all-electronic post production and image manipulation.

The concept of the electronic intermediate is a topic of considerable interest today. It can manifest itself in a "Super-HDTV" nonreal-time domain, as proposed by Eastman Kodak,[5] when a very high quality negative is the sought-for end result. It can also be implemented in a near real-time embodiment, using an HDTV production standard such as 1125/60 SMPTE 240M.[6] A great many of the considerations outlined in the following discussion concern the latter approach.

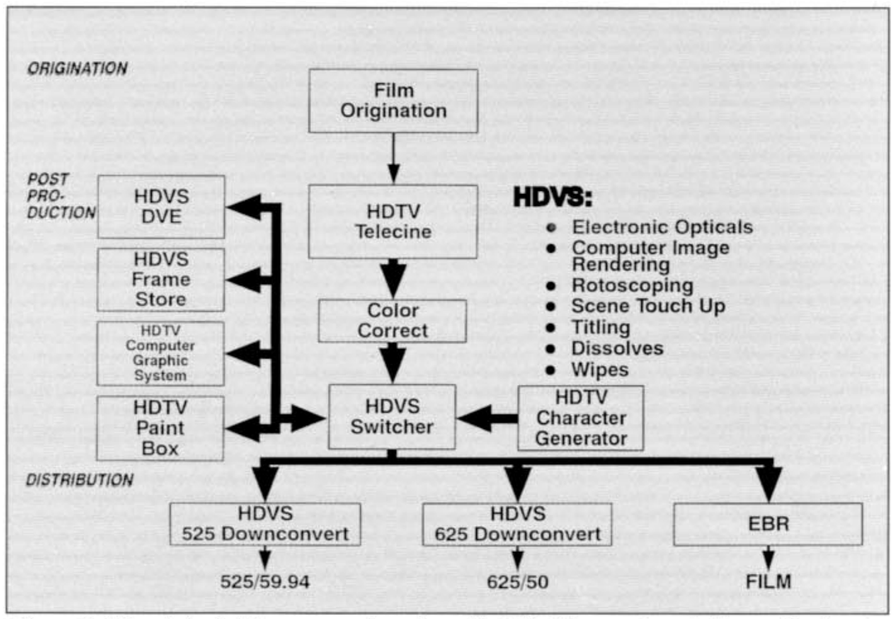

Figure 3. Film-originated images are transferred to HDTV for post-production, opticals, etc., and later transferred back to film or distributed directly as 525/625 television programming.

Combining HDTV and Film Images

It is often forgotten, in the heat of debate about horizontal pixel counts and scanning line numbers, that the HDTV image is, in fact, multidimensional; that is, the aggregate aesthetic quality of the picture is a complex combination of all of the following:

- Aspect ratio
- Horizontal resolution
- Vertical resolution
- Colorimetry
- Gray scale
- Total dynamic range (including overexposed portions of same) for still images, including other "dimensions" such as temporal resolution and lag for moving images.

The historic difficulty in finding

Table 1 — Different Terminology Used (in Video and Film Industries) for Comparable Imaging Parameters

Video	Film
Sensitivity	• Exposure index (EI) • Speed
Resolution	• Resolving power • Sharpness • Modulation transfer curves
Colorimetry	Color Reproduction
Gray scale	
Dynamic range	Exposure latitude
Noise	Diffuse rms granularity

Table 2 — The Need to Examine HDTV Video Dynamic Range

HDTV has broadened the scope of electronic imaging:

- Psychophysical relationship between wide display contrast ratio and sense of depth in the HDTV image
- Future HDTV display technologies may exhibit greater contrast handling range
- The desire to properly handle "overexposed" signals for creative image making
- Possibly we might require more RGB video "headroom" and "footroom" to realize wider color gamut
- The need to transform between media:
  - HDTV
  - Film
  - Computer graphics
  - Hard copy

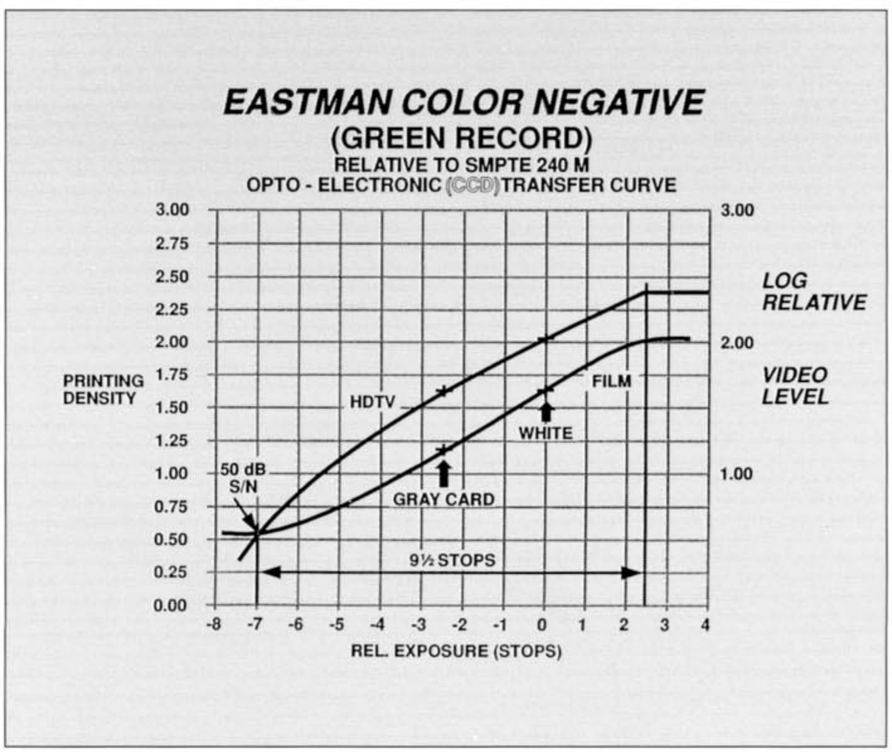

Figure 4. A comparison of the transfer characteristics of EXR-5245 color negative (green emulsion) and that of the HD CCD sensor output (modified by SMPTE 240M nonlinear pre-correction).

common ground for discussing HD video and film imaging lies as much in the disparate terminology used (Table 1) as it does in the task of properly quantifying some of these technical parameters. The particular mood or ambience that a program director might want to achieve in a given scene is often created in real time by artificial manipulation of one (or a number) of these dimensions, in both HDTV and film origination:

- Spatial resolution — by employment of fog filters on the lens or electronic image enhancement
- Colorimetry — by use of color filters or electronic painting
- Gray scale — by manipulation of an electronic camera transfer characteristic or a film process

When images, originated either on film or by an HDTV camera, are to be brought together in the electronic domain for possible integration (a composite or an intercut), it is desirable that they match each other as closely as possible in all of the imaging dimensions discussed earlier. The seamless operation of an electronic intermediate system is dependent upon the degree of this matching. A disparity in any one of the dimensions can, depending upon the scene content, easily mar the realism of the sought-for image composite.

A good match between the gray-scale characteristics of separate images originated on film and HDTV is dependent upon the specific transfer characteristics of each, and upon the exercise of certain operational discretionary practices during their separate shooting.[7] Some fundamental disparities have traditionally existed between film and video that mitigated against such an ultimate matching of gray scale. However, substantial improvements in video-camera sensors,[8,9] a better understanding of the respective media's transfer characteristics, innovations in manipulation of the video camera transfer characteristics,[10] and the recent spurring of interest in HDTV possibilities within the program creative community, have all added impetus to optimize the degree of match between the overall operational transfer characteristics of the two media.

HDTV Camera Dynamic Range

With the advent of the HD CCD sensor,[11] a giant step forward in electronic imaging, particularly from the viewpoint of total operating dynamic range, has been achieved. Figure 4 shows the transfer characteristic of the prototype HD CCD sensor, a 2-Mpixel frame-interline transfer (FIT) array, developed by Sony. Utilizing Hyper-HAD technology,[12] this imager has achieved a remarkable degree of control over dark noise and yields an effective total dynamic range of some 66 dB (2000:1). This range compares most favorably with that of the best 35mm negative film today (Fig. 4.) In this diagram, a direct comparison with the log scale (logarithm of printing density) of the film-transfer curve is structured by plotting the logarithm of the video-voltage level (which, in turn, is the signal-voltage output of the sensor multiplied by the precorrecting optoelectronic-transfer characteristic specified in the SMPTE 240M standard for 1125/60 HDTV).

Thus, the very good news is that already, in these early days of HDTV origination, camera technology has advanced quite dramatically in its ability to handle a large dynamic range. This excellent imaging capability demands that a careful examination of the subsequent entire HDTV system is warranted, on the basis that dynamic range has broad system implications. Some of these implications are summarized in Table 2.

The more sobering news is that, in fact, a great deal remains to be improved upon in the subsequent video system to fully capitalize on this new prowess of the HD imager. This squarely confronts us with the issue of digitization and video dynamic range — namely, the digital dilemma.

The Digital Dilemma

The dilemma of digital implementation of video studio signals is not unique to HDTV. We have been wrestling with this for almost a decade — in fact, ever since the marvelous achievement of CCIR Rec. 601 provided a solid basis for the ensuing global developments in 525/625 digital studio signal processing and recording.

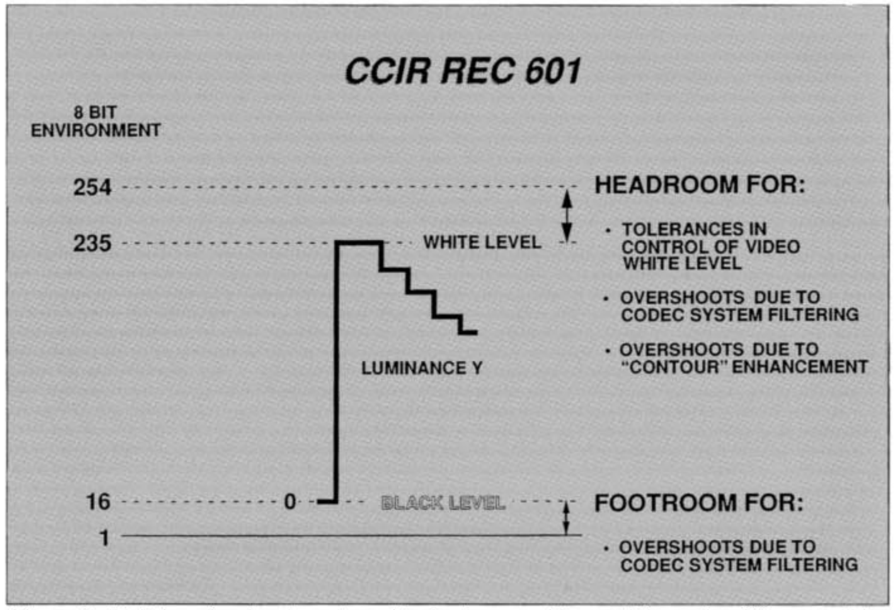

Figure 5. Digital coding levels assigned to luminance video and two color-difference video signals according to CCIR Rec. 601 (now defined for 525-line television within SMPTE RP 125).

Digital video is very demanding. It consumes digital bytes at a voracious rate. To record the three video components (luminance Y and two color-difference signals, Cr and Cb) specified in Rec. 601 requires a real-time total digital data rate of 216 Mbits/sec. A significant decision made in the early 1980s underlying Rec. 601 (following extensive examination and testing) was the choice of 8-bit linear quantization for each of the three component video signals. This was a good choice at that time, based on a pragmatic recognition of the challenges involved in actually implementing the high data rates of real-time digital VTRs and digital video effects, etc., and it remains today a sensible compromise for most television programming.[13]

However, the rapid advance of sophisticated high-speed computer graphic technologies, coupled with extraordinary breakthroughs in television-camera imaging, heralded in the mid-1980s by the arrival of the CCD imager, allows video image creation (in terms of both dynamic range and signal-to-noise performance) that outpaces the limitations of 8-bit linear quantization. A contemporary 525-line NTSC CCD camera has an unweighted luminance signal-to-noise ratio (SNR) in excess of 62 dB and utilizes an imager having a dynamic-range capability in excess of 78 dB (about 8000:1). This level of performance is readily available today in portable electronic newsgathering (ENG) cameras in addition to high-end studio cameras.

The situation is compounded by the further decisions within CCIR Rec. 601 relating to coding-level assignments. The exigencies of the real world of studio systems preclude the full exploitation of all 8 bits being applied to a video excursion from nominal black to peak white. The upper bound,imposed by level 255 in the 8-bit system, in fact functions as a highly effective (and quite unforgiving) "hard clipper."

Recognizing this, we must be cognizant of other system realities within a total television studio complex, including the following:

- Difficulties in maintaining precision in video levels throughout extensive routing and distribution systems
- Unpredictable excursions of video level due to low-frequency transient anomalies
- Vagaries of amplitude and group delay characteristics of multiple-system elements
- Operational controls on video levels (with operators in attendance)
- Video overshoots inherent in the coder/decoder filtering associated with signal conversions between analog and digital equipment
- Video overshoots associated with image-enhancement systems

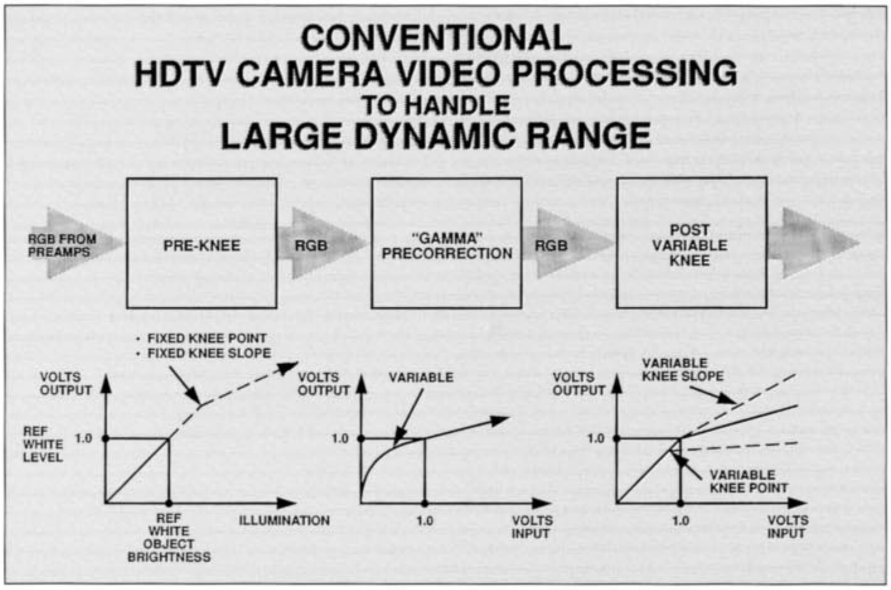

Figure 6. Three primary stages of nonlinear video compression typically employed in current television cameras to facilitate capture of wide dynamic-range scenes.

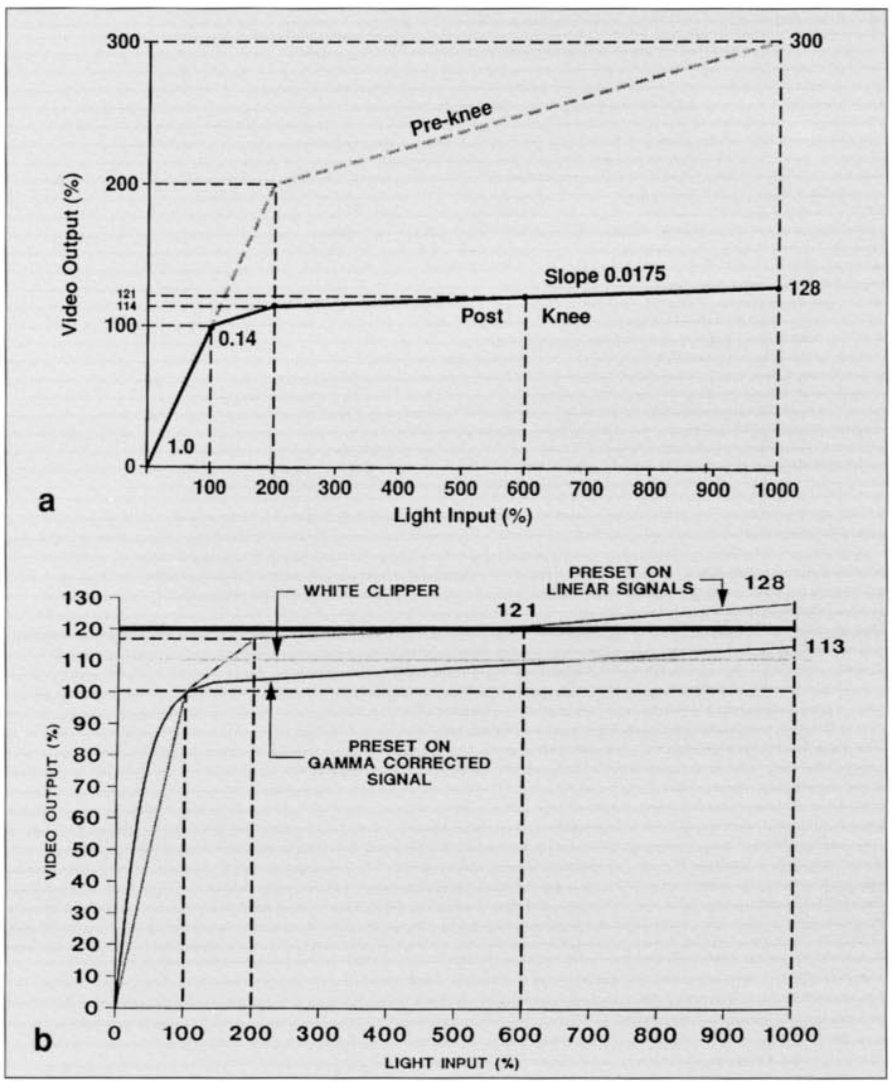

Figure 7. Nonlinear precorrection in Sony HDC-300 camera: (a) fixed pre-knee compressor following preamplifier and nominal setting of post-knee compressor following SMPTE 240M gamma correction; (b) overall extended transfer characteristic for linear signals and SMPTE 240M precorrected signals.

- Subjectively disturbing visual artifacts associated with hard clipping of video signals and/or their associated overshooting
- Subjectively unattractive appearance of clipped noise in the vicinity of video black level (black axis shift as a result of rectification)

These considerations led to the establishment of prescribed coding levels for the luminance and color-difference signals (Fig. 5), which allowed a small linear range beyond normal black-and-white video levels. A considerable accumulated global experience testifies to the wisdom of incorporating the "headroom" and "footroom" shown in this figure. The dilemma introduced, however, revolves around the consequent erosion of available quantization levels applied to actual video images — some 220 rather than 256 — in order to accommodate these "safe" areas.

The real dilemma in current 525/625 television, however, lies less with this modest restriction in quantization levels than with the fact that proper exploitation of the available video signal levels is difficult to ensure in practice. This can only be guaranteed when analog video signals applied to an analog-to-digital (A/D) converter are disciplined to align nominal peak white to level 235 and nominal black to level 16 (for luminance). In practice, such precision is rare, and nowhere is the imprecision more precarious than in the camera-to-recorder interface. Today, cameras are still largely analog in their video processing, but all recorders (even analog) utilize digital techniques within to implement time base correction and time compression multiplexing.

Unfortunately, both camera and VTR technology incurred dramatic technological advances throughout the decade of the 1980s, and the coordination between the two fell victim to the separate driving imperatives of each. Fierce competition between camera manufacturers, and also between VTR manufacturers, was fueled by the explosive growth in worldwide ENG and electronic field production (EFP).

Even before the advent of the CCD, broadcast camera design had made remarkable strides in lowering SNR (60 dB being contemporary for 2/3-in. pickup-tube portable ENG

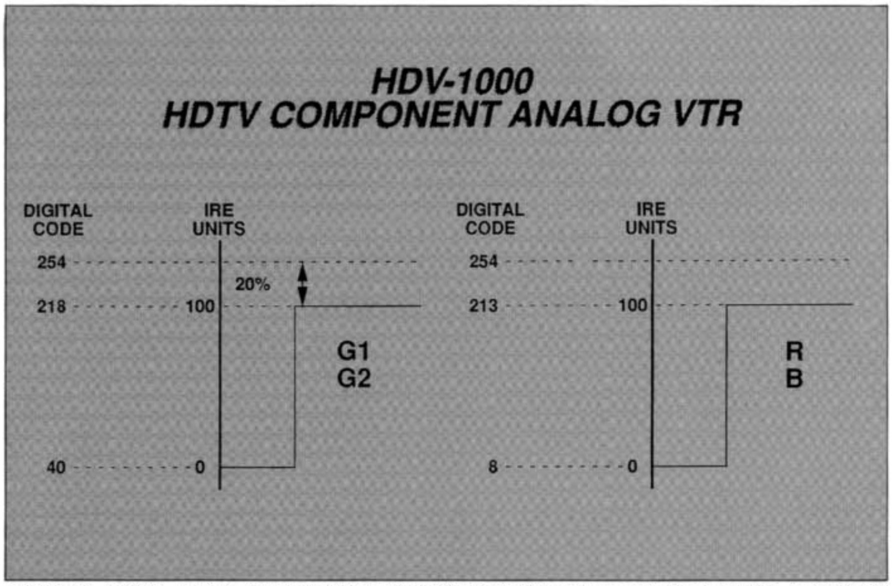

Figure 8. Digital coding-level assignment in first HDTV component analog VTR HDV-1000 (1984). Note generous headroom.

cameras in 1986) and extending dynamic range to better cope with the always present, uncontrollable highlights encountered in news gathering on location. Dynamic ranges of the order of 56 dB (about 600:1) were typical. Innovative schemes to exploit such dynamic range imaging capabilities proliferated, and virtually all cameras were soon endowed with novel combinations of video "knee" controls to allow subjectively acceptable compression of signals beyond nominal white exposure (Fig. 6). Controls such as "knee slope," "knee point," "knee gain," and "highlight contours" abounded, with each camera manufacturer separately seeking incremental competitive innovations in the handling of wide scene contrast ranges. The automatic beam optimization (ABO) schemes to dynamically control pickup-tube beam reserve also progressively improved.

The final incorporation of automatic knee controls (auto contrast) that would attempt crude analysis of actual scene content to facilitate closed-loop dynamic control of the degree of video compression in the highlight region were being widely implemented at about the time the CCD imager burst on the scene. The 66 to 72-dB (and today, even more) dynamic range of these sensors, coupled with their breakthrough elimination of the vexing highlight artifacts (i.e., comet tailing, blooming, sticking) of the pickup tube sparked a redoubling of efforts to exploit the enhanced dynamic range capabilities of the camera.

Attendant with such technical activities there arose an understandable thirst for creativity in image-making among the large numbers of videographers now out on location on a daily basis. The ever-present desire to emulate the "film look," particularly in outdoor shooting, stimulated widespread experimentation with camera processing controls. An unfortunate consequence of this was the new laxity that entered regarding the setting of the output white clipper control in video cameras. Formerly disciplined by technical personnel to be carefully set at a level commensurate with television transmitter capabilities (102 to105% above nominal white), now the setting of this control became an intimate part of the overall camera "highlight compression" alignment. It was not (and still is not, today) unusual to see white-clipper settings hovering in the 120 to130% level referenced to nominal white exposure, with attendant knee-compression curves set to handle 600:1 dynamic range in video levels.

Indeed, this current practice in 525/625 television broadcast cameras directly influenced the design philosophy of virtually all emerging HDTV cameras. As one example, Fig. 7 depicts the nominal design parameters of the Sony HDC-300 HDVS camera based upon its attempt to exploit most of the 60-dB dynamic range capability of the new all-electronic pickup tubes employed in that camera. The white clipper, when set to 110% of nominal white, and with all knee-related controls in their nominal position, allows about 800% (or 58 dB) of video dynamic range to pass under the clipper to the camera output. The analog HDV-1000 HDVS VTR in use at the time of the HDC-300 design embodied its own independent choice of digital coding levels for the built-in time base compressor (TBC) and signal

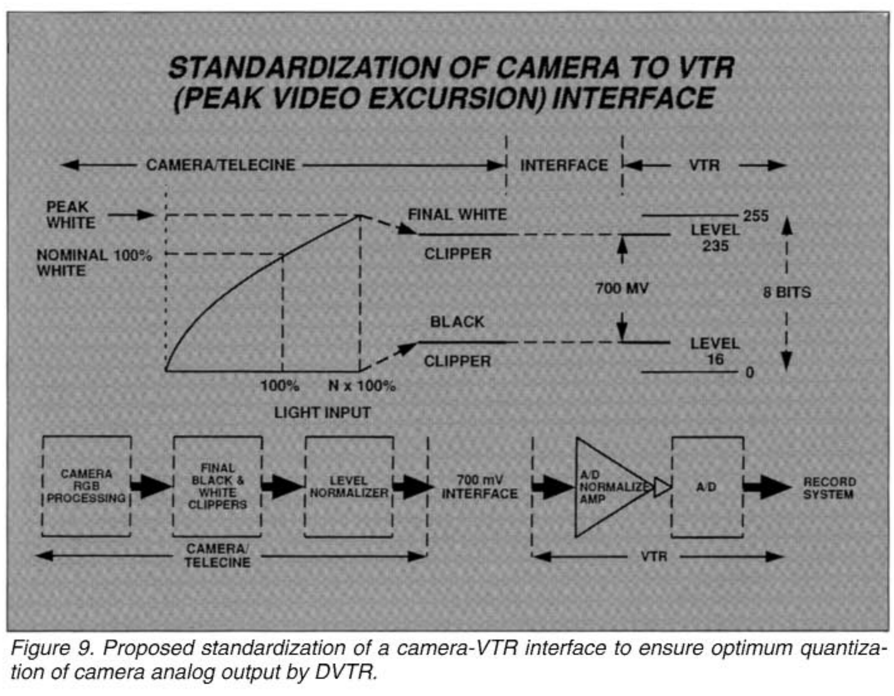

Figure 9. Proposed standardization of a camera-VTR interface to ensure optimum quantization of camera analog output by DVTR.

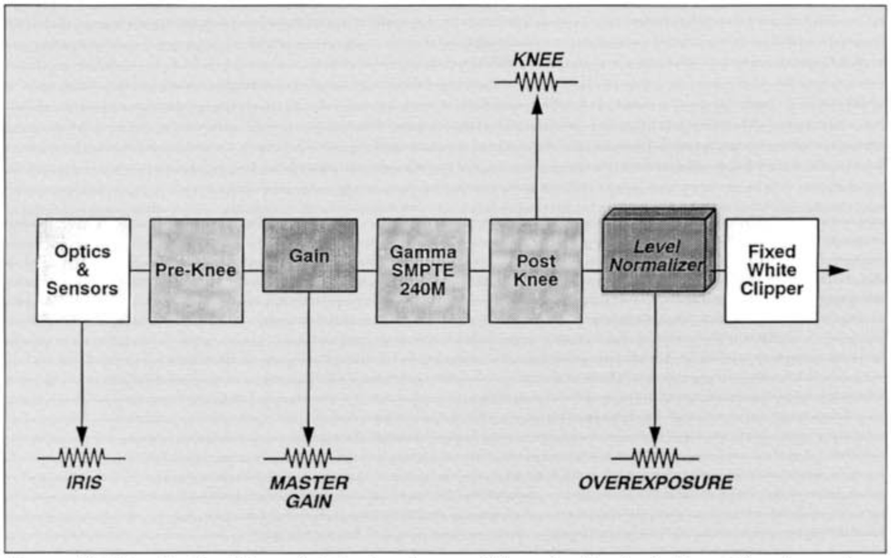

Figure 10. Required addition of video level normalizing circuit prior to the white clipper (now fixed setting) to facilitate operational control for handling extended dynamic range.

processor (Fig. 8). These coding levels were essentially a pragmatic judgment on the part of VTR engineers, who fully expected HDVS camera clippers to be set to nominal 110% (approximately) but also felt the need for additional protective headroom to handle further operational adjustments.

The experiences gained in early international HD program production brought home to both the Sony camera and VTR engineers the undisciplined vagaries of shooting in the real world. It also clearly exposed the urgent need for some form of standardization to reconcile the innate and strong desire among HD program producers for wide dynamic range image capture with a sensible digital level coding structure and camera-recorder interface. Unable to rationalize this internally in the absence of such a standard, Sony (in 1987) appealed to both the Broadcasting Television Association (BTA) in Japan and the SMPTE Working Group on High-Definition Electronic Production (WG-HDEP) to address this issue on an urgent basis.

The Rationalization of the Camera VTR Interface

In a 1989 paper submission to the SMPTE WG-HDEP, Sony proposed for committee consideration a strawman HDTV camera-recorder interface. It argued that there is only one sensible solution to the analog camera output interface to the digital VTR input — namely, control of the camera output signal level, definition of the camera white clipper setting, and reconciliation of the two with the VTR A/D converter in the manner shown in Fig. 9.

If the camera white clip level is set to a fixed nominal level — say 700 mV, as measured at the camera output — then the VTR can regard 0 to 700 mV (black clipper set to 0 V) as the standardized interface level and base its A/D settings on such a signal being presented at the VTR input. These settings can now align the quantization level 235 with the camera white clipper and quantization level 16 with the black clipper and ensure that no camera source signal will intrude into the "headroom" area reserved for subsequent system uncertainties (level variations, ringing, etc.) This proposal is now part of the proposed standard for the digital representation of SMPTE 240M, i.e., SMPTE 260M.

A new control mechanism must be added to the camera operational control panel — an output level normalizer, as shown in Fig. 10. Presently, if an HDTV camera is aligned in the normal manner on a gray-scale chart, this operational control would be set to an indent position corresponding to 0-dB master gain and 0-dB normalizer gain. If, however, in the course of actual shooting, a scene is creatively imaged with a specific intent to capture compressed overexposed signals, then the output level normalizer control must be readjusted to attenuate the output signal so that it passes "under" the now-fixed white clipper. Operationally, all the video operator need do is set this new output gain control to align the chosen overexposed peak white level to the clipper.

In actual practice such a normalizing control can be manually operated as described, or automated in some manner (automatic scene-contrast controls in some present-day cameras in one sense behave like this control).

The Digital Dilemma in Relation to Post-Linearization of the Camera Signal

When the SMPTE WG-HDEP mathematically defined an HDTV camera nonlinear precorrection

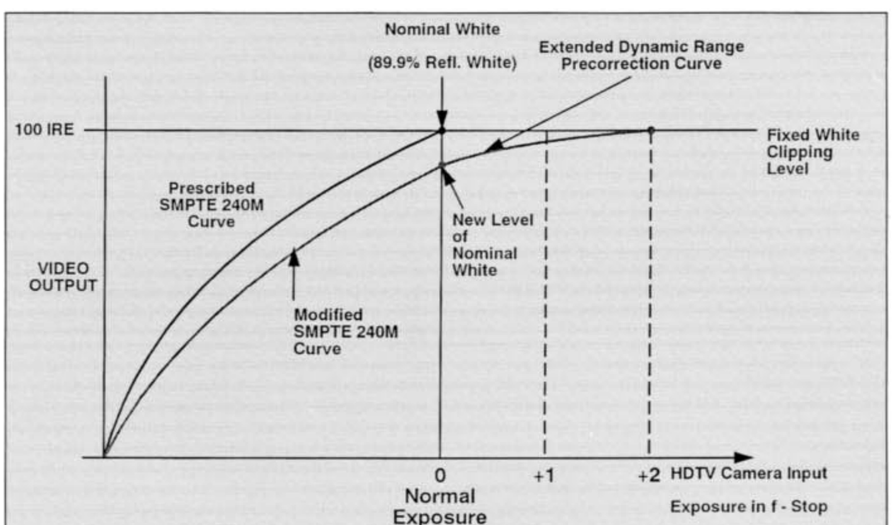

Figure 11. Modification introduced by an extended dynamic-range HDTV camera precorrection curve on original prescribed SMPTE 240M curve.

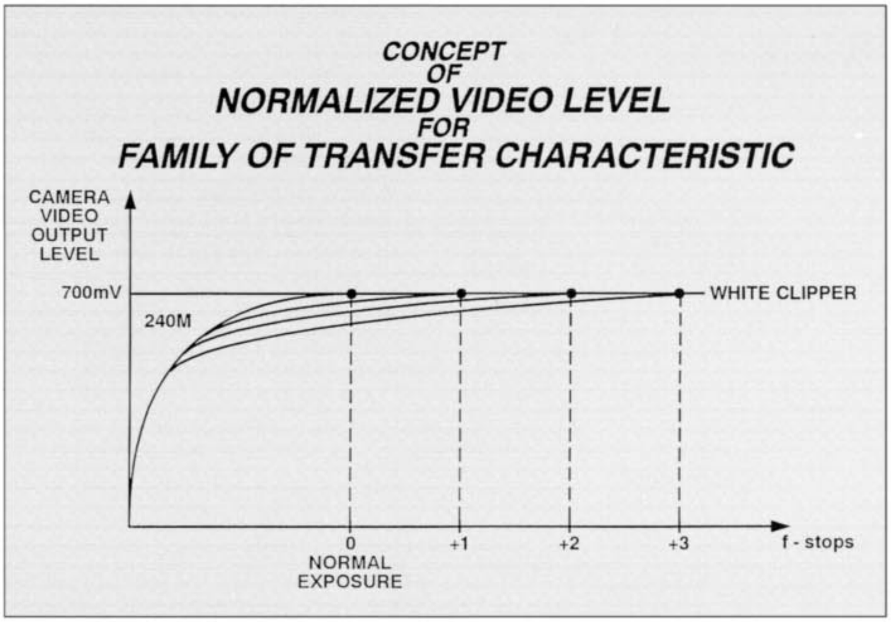

Figure 12. Concept of a family of prescribed curves allowing predictable operational control of camera extended dynamic-range shooting.

curve, it was with the intent to accomplish three things:

• Encourage a convergence in HDTV camera design to ensure better picture matching between cameras from different manufacturers.

• Allow a more predictable linearizing of the camera signal at any subsequent point in the system.

• Decouple the HDTV camera origination entirely from the final viewing image (from the viewpoint of colorimetry and transfer characteristic), allowing display gamma correction to be implemented in the viewing system itself.

This linearization might be required in a display system that incorporates a linear matrix for color transformation appropriate to that particular technology; or it might be to facilitate appropriate color linear transformation in an HDTV tape-to-film recorder (the particular transformation matrix being determined by the film stock employed; or it could be to allow complex digital manipulation of the image in post-production (such manipulation is generally conceded to best be accomplished on linear signals). The primary role of the camera nonlinear precorrection should become that of a compression circuit to optimize the signal-to-noise performance as the HD signal passes through the entire system. However, the latter facility introduces a new anomaly into the digital dilemma: the SMPTE 240M standardized transfer characteristic is specified only for video levels ranging from black to nominal reference white (that is, the 89.9% reflectance white chip on a reference gray scale). It does not deal with overexposed signals.

As described earlier, each camera manufacturer will implement the SMPTE 240M curve as prescribed, but for signals captured above nominal white, the form and shape of the compression knee curves are left to the discretion of each manufacturer. Assuming the new camera level normalizer control is used to ensure that a given scene (with compressed overexposed signals) is faithfully captured and properly applied to the A/D of the VTR, then the camera output transfer characteristic is no longer that of SMPTE 240M (Fig. 11). We are now dealing with an extended transfer characteristic and a special difficulty is presented to any downstream "linearizer" that might seek to restore the original camera imaging linear signals to allow digital manipulation. The question is: What shape linearizing curve should be employed?

This problem has been recognized within the SMPTE WG-HDEP, and its Ad Hoc Group on Colorimetry has undertaken the task of developing some rational approach to standardizing nominal extended transfer characteristics for handling extended dynamic range operation by an HDTV camera. Among the issues being examined are:

• The possibility of defining a family of nominally prescribed curves to handle progressively increasing dynamic range.

• A means of introducing more user-friendly camera controls to handle overexposed signal handling

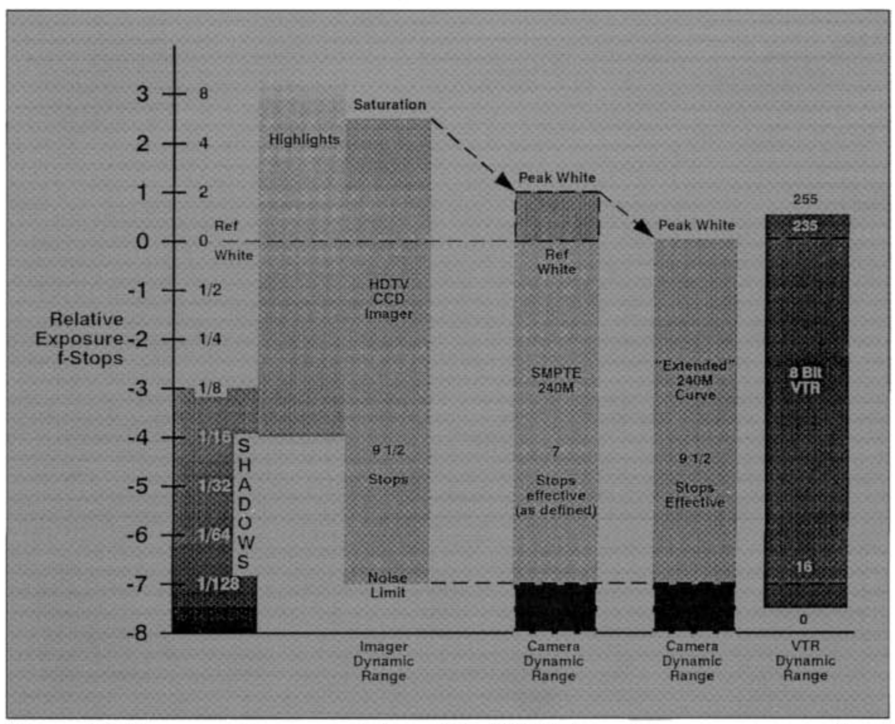

Figure 13. Diagrammatic representation of the HDTV dynamic range system problem, notably the restriction of 8-bit digital sampling.

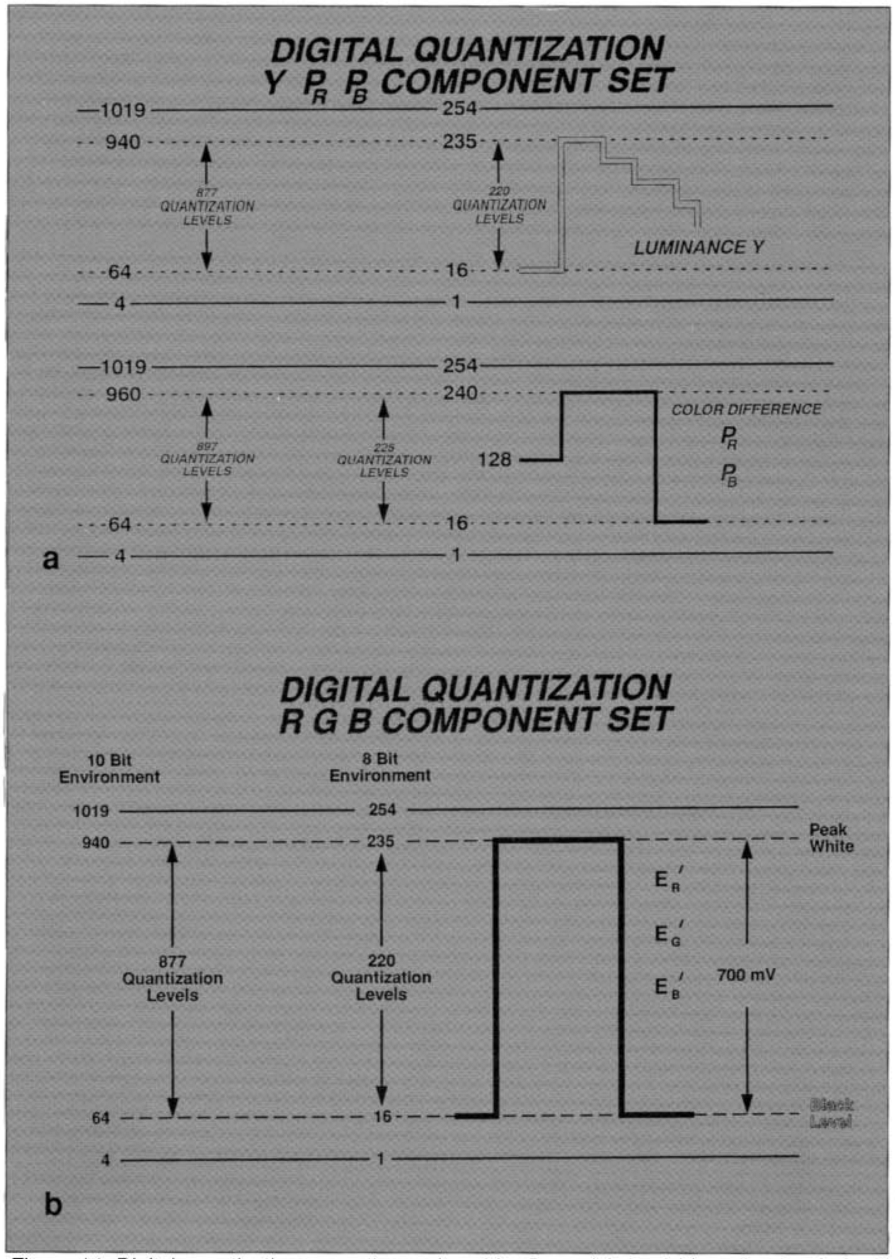

Figure 14. Digital quantization currently employed by Sony: (a) the 8-bit coding levels for Y Pb Pr used in the HDD-1000 digital HD VTR and the 10-bit coding levels used in other equipment under development; (b) the 8-bit and 10-bit coding levels applied to the RGB component set.

from a video operational viewpoint; that is, how to make implementing the family of extended curves as easy as current film cinematographers' modus operandi when they creatively exploit the various portions of a negative film's nonlinear transfer characteristic to capture a broad range of quite varied scenes and environmental moods (Fig. 12).

• Means of carrying information relating to such HDTV camera creative operational adjustments through the system (such as a video index signal carried with the video information itself) so as to allow appropriate (perhaps even automatic) selection of a complementary linearizing circuit anywhere downstream in the HDTV post-production system.

Return to the HDTV Digital VTR

Of course, overshadowing all of these considerations is the potentially restrictive quantizing capability of an 8-bit sampling system. Coding dynamic range is defined as the ratio between the peak power of a sinusoidal signal being quantized and the power of the quantizing noise.[14] On this basis the maximum dynamic range for an 8-bit system is 49.6 dB. But what does this mean when we are dealing with a nonlinear video HDTV waveform? The following sections will attempt to shed a little light on this. But first let us take a macro-view of the system.

If a diagram similar to that shown for the total film system is now drawn for the HDTV system, it is seen that the marvelous dynamic range capture by the sensors might be quickly eroded when the first 8-bit digital VTR is encountered (Fig. 13). Clearly, a similar dilemma is encountered when transferring from motion-picture film to HDTV in a digital telecine.[15]

In 1987, Sony was in the final stages of development of the first all-digital HDTV VTR capable of recording an HDTV video component set conforming to the new SMPTE 240M production standard. While some key digital encoding parameters had emerged within ATSC and SMPTE discussions, no standard had yet been developed. However, based upon industry discussions to date, Sony elected to proceed on the basis of these tentative parameters[16] and to employ an 8-bit coding level structure conforming to CCIR Rec. 601. A 10-bit coding-level assignment had also been chosen, based upon a direct scaling from the 8-bit set for work being done on a digital version of the Sony electron beam recorder (EBR) and other HDVS film-related systems currently under development. These coding structures are shown in Fig. 14.

The issue of coding levels applied to HDTV signals has been studied intensively over the past two years within the SMPTE Ad Hoc Group on Digital Representation of SMPTE 240M.[17]

Clearly, the 10-bit HDTV digital system option under consideration by SMPTE represents a highly desirable goal in the convergence of HDTV and film imaging. The effective quantization levels of such a system are 877, yielding a dynamic range upper limit of about 59 dB (based upon our earlier definition). However, the challenge posed by a 10-bit system to HD VTR technology remains. The current studies of such videotape recording also encompasses the even larger topic of bit-rate reduction of HDTV production signals, although the precise capability of the 8-bit digital HD VTR needs to be clarified.

Dynamic Range of the HDTV Camera and Digital VTR

The central question becomes: What dynamic range (captured by the HDTV camera) can, in practice, be faithfully reproduced on playback by a digital 8-bit VTR? A simple analysis will reveal most of all that is involved.

Case 1. Normal Exposure

Consider an HDTV CCD camera imaging a 16-step gray-scale chart employing a $\sqrt{2}$ relationship between adjacent steps (thus producing a 2:1 brightness change, or 1 *f*-stop, for every 2 steps).[1] The linear voltage output of, say, the green CCD would then be as shown in Fig. 15. On a log scale the steps are equal amplitude and are shown here normalized to 1 V for easy reference. The 52-dB SNR of this CCD imager produces a root mean square (rms) voltage of 2.5 mV of noise relative to 1 V pp of signal. This rms value is essentially equal to the peak-to-peak amplitude of the 15th riser in this gray scale, that is, at step 15 the SNR has decreased to 0 dB, rendering this riser indistinguishable, and thus establishing a boundary to the dynamic range of the imaging system. As was shown in Ref. 1, the dynamic range of the HDTV cameras is determined by the SNR of the sensor (and has nothing to do with the camera nonlinear correction). Clearly, in this case, for a normally exposed signal the dynamic range is 7 *f*-stops.

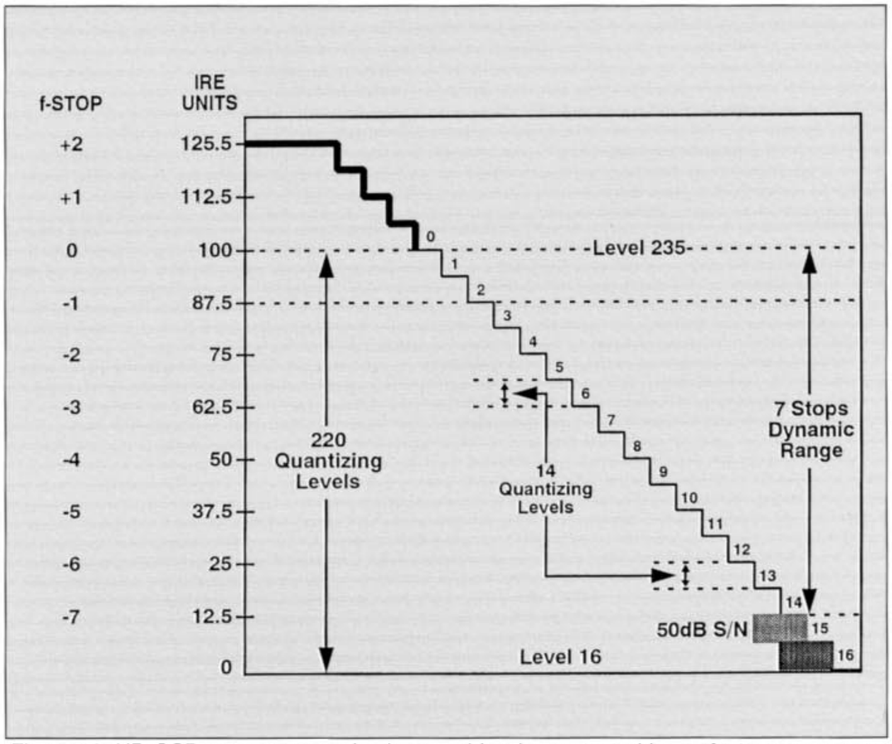

Figure 16. HD CCD camera output luminance video (precorrected by perfect gamma power curve) when viewing gray-scale chart with 2 f-stops overexposure.

If this signal were now passed through a camera "gamma" corrector having a perfect complementary power law (the video signal voltage at the output would be a linear stairstep having risers of equal signal amplitude) with the lower step 16 within a few tens of mVs of nominal black level. Following the nonlinear precorrection circuit it will now be assumed that the three RGB signals are matrixed (according to the equation in the SMPTE 240M standard) to produce a luminance Y signal. If the camera master black level control were very slightly lowered to place the 16th riser of the luminance signal at the black clipping lower boundary, the video signal levels (in IRE) would be, to a good approximation, as reproduced in Fig. 16. The effective dynamic range of the camera has now been precisely "fitted" to the two camera video processing boundaries defined by the two fixed black and white clippers.

Applying this Y signal output to an HDTV digital VTR according to the new interface rules of Fig. 9 means that all 220 quantizing levels are available to handle the 16-step gray scale, with almost 14 quantiz-

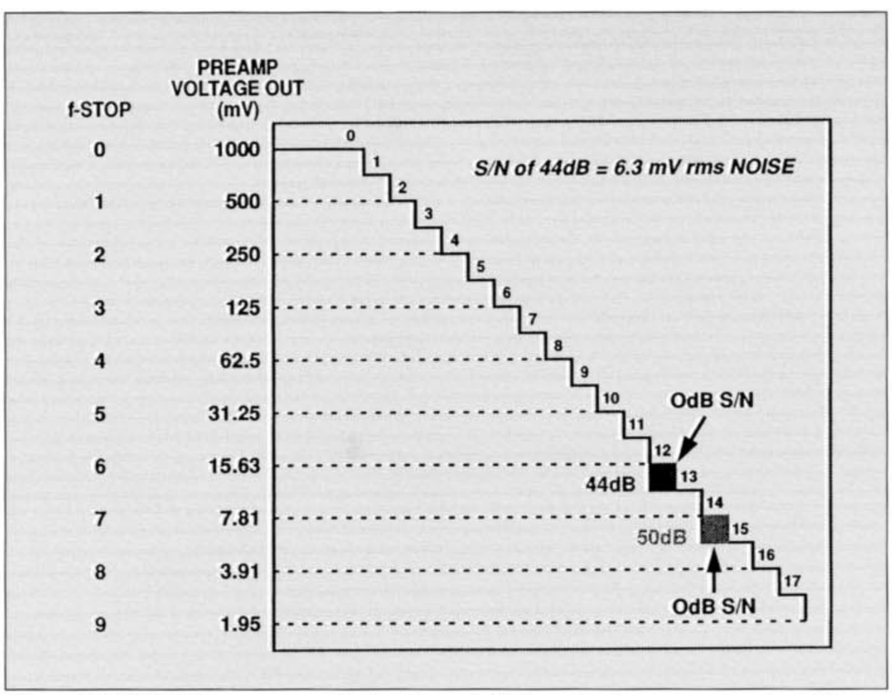

Figure 15. HD CCD imager normalized video output when viewing gray-scale chart with normal exposure.

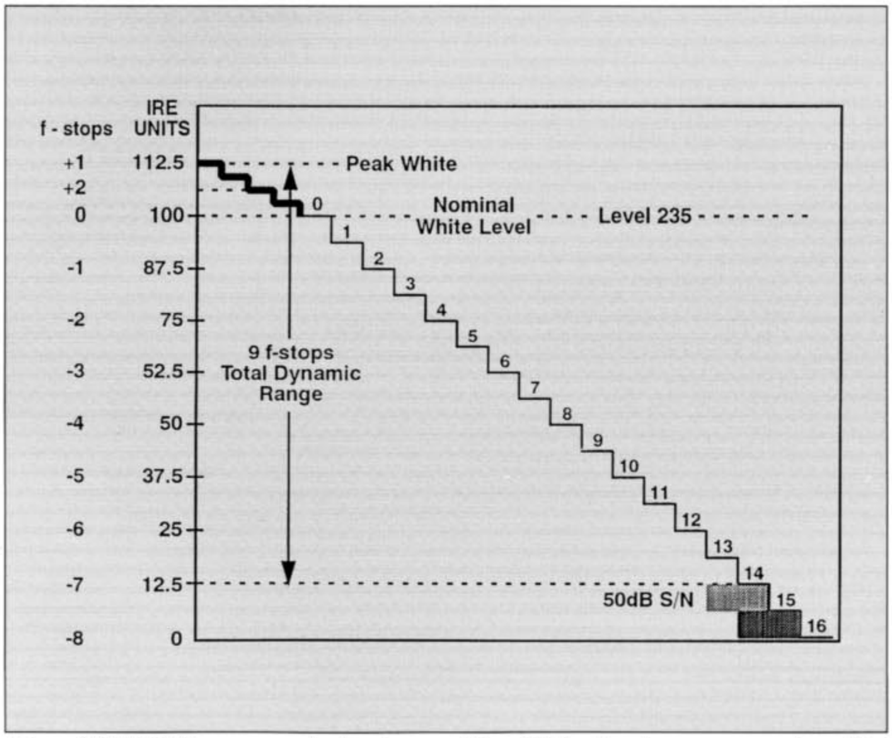

Figure 17. HD CCD camera output luminance video with fixed pre-knee compression added before gamma correction.

ing levels available for each of the 16 risers. Thus, the 8-bit VTR, with coding levels according to Fig. 14, is more than adequate to handle the 7 *f*-stop dynamic range capability of a normally exposed HDTV CCD camera.

Case 2. Overexposed Camera

If a gray-scale chart could now be postulated having the same 16 brightness levels discussed, but also having four additional steps, producing brightness levels brighter than the nominal white chip (level 0) according to the same $\sqrt{2}$ law, then the camera is imaging 2 *f*-stops of brightness "overexposure." Assuming the camera imager reproduces this in a linear manner (which the HD CCD can do), and that this signal is again applied to a perfect power law gamma corrector circuit having extended video handling capability above nominal white, then the output of this circuit would be the linear staircase reproduced in Fig. 17. The peak video level would now be 125.5% of nominal white level. Making a further assumption that a piece-wise fixed linear pre-knee circuit is introduced between the CCD outputs and the gamma-correction circuit in a manner designed to compress the two *f*-stops of additional video signal level to 112.5% (instead of 125.5%) at the gamma-corrector output, then a luminance voltage waveform would be shown on an IRE scale, as in Fig. 18.

Activating the new level normalizer control (following the gamma-correction circuit) to set the +2 *f*-stop luminance peak white signal level at precisely 100 IRE units would linearly readjust all stairstep levels to that shown in Fig. 18. Each of the linear risers below nominal white level (now reset to 88.8 IRE) has an amplitude of 5.5 IRE units. With 220 quantizing levels now applied to this total waveform, as shown in Fig. 18, there are now 12.2 quantizing levels for each riser in the linear portion of the stairstep video waveform and 6.1 quantizing levels for each of the compressed risers corresponding to the overexposed signals. Clearly, again, the 8-bit digital videotape recorder (DVTR) can unambiguously reproduce the total 9 1/2 *f*-stops. It would only take a small amount of work to see that a further increase of another *f*-stop of overexposure would also be readily accommodated by 220 linear quantizing levels. Thus, the full 10 *f*-stop capability of an HD CCD imager can be properly captured on an 8-bit HD VTR, provided an appropriately designed nonlinear compression circuit operating over the total dynamic range capability of the imager is employed.

The above analysis was based on two key premises:

- Perfect power law gamma pre-correction circuit of wide video dynamic range

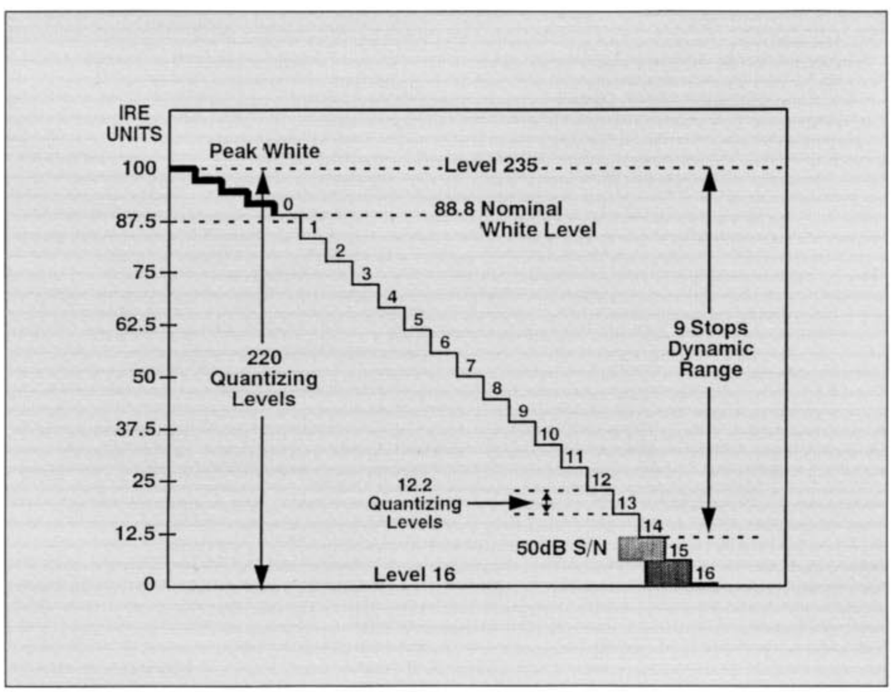

Figure 18. HD CCD camera luminance output with video level adjusted to meet proposed camera output standard.

• Careful design of a pre-knee compression circuit to ensure adequate signal amplitude in the overexposed signals

The SMPTE 240M precorrection curve, however, is not a perfect power law.[1] It was instead carefully designed to deviate from the high gain of such a law in the lower signal regions. If the SMPTE 240M curve is used, the output stairstep voltage, instead of having linear equal amplitude steps (Fig. 16), will exhibit an increasingly compressed series of steps at the lower regions and somewhat uncompressed steps at the upper regions. This will, in turn, progressively reduce the quantizing levels (for an 8-bit A/D converter) available to each of the lower risers and increase them slightly for the top risers.

The amplitude of each of the risers can be easily calculated for the case of precorrections according to the SMPTE 240M curve. It would show that a compression of about 4:1 would affect the riser for level 16 and a compression of slightly in excess of 3:1 for the riser of level 12. Thus, instead of 14 quantizing levels being available for the lower steps, a worst case of about 3.5 quantizing levels is available. However, this is quite enough for reliable representation of this step. Thus, in total, the full 7 *f*-stops of dynamic range is well satisfied by the 220 quantizing levels.

Clearly, it is in dealing with the extended dynamic range signal that the limits of 8-bit recording are approached. Referring back to Fig. 18, it is obvious that the SMPTE curve will further compress the lower steps. Instead of 12.2 quantizing levels for the lowest riser (that for level 16) there are now only 3.0, and about 4 for the riser in level 12.

Conclusion

Capitalizing on the already splendid dynamic range capability of the HD CCD imager is very important to a system that intends to use HDTV (according to the SMPTE 240M standard) as an electronic intermediate to bring together images from film, computer-generated imaging systems, and live HDTV imaging. Dynamic range capabilities encompass a great deal of the creative flexibility open to cinematographers, regardless of the medium in which they might choose to shoot.

The current state of the art in HDTV digital VTRs (and indeed other real-time HD digital processing systems) hovers slightly in excess of 1 Gbit/sec. With the spatial resolution of the SMPTE 240M standard, this relegates sampling resolution to 8 bits. Ten-bit operation has been identified as a resolution more suited to fully exploiting all the dimensions of the HDTV image, particularly as camera imagers continue to improve. However, until HDTV recording technology allows reliable operation at 10 bits, we must make do with 8 bits.

Eight bits can, nevertheless, do an impressive job, provided appropriate precautions are taken, notably in the HDTV camera. While the SMPTE 240M camera nonlinear precorrection curve may not be ideal for optimizing the "fit" of a video signal of wide dynamic range to the digital coding levels specified for the VTR 8-bit A/D converter, it still strikes the best compromise between the required compression and defining the SNR performance of the HDTV production system. However, the SMPTE 240M precorrection curve applies only to a normally exposed camera; it does not deal with overexposed signals. The proper handling of an extended dynamic range must therefore go further than the very useful precedent that was set in the writing of the SMPTE 240M standard. The SMPTE WG-HDEP is currently studying all aspects of HDTV dynamic range. Central to the task at hand are considerations of:

• Design of an overall compression curve for handling overexposed HDTV camera signals.

• Means of predictably controlling this transfer characteristic, from an operational viewpoint.

• Reconciliation of HDTV camera output video levels with the A/D converter of the digital VTR (as proposed by the SMPTE Ad Hoc Group on Digital Representation of SMPTE 240M).

• Optimization of the shape of the overall transfer curve (for maximum camera output dynamic range) for 8-bit A/D conversion.

• Means of transmitting information within the HDTV signal itself that describes camera operational adjustments for exposure and extended dynamic range capture to allow complementary downstream linearization when required.

References

1. L. Thorpe, "HDTV and Film — Issues of Dynamic Range," *SMPTE J., 100:*780-795, Oct. 1991.
2. SMPTE RP 41-1983, "Evaluation of Color Films Intended for Television," *SMPTE J., 93:*543-544, May 1984; and ANSI/SMPTE 196M-1986, "Motion Picture Film Screen Luminance and Viewing Conditions — Indoor Theater Projection," *SMPTE J., 95:*1124-1125, Oct. 1986.
3. M. Ogomo, T. Yamada, K. Ando, and E. Yamazaki, "Considerations on Required Property for HDTV Displays," *Proc. of HDTV 90 Colloquium, Vols. 1, 2B:* 1.2 28.1.18.
4. H. Tanaka and L. J. Thorpe, "The Sony PCL HDVS Production Facility," *SMPTE J., 100:*404-415, June 1991.
5. B. Hunt, G. Kennel, L. DeMarsh, and S. Kristy, "High-Resolution Electronic Intermediate System for Motion-Picture Film," *SMPTE J., 100:*156-161, Mar. 1991.
6. J. Galt and R. West, "A Comparison of High-Definition Electronic Optical and Film Optical Processes," presented at 132nd SMPTE Technical Conference, New York, Oct. 1990.
7. H. Mathias, "Gamma and Dynamic Range Needs for an HDTV Electronic Cinematography System," *SMPTE J., 96:*840-845, Sept. 1987.
8. L. Thorpe, E. Tamura, and T. Iwasaki, "New Advances in CCD Imaging," *SMPTE J., 97:*378-387, May 1988.
9. L. Thorpe et al., "New High Resolution CCD Imager," *1988 NAB Engineering Conf. Proceedings,* pp. 334-345.
10. M. Favreau, S. Soca, J. Bajon, and M. Cattoen, "Adaptive Contrast Corrector Using Real-Time Histogram Modification," *SMPTE J., 93:*488-491, May 1984.
11. Y. Ide, M. Sasuga, N. Hamade, and T. Nishizawa, "A Three-CCD HDTV Color Camera," *SMPTE J., 99:*532-537, July 1990.
12. L. Thorpe and D. Dahlberg, "On-Chip Lens Hyper-HAD Sensors for Increased CCD Sensitivity," *Advanced Imaging,* pp. 43-45, June 1991.
13. D. P. Owen, "Dynamic Rounding in Digital Video Processing: An Update," *SMPTE J., 98:*447-450, June 1989.
14. M. Bellanger, "Digital Chapter 1, *Processing of Signals — Theory and Practice,*" pp. 26-29, John Wiley & Sons Ltd.: New York, 1984.
15. R. Lees et al.,"High-Performance CCD Telecine for HDTV," *SMPTE J., 99:*837-843, Oct. 1990.
16. L. Thorpe et al., "HDTV Digital VTR," *SMPTE J., 98:*738-747, Oct. 1988.
17. H. Gaggioni, "Work in Progress on the Development of the Digital Representation of the 1125/60 HDTV Studio Signal," presented at the 132nd SMPTE Technical Conference in New York, Oct. 1990.