SpectraCal

Poynton's Vector

Poynton's Vector comprises a collection of short notes, issued approximately monthly, on topics of interest to those who seek faithful presentation of video, preserving the content creators' intent (and especially home theatre enthusiasts and calibrators). The series is sponsored by SpectraCal.The present volume is a collection of the notes issued to date: Check the SpectraCal web site for issues, or the collection, dated later than this document's date. This collection is Copyright © 2011-12-02 Charles Poynton.

Charles Poynton
tel +1 416 535 7187
charles @ poynton.com
www.poynton.com

POYNTON'S VECTOR

1 CONTRAST, BRIGHTNESS, and the naming of things

Video professionals face a serious, long-term issue: Consumers don't know what knob to turn to make their pictures brighter!

For more than half a century, the two primary image controls have been called CONTRAST and BRIGHTNESS. That these controls are misnamed was observed half a century ago by the preeminent electronics engineer Donald Fink:

FINK, DONALD G. (1952), *Television Engineering*, Second Edition (New York: McGraw-Hill)

> "Unfortunately, in television systems of the present day, ... the separate manipulation of the receiver brightness and contrast controls (both of which are misnamed, photometrically speaking) by the nontechnical viewer may readily undo the best efforts of the system designers and the operating technicians."

In the 1950s, CONTRAST (which controlled video gain, as it does now) was apparently recognized by CE engineers as an important operational control – the CONTRAST knob was often concentric with VOLUME. On the other hand, BRIGHTNESS (which introduced an offset or bias) was apparently recognized as being an unfortunate technical necessity: On several television sets from that era, the BRIGHTNESS knob was placed between FOCUS and VERTICAL HOLD! I don't know if consumers in 1950 knew the difference between BRIGHTNESS and CONTRAST, but today the consumer that wants a brighter picture is just as likely – perhaps even more likely – to crank up BRIGHTNESS as CONTRAST, and thereby impair contrast ratio. This is the crux of Fink's complaint.

Fink also complained parenthetically about misnamed controls. It seems to me that if we retain the BRIGHTNESS control, we should relabel it as BLACK LEVEL, the term that is used on processing equipment and on many professional displays. However, with the functions of FOCUS and VERTICAL HOLD today taken care of by design, and those two controls abolished, we have to ask whether it is time for BRIGHTNESS to be abolished.

Consider a digitally encoded signal that is conveyed to the consumer in digital form – ATSC digital broadcast, DVD, or Blu-ray physical media. We must assume that black is correctly mastered, because we cannot distinguish creative intent from faulty production. No black level impairments are introduced by modern transmission systems, and the intrinsic display circuitry doesn't impair black or induce drift. BLACK LEVEL is not required to correct any of these historical issues.

One remaining potential justification for a display BLACK LEVEL control arises from the degradation of black-level luminance caused by ambient light. The degree of that degradation is determined by the diffuse faceplate reflectance. Considering the screen as a passive reflector, luminance is the ambient illuminance (in lux), divided by π, times the diffuse reflectance factor. Ambient illuminance of 10 lx reflected from a perfect diffuse surface produces luminance of about 3 nt. Twenty years ago, with a 20% reflective CRT faceplate, ambient reflectance would have produced 0.6 nt. With 100 nt white – bright at the time – faceplate reflectance limited contrast ratio to 100/0.6, or 160:1. Diffuse reflectance declined from perhaps 20% two decades ago to 10% one decade ago (at the pinnacle of CRT display technology); with the introduction of plasma displays it declined to about 2%, and for today's LCDs the value is about 1%. Today, ambient illuminance of 10 lx produces black luminance of about 0.03 nt. With white at 100 nt, diffuse reflectance alone can't reduce contrast below 3,000:1. Delivered contrast ratio is no longer dominated by diffuse faceplate reflectance.

Apple has apparently concluded that ambient-induced black shift is so insignificant today that they have abolished the BLACK LEVEL control entirely. Apple avoids the naming problem by labelling the control with a symbol and no words! However, some confusion remains because Apple uses the historical BRIGHTNESS icon to label the remaining control.

With an Apple display, it's clear to the consumer what knob to turn to make the display brighter: There's only one knob! I suggest that we consider the same idea in the video and HDTV arena, and abolish BRIGHTNESS (or BLACK LEVEL) from consumer displays. For the remaining control, I've been thinking about a better name than CONTRAST. My tentative suggestion is WHITE LEVEL.

It's another topic what luminance should be produced for an encoded zero-unit signal level. Perhaps I'll tackle that question in a future column. Meanwhile, I welcome your comments and suggestions!

p.s. *After first publication of this note, my colleague Cam Morrison pointed out my implicit assumption that faithful image portrayal is the goal. Cam correctly points out that manufacturers of television receivers do not usually have that goal: Their goal is typically to succeed in selling television receivers. Manufacturers have found that consumers respond to attributes other than faithful portrayal; manufacturers distort tones and colours according to what they think is most effective in attracting consumers. Among home theater enthusiasts, an important use of the* BLACK LEVEL *adjustment is to dial-out the manufacturers' preference in order to achieve the content creator's intent. Apple was able to abolish the* BLACK LEVEL *adjustment because Apple's products portray imagery rather faithfully, with a minimum of signal processing.*

POYNTON'S VECTOR

2 COLOUR, TINT, and the naming of things

In the previous column, I explored the controls called CONTRAST and BRIGHTNESS. This time, I'd like to explore COLOUR and TINT. As before, I'll typeset the names of controls in small capitals.

You may find it idiosyncratic, but I'll spell *colour* the way I did as a child – and in the manner that my wife and daughters insist – instead of the way I spelled it when I lived in California. Hey, we're all quirky in one way or another!

COLOUR and TINT controls arose historically from the mechanism of NTSC encoding. In analog NTSC, poor frequency response characteristics and differential gain errors often led to reduction of the amplitude of the modulated chroma signal. Broadcast technicians corrected those impairments manually by increasing chroma gain. Comparable facilities were introduced to consumers, labelled as COLOUR.

I argue that COLOUR is misnamed because the consumer can't be expected to know whether COLOUR means *which* colour or *how much* colour! Some professional equipment uses the name SATURATION. That name is a poor choice in my opinion because *saturation* refers to many other phenomena – for example, clipping of overexposed scene elements in a camera's image sensor. It seems to me that we should adopt the name CHROMA, as is used on some receivers: This name clearly suggests the amount of colour. It is intuitive that setting CHROMA to zero yields a greyscale image.

Going back to analog NTSC, differential phase errors often led to shifts in phase of the modulated subcarrier. Such shifts produced visual hue errors. In the worst cases the intended hue could only be established by manually rotating the decoder's subcarrier phase reference. Some professional decoders still today have PHASE adjustment; in consumer equipment the control came to be known as TINT.

WIKIPEDIA (2010), *Tints and Shades*, <http://en.wikipedia.org/wiki/Tints_and_shades>).

I argue that TINT is misnamed: To an artist, "to tint" means to add white, thereby lightening a colour *without* changing its hue! A quick check on Wikipedia or Google confirms the popularity of that interpretation. PHASE refers to the underlying technical mechanism, but we should not burden the consumer with a term dependent upon the implementation; rather, we should use a perceptual term. The obvious perceptual name appropriate for this function is HUE.

So, although COLOUR and TINT are popular among consumer receivers, CHROMA and HUE are, in my view, far preferable.

The BLUE ONLY feature of professional NTSC displays provided the video technician with a simple way to disable the red and green components of the displayed colours. In the colourbar test signal, the cyan and magenta bars both contain the same amount of the blue primary, and when displaying BLUE ONLY they should display identically. However, modulated subcarrier phase differs between the two;

only if HUE is set correctly will the decoded blue component values match. The white and blue bars both contain the same amount of the blue primary, but white has no modulated subcarrier. Only if chroma gain is set correctly will the blue decoded from the blue bar match the blue decoded from the white bar.

The HUE adjustment is meaningful only when decoding composite video. In a professional broadcast video monitor – "BVM," the subject of a future note – the CHROMA (or SATURATION) and HUE (or PHASE) controls are typically disabled when viewing a component input.

Providing a HUE control may be useful in program creation, but is highly unlikely to be useful as an expression of a consumer's viewing preference. My recently purchased LCD computer display has only $R'G'B'$ digital inputs. There's no composite NTSC input, and therefore no modulated chroma signal to correct; however, the display provides not only a CHROMA adjustment (there labelled SATURATION) but also HUE (labelled TINT)! Apparently this display's signal processing chain takes perfectly good $R'G'B'$, encodes to $Y'C_BC_R$, applies chroma gain and C_B/C_R rotation, then matrixes back to $R'G'B'$! In my view this "feature" is design engineering gone amok, or perhaps symptomatic of poorly informed marketing. HUE should simply be made correct by design. No useful perceptual attribute is addressed by rotating hue, and in component or digital video, no useful purpose is served by providing the consumer with a HUE control.

Poynton's Fourth Law:
Once a program is mastered, errors in mastering are, in principle, indistinguishable from expressions of creative intent.

Some people may argue that a HUE control enables correction of poorly mastered program material. To them, I assert Poynton's Fourth Law (in the margin). If you "corrrect" hue, what are you correcting to? In *The Hulk,* the protagonist's face is supposed to be green. Admittedly a HUE control could be used to render Hulk's face with normal skin tone, but isn't that detrimental to creative intent?

There is a minor reason that argues in favor of providing a CHROMA control to the consumer. A bright display (more than 100 nt), a bright ambient, and/or a bright surround may cause a systematic increase in colourfulness. If the consumer's viewing situation differs from that at mastering, maintaining creative intent may require dialing-back some of the colourfulness increase. This topic will be the subject of a future piece. I welcome your comments and suggestions!

POYNTON'S VECTOR

3 Gamma – 2.0, 2.2, or 2.4?

ITU-R BT.709-5 (2002-04), *Parameter values for the HDTV standards for production and international programme exchange*.

BT.709 refers to an international standard, first adopted in 1990, for HDTV. The standard defines the colours (chromaticities) of the red, green, and blue primaries and the white point (CIE D_{65}). These specifications are well established and widely used. How to accommodate wide color gamut is a challenge that might require some changes to BT.709 in the near future, but let's leave that for a later installment.

POYNTON, CHARLES (2010), *Picture rendering, image state, and BT.709* (unpublished, available at <http://www.poynton.com/notes/PU-PR-IS/Poynton-PU-PR-IS.pdf>).

In addition to specifying chromaticities, BT.709 also purports to specify nonlinear image coding. I have written extensively elsewhere about this aspect of BT.709. The short story is that the BT.709 story is wrong. As written, BT.709 documents *camera* characteristics, but what is needed is specification of a reference *display.* Without a reference display standard, there is no reliable mechanism to establish creative intent. We need a new standard that specifies characteristics comparable to those of a broadcast video monitor (“BVM"). A classic BVM has a gamma of about 2.4: Input $R'G'B'$ signals from 0 to 100 units are scaled by $^{1}/_{100}$, raised to the power 2.4, then scaled to the absolute luminance established for white.

If you are a home theatre calibrator, you might at this point be saying: “But I align my customers' displays to gamma of 2.2, not 2.4! Am I doing it wrong?" Well, perhaps not, but some explanation is in order.

HD programming was historically approved on “BVMs” – *B, V,* and *M* are the first three letters of the part numbers of a series of Sony displays: A Hollywood studio might routinely master on a Sony BVM-D32E1WU. Such a reference display would historically produce white luminance of 100 nt. Twenty years ago, content would have been approved in an ambient illuminance of perhaps 6 lx with a “dim" surround of perhaps 10% of reference white. Today, final approval is done in very dark conditions, with ambient illuminance 1 lx or less, and surround luminance perhaps just 1% of white luminance.

Colors change appearance depending upon absolute luminance, and upon their surroundings. A very dark surround at mastering will “suck" color out of a presentation previously viewed in a light surround. A colorist will dial-in an increase in colorfulness (for example, by increasing chroma gain). The intended appearance for an HD master is obtained through a 2.4-power function, to a display having reference white at 100 nt, with 1 lx ambient, and 1% surround – but that appearance will not be faithfully presented in different conditions!

The key point concerning the consumer's gamma is this: What we seek to maintain at presentation is the appearance of the colors at program approval, not necessarily the physical stimuli. If the consumer's display and viewing conditions differ from those at mastering, we may need to alter the image data to preserve appearance.

In a home theater environment, you might set the consumer's display to 100 nt, matching the approval luminance. However, ambient conditions in a consumer environment – even a rather dark home theater – are somewhat lighter than typically used for mastering today. The lighter conditions cause a modest increase in contrast and colorfulness, beyond that witnessed at content creation.

If the power function on *R'G'B'* – display gamma – is dialed back a little, that contrast and colorfulness are reduced. At about 300 nt, with ambient illuminance of 5 or 10 lx, and with a surround of say 5%, decreasing gamma from 2.4 to 2.2 will visually compensate the effect. So, if your consumer has such an environment, I recommend gamma of 2.2. If your customer preferred to display the same imagery at 48 nt, in darkness (zero ambient illuminance), in a 0% surround, then gamma of 2.6 (as in digital cinema) might be appropriate. In a really, really bright environment, or with a really bright display (say 400 nt or 500 nt), decreasing gamma to 2.0 might be appropriate.

EBU Tech. 3325 (2008), *Methods for the Measurement of the performance of Studio Monitors,* Version 1.1 (Sep.).

It is another issue how ten or twenty measurements of the greyscale curves can be distilled down to a single gamma number. I'm not enthusiastic about the EBU recommendation for the calculation: EBU Tech. 3325 (on page 34) calls for subtraction of the base luminance L_{MIN}; subtracting that bias is a mistake in my view. In the EBU document – and in most home theatre calibration packages – all of the measurements are effectively normalized by the luminance at reference white; however, that normalization gives the 100% measurement undue weight: That particular measurement is very likely to exhibit some saturation droop. On the other hand, we can't complain too much about the EBU technique, because SMPTE and ITU fail to give any guidance on computing effective gamma. I'll save the remainder of this argument for a future piece.

As always, your comments are welcome. ⁙

p.s. *After original publication of this note, ITU-R established (in March 2011) Recommendation BT.1886, entitled "Reference electro-optical transfer function for flat panel displays used in HDTV studio production." That document codifies a display power function exponent of 2.4. However, there are in essence two knobs – the Recommendation calles them "legacy BRIGHTNESS and legacy CONTRAST" that are open to setting by the viewer. In addition, reference white luminance is not specified, and viewing conditions are not specified. So, a small step has been taken in the right direction, but much work remains to be done.*

POYNTON'S VECTOR

4 Legal, valid, and sensible colors

Studio HD technicians and home theater calibrators are familiar with 10-bit HD coding: Reference black is placed at interface code 64 and reference white is placed at 10-bit interface code 940. (The comparable 8-bit codes are 16 and 235.) These communities are less familiar with the origins of footroom and headroom, and are generally unfamiliar with proper treatment of codewords that lie in these regions. I will address these topics in this column.

$$V_{709} = 64 + 876 \cdot \frac{units}{100}$$

Eq 1 Mapping from units to 10-bit digital video interface code

$$\left(\frac{1019-64}{940-64}\right)^{2.4} \approx 1.23$$

Eq 2 Peak luminance as a fraction of reference white luminance

Codes 64 and 940 are meaningful across a digital video interface; however, for signal processing it is much more convenient to declare reference black and white to be zero and unity respectively. Call those levels 0 units and 100 units if you like. In standard HD, headroom extends to about 109 units; the corresponding luminance is about 1.23 times reference white. *Reference* and *peak* are different! Many ITU-R, EBU, and SMPTE standards get this wrong, mistakenly using *peak white* when *reference white* is meant. Codes 0–3 and 1020–1023 are prohibited across an HD-SDI interface.

The original digital studio video standard was CCIR Rec. 601, established in 1984. Analog studio video signals drifted somewhat; to introduce digital video required accommodation of analog signals having imperfect reference levels. Footroom and headroom were necessary. That reason has now vanished. However, several good reasons for footroom and headroom remain.

When presented with an input signal containing high frequency content, any practical filter – whether analog or digital, or lowpass, bandpass, or highpass – necessarily involves some degree of undershoot and/or overshoot. Premature clipping of undershoots and overshoots is detrimental to image quality: Clipping should be deferred to the last possible point in the signal chain. Footroom and headroom accommodate undershoots and overshoots. Historically, it has been generally agreed – if not properly documented – that signals shouldn't dwell in the footroom or headroom region for longer than half a dozen samples. I'll bring this assumption into question below.

A relatively recent reason for footroom is to convey the negative excursions of camera noise. All sensors generate noise, even around black. When a camera is sensing true optical black (e.g., capped), it is sensible to set the average signal value to reference black. However, noise has excursions above and below that level. If the negative-going excursions are clipped, then the noise is said to be “rectified”: the average value of the noise then rises above reference black. It is important to defer the rectification to the latest possible point in the signal chain.

A final reason for footroom is that it allows coding of the blacker-than-black (–2%) PLUGE signal element commonly used to set black level in studio displays.

A camera engineer typically aligns an HD camera to produce 100 units for a near-perfect white reflector in the scene. However, when the camera is turned over to the cinematographer, he or she may wish to convey specular reflections or light sources in the scene, and he or she may therefore reset exposure to place the white card well below 100 units. Light sources and speculars may lie anywhere above the cinematographer's white reference, and the speculars and sources won't let up until they clip! The headroom region ends up carrying these elements. Typically these elements have momentary excursions, but they may well be sustained across more than half a dozen samples.

All of the reasons that I have mentioned are arguments against clipping anything in the footroom and headroom regions! Indeed, to clip is bound to introduce some degree of visual artifacts. Movie studios use commercial “QC” outfits to review commercial content prior to mastering. Many QC outfits are motivated to report “violations.” Most post houses take the easy way out and clip to “legal” before shipping content out to the QC houses. The potential visual quality of such material is compromised, and in the long term, the QC houses should be educated. Once the QC houses are well informed, the post houses will stop clipping.

Martindale, David and Alan W. Paeth (1991), “Television Color Encoding and 'Hot' Broadcast Colors," in Arvo, James (Ed.), *Graphics Gems II:* 147–158 (Academic Press, Boston).

Apart from footroom and headroom in the component digital environment, in the NTSC days, limiting chroma was important to prevent a transmitter from exceeding 120% power. That was generally accomplished through “NTSC legalization,” a process too complicated to be detailed here. Suffice to say that over-the-air NTSC analog transmission no longer occurs in the USA. No studio or consumer equipment operates under the 120% carrier constraint of the now-absent broadcast transmitters. No NTSC legalization is required – indeed, none is appropriate – for any signals today, even those in NTSC form!

Unfortunately, in the consumer domain, manufacturers are motivated to compromise the headroom region for two reasons. First, consider a display with maximum luminance of 250 nt. By properly following the gamma curve all the way up to 109 units (historically, “IRE”), peak luminance is 250 nt, but reference white luminance is 200 nt. If the display follows the gamma curve up to 100 units then clips, the manufacturer can claim reference white at 250 nt! Consumers tend to think bigger numbers are better – but the clipped picture will suffer. A home theatre calibrator, on assessing the greyscale response, will be savvy to this trick, and will adjust the display to follow the gamma curve all the way up to peak. Second, a manufacturer may be motivated to push the curve up in the midscale, then roll-off the higher regions in an effort to deliver higher average luminance. The well-equipped calibrator will similarly reverse the trick.

A well-known “standard” for wide-gamut, xvYCC, uses $Y'C_BC_R$ codewords outside the $R'G'B'$ unit cube to convey wide-gamut colors. To use such a system, gamut legalizers must be disabled. The wide-gamut arena will be the subject of a future column.

5 *B, V,* and *M* are obsolete

Broadcast video monitor: BVM.

Program production culminates with conversion of video data representing the finished program – think of it as a massive amount of *R'G'B'* image data – to colored light on the surface of a display. To enable consumers to witness a program as it was intended, the final production display must be characterized. To achieve faithful presentation to the consumer, the consumer's display must approximate the final production display. *B, V,* and *M* are the first three letters of the part numbers of a family of popular (and expensive) Sony displays commonly used for production: A Hollywood studio might routinely approve and master on a Sony BVM-D32E1WU.

I argue that in the modern age the notation "BVM" is seriously wrong. At first glance you might think that a part number could hardly matter, but I contend that the term and the part number exposes serious philosophical issues that are worth discussing.

First, the B is wrong. "Broadcast" reflects the old scheme where a very small number of program producers, aggregators, distributors, networks, and television stations were all part of a highly centralized distribution system. But YouTube doesn't "broadcast;" the iTunes store doesn't "broadcast;" Hulu doesn't "broadcast," Netflix doesn't "broadcast;" and when you watch a DVD or a Blu-ray disc, you're not experiencing "broadcast." None of the new, innovative, disruptive entrants in video distribution involve what I would term broadcasting.

Second, the V is wrong. Emergent displays such as RGB-LED backlit LCD displays offer primaries very close to the DCI P3 RGB primaries of digital cinema: These displays can be used for certain aspects of digital cinema creation, processing, and color approval. It's not just "video" that we're poised to approve.

Third, and most seriously, M is wrong. In BVM, M stands for "monitor." If you consult a dictionary, monitor is a passive verb: You are watching something as it goes by. A related concept is Quality Control (QC): If you're in a QC department, you examine imagery to make sure that no unintended impairments have been introduced, but you do not modify the content. The QC department may raise a red flag, but QC is not responsible for, and must not alter, the look of a show. The most important function of the display at the end of the content creation chain is approval, and *approval* is an active verb – program content is manipulated until its *R'G'B'* values create the intended visual impression. M for monitor is wrong.

So, if not BVM, what? I suggest “studio HD reference display.”

The word “studio” implies professional content creation. You may create content in your garage, but if you successfully distribute content at a reasonably wide scale, then that’s your studio.

I use the word “reference” because the display used at the end of the content creation chain establishes the intended reference for all downstream displays. If an identical display is present downstream in an identical environment, then it should present an identical picture. In the consumers’ premises, we don’t expect the tight tolerances of a studio display, but we do seek the same aim points.

My term uses HD instead of video because we don't expect any emergent SD displays. For consumer mastering, there's no need to refer to cinema because the consumer won't have access to movies encoded in digital cinema form for some time to come.

I use “display” because that is the generic term for the transducer that converts electrical video signal to light.

Color appearance is strongly influenced by surround conditions. My recent proposal for a new standard is entitled “studio HD reference display *and viewing conditions*.”

Some of you might have taken the title of this piece to suggest that BVM CRTs are dead. That conclusion is certainly true: The BVM-D32E1WU that I mentioned earlier has, in fact, been discontinued by Sony. I will address that general issue in a future piece.

Your comments are welcome.

p.s. *After first publication of this note, Sony demonstrated the BVM-E250 studio reference display, which uses AMOLED technology. See Poynton’s Vector Issue 11.*

POYNTON'S VECTOR

6 The demise of the CRT

Last month, I outlined my argument that the historical term "broadcast video monitor" is archaic and misleading; I proposed that we should say "studio reference display" instead. A more serious problem than terminology is that studio-grade CRTs are no longer available, having been discontinued by professional equipment manufacturers.

CRTs were heavy, bulky, and power hungry; alternatives to overcome those advantages have been sought for a long time. In the consumer market, many alternatives are available, primarily LCDs and PDPs.

WIKIPEDIA (2010), *Restriction of Hazardous Substances Directive*, <http://en.wikipedia.org/wiki/Restriction_of_Hazardous_Substances_Directive>).

The direct cause of the withdrawal of CRTs from the studio market was the introduction, in 2003 in Europe, of regulations to minimize or eliminate the use of lead in electronic devices, an initiative known as *Reduction of Hazardous Substances* (RoHS, often pronounced "roze" or "ro-haws"). Lead is a hazardous substance; lead was very widely used in solder in all electronic devices. Solder without lead has a high melting point, and electronic products have to be reengineered to accommodate changes to the manufacturing process in order to use lead-free solder. In addition to lead in solder, CRTs used lots of lead in the glass tube itself, to absorb X-rays (and incidentally to improve the optical properties of the faceplate). Presumably, studio CRT display manufacturers could have reengineered their CRT products to conform to RoHS standards. However, manufacturers apparently concluded that LCD displays achieving studio quality were right around the corner, so instead of reengineering their CRT-based products, they developed and commercialized what they deemed to be studio-grade LCDs.

Predictions of the performance of studio LCD displays were wrong. The first studio LCD products did not attain the visual performance of CRTs, and nearly all studio engineers agree that CRT performance is not reached even today in the latest models of LED-illuminated LCD studio displays. LCD contrast ratio isn't high enough (that is, blacks aren't dark enough), LCDs have poor luminance and colour uniformity on flatfields, and LCDs suffer viewing angle problems. Meanwhile, the gold-standard BVMs are aging and dropping, one by one, out of use. Studio engineers are faced with a vexing problem: On what display do you approve HD content?

The main purpose of a studio reference display is to mimic the best possible display in the consumers' premises. At first glance you may think that goal can be achieved by mastering on high-end consumer displays! However, even if consistency were to be achieved in consumer panels, there is a huge diversity in consumer-class signal processing. Also, consumer manufacturers are not motivated to provide faithful display of the image data; they are motivated to sell TV receivers. CE manufacturer seek to differentiate their products from those of their competitors. Consumer displays are inconsistent, so they cannot serve as reliable references for mastering.

Most studios are limping along at the moment with whatever BVMs remain operational, typically running at 80 nt white to extend the lifetime somewhat compared to the more desirable 100 nt. Some studios are mastering on the new breed of LED-backlit-LCD studio displays, but as I have mentioned most studio engineers declare these displays to have unsatisfactory performance. One prominent studio masters using a D-cinema grade DLP projector, set to BT.709 primaries; that's an excellent solution, although an expensive one. Some studios are now mastering using industrial-grade plasma displays, but custom colour-mapping machinery is required to bring the plasma gamma to a 2.4-power function; to bring the plasma primary colours into reasonably good conformance with BT.709; and to suppress any non-additive colour behaviour that may be found as a consequence of poor signal processing. Another complication is the luminance loading of plasma displays: When a full reference white flatfield is presented to the display, luminance drops according to a total power limit that corresponds to perhaps 35% of the small-field reference white.

I hope that acceptable studio-grade direct view displays become commercially available fairly soon. In the meantime, there's no viable technology-independent standard for a studio reference display. Disregard the potentially hundreds of variations of signal processing and display differences among consumer TV models, and consider a modest variation in consumer display and viewing conditions. Imagine three different consumer display luminances – say 100, 200, and 400 nt – and three different surround conditions – dark (1%), dim (5%), and average (20%). It is unreasonable to expect that programs should be mastered in nine versions, one for each of these conditions. Instead, content is mastered in one condition: 100 nt at 1% surround. That mastering condition should be built into a new studio standard. In my view, a display EOCF power ("gamma") of 2.4 is appropriate, and reference black video (code 16 at an 8-bit interface) should result in a luminance of about 0.03 nt.

A diversity of mastering displays are in use today, and it seems to me that no single display technology will dominate the future like the CRT did in the past. It remains a complex topic to determine how image appearance changes as a function of different display and viewing conditions, but it is clear to me that the HD community desperately needs a single, worldwide, technology-independent studio reference display standard. Your comments are welcome!

POYNTON'S VECTOR

7 Primaries for HD content

In last month's article I discussed the demise of the CRT as the definitive display device for approving colour content of HD programming. The inability of content creators to obtain CRT reference displays causes an immediate need for a replacement. However, even in the era when CRT reference displays were viable, there was a latent issue that no one talked about, and it's that issue that I'd like to raise this month: Worldwide standards for HD specify BT.709 primaries.

If you're a consumer video expert – perhaps a home theatre calibrator – you say, well, *of course, that's obvious!* But let's investigate.

ITU-R Rec. BT.709 is the main international standard for HD content. *R'G'B'* values are supposed to be displayed according to a particular set of primary chromaticities specified in the document. The Recommendation was adopted in 1990. In the years following, various national and continental standards followed suit. For example, SMPTE 274M, the main studio production standard in North America, states:

In 1990 the Recommendation was called CCIR Rec. 709.

> "Equipment should be designed in accordance with the colorimetric analysis … defined in this section. This corresponds to ITU-R BT.709."

Comparable provisions are found in ATSC, DVB, and EBU standards.

In 1993, the computer industry enthusiastically adopted the BT.709 primaries as one aspect of the sRGB standard. The sRGB standard is now ubiquitous for desktop computing, not just for stills (such as Exif JPEGs), but for video: sRGB video encoding is implicit in new video distribution channels such as YouTube, Skype, the Apple iTunes Store, Hulu.com, Netflix, and so on.

However, HD content creators hold a dark secret: Following international agreement on BT.709 in 1990, content creators in 60 Hz countries never switched from the SMPTE primaries of 480*i* SD to the BT.709 primaries, and content creators in 50 Hz countries never switched from the EBU primaries of 576*i* SD to the BT.709 primaries. Despite SMPTE, EBU, and DVB conformance to BT.709, HD content is generally *not* approved on BT.709 displays! SMPTE "C" (RP 145) primaries remain entrenched for 60 Hz HD in North America, Japan, and much of Asia; EBU primaries remain entrenched for 50 Hz HD in Europe and other parts of the world.

Sony's BVM-D32E1W CRT was the "gold standard" studio reference display. It was available with SMPTE or EBU phosphors – but was never offered with BT.709 phosphors! I am not aware of any studio-grade CRT that offered BT.709 phosphors.

BT.709 supposedly exists for international exchange; however, international HD program exchange in practice does not use BT.709. We find ourselves in the ironic situation that studio and broadcast HDTV does *not* use the “HDTV” (BT.709) primaries, but the computer industry has embraced them to the virtual exclusion of everything else! (Displays having Adobe RGB 1998 primaries are deployed in limited numbers for applications in high-end graphics arts.)

SD and HD luma coefficients differ, necessitating a transform of $Y'C_BC_R$ upon conversion between SD and HD. Within any geographical region the underlying $R'G'B'$ are identical, though, so the discrepancy in luma coefficients has absolutely no effect on colour gamut or “colour space.” By the way, there’s no such thing as “BT.601 colour space.” BT.601 is primary-agnostic. It is implicit that 480*i* content has SMPTE primaries and 576*i* content has EBU primaries.

High-end HD production and distribution standards conversion equipment (such as from Snell & Wilcox and Teranex) can be configured to perform colour primary transforms among SMPTE, EBU, and BT.709 primaries at the time that programme material is subject to standards conversion (say from 60 Hz to 50 Hz). However, if such a conversion is done, gamut clipping and gamut mapping issues are liable to arise. In my view, it is best to leave the $R'G'B'$ alone, and live with some small systematic colour errors (which are unlikely to be visible without comparison to the original) rather than risk gamut clipping (which is likely to be visible even without access to the original).

At the consumers’ premises, 60 Hz content has almost certainly been mastered with SMPTE primaries. If you’re a home theatre calibrator, you may be tempted to calibrate your display primaries to the SMPTE spec rather than BT.709. In my view that would be a mistake, for the reason that I mentioned a moment ago in connection with colour primary transforms in production: The risk of artifacts from colour clipping outweighs any advantage in colour accuracy – and in any event, when transmitted through an ATSC or DVB channel, the content effectively declares itself to be BT.709, so that’s how you should display it.

Studio display manufactures will face increasing difficulty in sustaining a pair of legacy chromaticity sets (SMPTE and EBU) for high end HD content, especially when the discrepancy with BT.709 confers no functional benefit. In the face of increasingly global distribution of television programmes, and with a single worldwide standard for desktop computing used in emergent video distribution technologies, it will be increasingly difficult for content creators, aggregators, and broadcasters to justify different primary chromaticities for 50 Hz and 60 Hz regions. The next 1 to 4 years will see replacement of CRTs across virtually the entire installed base of studio reference displays. Deployment of FPD studio reference displays (not “monitors”; see Issue 5) provides an opportunity for broadcasters throughout the world to migrate to BT.709 primaries. In my view, that’s the only sensible way forward is BT.709. If you’re a home theatre enthusiast, calibrating to BT.709 prepares you for this eventuality.

Wide-gamut television receivers are now commercially available that warp perfectly reasonable BT.709 colours into regions of colour-space that the content creators never intended: I call it *wild gamut*. Faithful presentation cannot be achieved with such processing. HDMI 1.4 and xvYCC/x.v.Color are implicated. The development of legitimate wide gamut will be the subject of a future issue. Meanwhile, I welcome your comments! ▦

Charles Poynton
tel +1 416 535 7187
charles @ poynton.com
www.poynton.com

POYNTON'S VECTOR

8 Wide gamut and wild gamut: xvYCC for HD

It's a complex topic, but I'd like to summarize – in two pages, without cheating the type size or the margins – the situation of wide gamut colour in consumer electronics, and specifically, xvYCC (x.v.Colour).

First, some history: The colour gamut of HD was agreed upon in 1990 in the standard (and therefore, colour space) now called BT.709.

In *Poynton's Vector* Issue 7, I described the disconnect between colourspace in content creation and colourspace in content presentation.

The colour gamut for digital cinema, agreed upon in 2006, is called DCI P3 RGB. The P3 gamut approximates the gamut of cinema film; it's wider than BT.709. Movies today are typically mastered in P3 gamut. When transferred to HD, a remastering step is undertaken, with "colour correction" performed by a skilled colourist who is sensitive to the creative intent of the movie and who is under the supervision of the director and/or cinematographer. Once the movie is mastered for HD, the creative intent is embedded in BT.709 colourspace. To be faithful to that intent, the movie (in HD) must be presented in BT.709 space. No consumer has access today – and no consumer is likely to have access anytime soon – to the P3 material.

NTSC and PAL video was negative-polarity amplitude modulated onto the RF carrier. The 120% limitation was to avoid undermodulation – you could call it "underloading."

Now, some deep background: In the old days of analog NTSC and PAL, video engineers limited *R'G'B'* values to the [0, 1] "colour cube." For 100% bars, the corresponding composite NTSC/PAL video limit is 133.33%. Transmitter power was set a little lower, to the equivalent of a composite limit of 120%. Video engineers came to use 75% colourbars so as to avoid exceeding the 120% transmitter limit. In the U.S., NTSC transmitters are now shut off. Analog transmission will soon cease in other countries. These limits are no longer relevant. (Certain broadcasters, post-production houses, and quality control outfits are creatures of habit; for no good reason, they still use "gamut alarms" and "legalizers" to clamp to the historical limits.)

That the BT.709 and sRGB gamuts are identical is no accident: Following agreement on BT.709, computer vendors quickly realized the benefits of adopting a single *RGB* coding worldwide. They incorporated the BT.709 primaries into the sRGB standard.

Fast forward to today: RGB LEDs are practical for use as backlights for LCD panels. They have colours close to the spectral locus of the CIE chart, so they enable wide-gamut displays. RGB LED backlights are commercially available in specialized displays for graphics arts, in some computer displays, and in some consumer television receivers ("LED TVs"). The question is: How should wide-gamut content be encoded and displayed? Clearly, sRGB content doesn't extend outside the sRGB colour gamut, and today's BT.709 content doesn't extend outside the 709 gamut.

One way to encode wide gamut colour is to adopt new primary chromaticities outside the current ones. That approach was taken with the Adobe RGB industry standard used in graphics arts. Many years ago studio HD engineers realized that another way to represent wide gamut – having a better compatibility story – was to leave the primary reference points fixed, and to allow negative signal excursions with respect to those primaries. That approach was documented for HD in 1998 in ITU-T BT.1361, but BT.1361 was never commercialized. A few years ago, the BT.1361 approach was resurrected in a somewhat modified form, and adopted in 2006 as IEC 61966-2-4, a sister specification to IEC 61966-2-1 (sRGB). That scheme is entitled *xvYCC* – xv for extended video, YCC for $Y'C_BC_R$. Sony trademarked *x.v.Colour* as its marketing term for xvYCC; when *x.v.Colour* is used for advertising, it is with Sony's permission.

The pitch from the xvYCC proponents in consumer electronics is this: At production, turn off the legalizers and gamut alarms. Convey wide-gamut through today's $Y'C_BC_R$ pipeline (digital broadcasting, DVD, Blu-ray, etc.), and we'll deliver wide-gamut colour to consumers.

No so fast. Consider little Nemo the fish. In the cinema, he exceeds BT.709 gamut. If the producer were to take *Finding Nemo* in P3 RGB and simply use textbook xvYCC encoding for delivery to consumers, and if that xvYCC data is decoded and displayed by legacy consumer equipment, Nemo might clip. His scales could be lost. He'd be liable to look like a blob of plastic instead of a cute cartoon character. Taking wide-gamut content to a legacy display presents a big problem.

On the other hand, consider legacy content on a new wide-gamut display. Upon introducing wide-gamut displays, marketers were not content to present legacy BT.709 content faithfully. Instead, they elected to "warp" legacy content to create new, saturated colours that they thought would entice consumers, but which content producers never experienced upon mastering the material. I call this *wild gamut.* The mapping can be defeated in various poorly documented ways – for example, by setting FILM, MOVIE, or THEATRE mode – but most consumers are unaware of these modes or don't understand them.

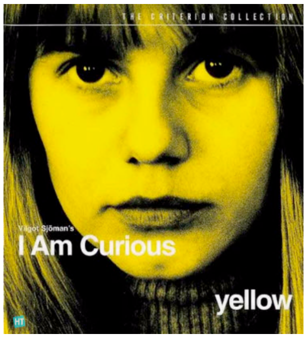
HDMI 1.3 endorsed xvYCC, and in addition included provisions that seem at first glance capable of conveying colour gamut boundary data (GBD) across the HDMI interface. However, a gamut boundary does not define a colour mapping. What is needed is a way for content producers to have control over – or at the very least, have knowledge of – colour mapping that will be applied to their content when it encounters wide-gamut displays. Imagine a CE design engineer excitedly asking a cinematographer, "How could I map your colours so as to improve your movie?" The cinematographer would in all likelihood answer something like, "Get your grubby hands off my movie!"

At the moment, content producers viewing legacy displays see their BT.709 material displayed reasonably faithfully. However, viewed on new, wide-gamut displays, their carefully chosen colours are being screwed up. Meanwhile, CE manufacturers are in effect saying, "Please give us your cinema-grade, P3 wide-gamut images!" Content creators reply, "Why, so you can screw *them* up, too?"

Resolution to the problem won't be soon in coming. Development of technology to maintain creative intent, and availability of wide-gamut content, are perhaps 3 or 4 or 5 years away. If you're a home theatre calibrator, endeavour to disable colour mapping for now, and hold out until things settle down. And watch this space!

POYNTON'S VECTOR

9 I Am Curious (yellow)



Sorry, I couldn't resist the title. It refers to a 1967 Swedish film by Vilgot Sjöman. There's a companion movie, *I Am Curious (blue).* The colours in the movie titles are those of the Swedish flag.

Sharp recently introduced several AQUOS Quattron™ television models, featuring "Quad Pixel Technology." Each display pixel has the usual three colour components (red, green, and blue) but adds a fourth component, yellow. I'm curious about how and why.

You may have thought that all colours could be formed from three additive primaries, but that was never completely true. When plotted on a CIE [*x*, *y*] chromaticity diagram, the path traced out by pure monochromatic source as it sweeps across the visible wavelengths from 400 nm to 700 nm takes on a "horseshoe" shape – the *spectral locus*. In additive colour formation with three primaries, it is possible to reach all of the colours whose coordinates lie within the triangle formed by the vertices. That's *most* of the colours, but no set of three primaries can reach *all* of the colours. Over the half-century from 1953 to about 2003, the television industry used a specific set of three RGB primaries (now standardized in BT.709) that covered most of the colours – let's call them the colours economically important in those times. In about 1993, computing industry reached *de facto* agreement to use the same set of primaries, which they called sRGB.

During the last 5 years or so, LED backlights have emerged that enable a wider range of colours – a wider *colour gamut* – than conventional backlights. One or more of the R, G, and B primaries is moved out closer to the spectral locus. The LED spectra are narrow, just 30 or 40 nm wide, a point to which we'll return in a moment.

Sharp's Quattron apparently uses a white LED backlight, instead of an RGB LED backlight. The white LED by itself offers no increase in gamut compared to a conventional fluorescent (CCFL) backlight; however, adding the fourth (yellow) component allows an increase in gamut in the yellow region of the spectrum (and a bit more in green).

Video content is mastered on studio-grade displays – typically CRTs – that exhibit near-perfect additive mixing. For faithful display, you also want near-perfect additive mixing! Any 4-primary display (or 5 or 6) needs signal processing to display colours reasonably close to those experienced by the content creators. Quad Pixel Technology involves signal processing that roughly mimics additive mixture of RGB, but substitutes a yellow stimulus for certain mixtures of red and green. The non-additive nature of the Quattron departs from the

Poynton's Vector Issue 8, "Wide gamut and wild gamut: xvYCC for HD".

creative intent of the material, a subject that I took up in Issue 8. When such a new display is driven with "legacy" image data, the wrong colours are displayed.

CE vendors are generally more interested in colorful images than in accurate images, but as far as I'm concerned, the mismatch between content creation and content display is a problem. I want faithful display of what the director experienced. Faced with a Quattron, a home theatre calibrator has to dial-out the enhanced yellow gamut in order to maintain the director's creative intent. As mentioned in Issue 8, we await a coherent industry-wide approach to mastering and displaying wide-gamut content.

On a related topic, I recently got e-mail from a studio engineer colleague who had dusted off his ancient Minolta TV-2150 TV Color Analyzer II, a colorimeter, to calibrate some studio reference displays. He discovered that a fairly new LCD professional display measured off-the-scale according to his meter.

Colorimeter filters aren't a perfect match to the ideal CIE spectral functions (*colour matching functions,* CMFs): A particular instrument might have response slightly too high in some regions of the spectrum and slightly too low in other regions. When measuring spectrally broad and smooth stimuli, the errors tend to average out; colorimeters do reasonably well. But some modern displays have narrow or spiky spectral content. A spike of the display primary might sample the instrument's response over a narrow range where there happens to be a mismatch with the CIE standard. This was the issue faced by my colleague; erroneous readings resulted. The worst case occurs with laser primaries, which effectively point-sample the instrument curves with no spectral averaging at all. If the instrument curve isn't accurate at *every* wavelength, problems arise. The issue isn't academic: Lasers were used to illuminate a DLP in a recently withdrawn Mitsubishi rear-projection TV model; Kodak recently demonstrated a laser-illuminated digital cinema projector.

Some colorimeters can be calibrated (typically using a 3×3 matrix) to achieve higher accuracy on a particular type of display. The approach is useful when measuring a set of displays that have the same spectral characteristics, say a bunch of Sony BVM CRTs, but is of little use in applications such as home theatre calibration that are faced with a diversity of display types.

Poynton's Vector Issue 6, "The demise of the CRT".

Returning to RGB LED backlights, they have narrow spectral bandwidths: They're "spiky." They are potentially incorrectly measured by classic colorimeters. The lesson for the systems calibrator is to use either a really accurate colorimeter or a spectroradiometer. I fear that my colleague's old TV-2150 TV Color Analyzer II will have to be retired. It was fine for CRTs, but is bound to be inaccurate for newer displays. And CRTs are disappearing fast; see Issue 6.

A final comment on the Quattron: If its white LED backlights resemble others common in the industry, they are essentially blue LEDs augmented by phosphor that converts visible blue light to what you might call yellow. The blue from the LED adds to the yellow from the phosphor to yield white. But the white has a blue spike that is liable to cause measurement trouble on classic colorimeters.

I welcome your comments and suggestions! ▪

POYNTON'S VECTOR 10 Video quality (VQEG)

Video Quality Experts Group, <http://vqeg.org>

WINKLER, STEFAN (2005), *Digital Video Quality – Vision Models and Metrics* (Chichester, U.K.: Wiley)

WU, HONG RENAND and RAO, K.R. (2005), *Digital Video Image Quality and Perceptual Coding* (Boca Raton, Fl.: CRC Press)

WANG, ZHOU and BOVIK, ALAN C. (2006), *Modern Image Quality Assessment,* Synthesis Lectures on Image, Video, and Multimedia Processing (San Rafael, Calif.: Morgan & Claypool, Publishers)

ITU-T REC. J.247 (2008), *Objective perceptual multimedia video quality measurement in the presence of a full reference*

If you're a video enthusiast – especially if you're interested in home theatre, or you're a home theatre calibrator – you're interested in video quality. A group called *Video Quality Experts' Group* (VQEG) sounds promising: VQEG has, for about 15 years, been studying objective estimation of video quality. There are hundreds of research papers, several books, and a few international standards.

The main concern of VQEG is the evaluation of the perceptibility of MPEG-class coding errors. We can use a panel of observers and perform a subjective test to evaluate MPEG compression, transmission, and decompression – for example, according to the protocol of BT.500 – but that's time-consuming and expensive. We can use "golden eyes" – trained observers – but that's also time-consuming and expensive. VQEG seeks what I call an "MPEG-o-meter."

The MPEG-o-meter, in one form, has two video inputs. One input presents a pristine, original "reference" video sequence. Another input presents the result of compression, transmission or recording, decompression, and perhaps some processing – the "test" sequence. On its (metaphorical) front panel is a meter that estimates the quality of the reproduction, compared to the original. More specifically, the device attempts to estimate the quality impairment that a human observer would report in a subjective test session. This technique, where the comparison has access to the original video and the impaired video, is called *full-reference* (FR). A full-reference algorithm starts by subtracting, pixel-by-pixel, the test from the reference. The difference reflects errors, which are then processed in a manner intended to mimic the visual system's sensitivity to image features.

If video transport or decompression shifts the picture six samples to the right and two lines down, a significant visual difference would be estimated by such an analysis algorithm. Similarly, if a compression system offsets luma +10 codes, scales luma down by 0.92, and scales chroma up by 1.08, full-reference analysis would start with a significant image difference. The MPEG-o-meter would report poor quality. I say, "Correct!" Picture shifting, scaling, and offsetting are errors. There's no good reason that transport or decompression should distort those aspects of the signal; such processing is properly deemed erroneous, and the estimated opinion score should be penalized.

However, VQEG offers bad news: The VQEG algorithms "normalize" – or, in what I consider to be a bizarre choice of words, "calibrate" – out such image data modifications! An active participant in the VQEG process told me that such "calibration" is necessary because many encoders reposition the image and introduce luma or chroma gain or offsets. VQEG experts apparently think video looks ok when subject to such changes; they take steps to null-out such modifications in their tests because if such changes were allowed to influence their results, their "measured" quality level would go way down.

People familiar with approval and mastering of high-value content know that scaling and levels are carefully controlled. Broadcast-grade compression systems don't arbitrarily reposition the picture, and they don't arbitrarily scale or offset luma or chroma. Studio engineers and home theatre enthusiasts are disturbed when such modifications take place in transport or at decompression. Take the notorious "DVD chroma upsampling error": Chaos ensued when certain consumer DVD player manufacturers decoded in a manner nonconformant with encoding standards. VQEG fails to penalize systems that introduce comparable errors.

It seems to me that we want to encourage natural – or should I say technological – selection: We want poor picture quality to cause selection pressure for improvement. If poor processing isn't penalized by poor scores, how will compression system engineers learn to do processing correctly? So far, VQEG's efforts have been used in multimedia, IP video, and teleconferencing; however, their efforts have – too put it bluntly – fallen flat in high-end space. Perhaps this is why.

VQEG is also evaluating "no reference" (NR) techniques, where quality assessment is attempted without access to the original material. With no reference, any VQEG NR method will inevitably estimate *Blair Witch Project* by as having horrible quality. As far as I'm concerned – and content creators presumably agree – the "no reference" idea is a non starter: If you remove the noise from *Blair Witch Project,* you destroy the movie! Many movies have unusual visual features. Content creators want flexibility to create whatever visual stimulus they like. From the point of view of telling a story through visual means, it seems to me that there can *never* be an algorithmic measure of what constitutes a "high quality" picture.

ITU-T Rec. BT.500-12 (2009), *Methodology for the subjective assessment of the quality of television pictures*

There's further bad news. The ITU BT.500 standard specifies subjective testing protocols and test conditions, but BT.500 prescribes display and viewing conditions wildly different from those used today to master high-quality program material. BT.500 specifies contrast ratio of 50:1 or 100:1, but approval and mastering today is typically performed at around 1000:1. BT.500 specifies a surround ratio of 0.15, but approval and mastering typically has a surround ratio of 0.01 or 0.02. BT.500 specifies ambient illuminance of around 200 lx (!), but approval and mastering facilities are typically illuminated between 1 and 5 lx. These are factor-of-10 or factor-of-100 differences!

Perhaps we can convince VQEG that "normalizing" luma and chroma is a relic of analog interfaces; that such a practice is not only unnecessary but flat-out wrong in the digital age. Perhaps we can convince VQEG to use, for its subjective tests, viewing conditions that are representative of the way in which high-quality material is mastered today. In the mean time, use your instruments, but take measured readings with a grain of salt. Learn to evaluate pictures visually; trust your eyes. I welcome your comments and suggestions!

POYNTON'S VECTOR 11 AMOLED displays

I recently attended the Hollywood Post Alliance (HPA) Tech Retreat; I presented a half-day seminar, *Studio Reference Displays.* Some of the topics that I discussed will be familiar to readers of *Poynton's Vector.* In Issue 5, I described how I am retraining myself to use the phrase *studio reference display* instead of *broadcast video monitor.* In issue 6, I lamented the demise of the CRT as a reference display. However, things are looking up. At HPA, Dolby (again) demonstrated their Professional Reference Monitor, which incorporates a spatially modulated backlight (which in Asia would probably be called local dimming). They have demonstrated it at previous events; this time, it's close to commercial availability.

Considering the part number prefix *BVM* (*broadcast video monitor*), Sony apparently didn't take to heart my suggestion to adopt a new acronym. See Issue 5.

As exciting as the Dolby display was Sony's demonstration of a 24.5-inch AMOLED studio reference display. It has a part number, BVM-E250, and it is expected to be commercially available in May, for a few tens of thousands of dollars. It comes from a new factory that has been quoted as requiring an investment of $200M (US). Clearly, Sony can't recoup that investment from professional markets alone: We can expect this display technology to be introduced into consumer electronics.

About 3 years ago, shortly after discontinuing CRTs, Sony introduced the BVM-L230 reference display. The L230 comprised an LCD panel with an RGB LED backlight unit. The introduction event took place in a dark room where two L230s were placed alongside a reference-grade CRT BVM; all were displaying the same picture. The visitor was challenged to identify which was the CRT. You could argue that the source material was carefully chosen, but in my opinion the Sony engineers quite successfully managed to get the L230 to mimic CRT behaviour. However, the L230 emitted about 0.15 nt when driven with reference black signal level (equivalent to a contrast ratio of about 700:1): The blacks weren't very black, and studio users weren't too impressed. The L230 model didn't do very well, and it was shortly followed by a somewhat improved model – the BVM-L231 – which still didn't produce blacks adequately dark for mastering content.

Sony's 3-display setup was repeated at HPA for the E250 demonstration; however in this case it was interesting that Sony made no attempt to mimic the studio CRT black level of about 0.03 nt (a contrast ratio of about 3000:1): When the video signal goes to reference black level, an AMOLED emits no light; the display is pitch black. The AMOLED images were quite different from those on the CRT.

The luminance produced by code 0 is relevant to the appearance of displayed pictures. Code 1 is also relevant. Let's compute:

Eq 11.1

$$\left(\frac{1}{255}\right)^{2.2} \approx 0.000\,005\,077 \approx \frac{1}{197\,000}$$

Eq 11.2

$$\left(\frac{1}{219}\right)^{2.4} \approx 0.000\,002\,415 \approx \frac{1}{400\,000}$$

Eq 11.3

$$\left(\frac{1}{858}\right)^{2.4} \approx 0.000\,000\,091 \approx \frac{1}{11\,000\,000}$$

The first equation calculates the relative luminance expected from the 8-bit sRGB code value used in personal computers, assuming 2.2-gamma and no ambient light. Code 1 (on a scale of 0 to 255) has equivalent contrast ratio of about 200 000:1. In any reasonable viewing environment, no one could be expected to see that light.

Here I use video code level 0 for reference black and code 219 for reference white. You might be expecting codes 16 and 235, but those are *interface* codes. Calculations such as mine are greatly simplified if the interface offset of +16 is first subtracted. See Issue 4.

The second equation calculates the relative luminance expected from the HD or SD 8-bit code values typical of consumer equipment, assuming 2.4-gamma, and again assuming no ambient light. Code 1 (now on a scale of 0 to 219) has equivalent contrast ratio of about 400 000:1. Again, no one could be expected to see code 1.

The third equation calculates the relative luminance expected from 10-bit studio video coding, with the same assumptions. Equivalent contrast ratio is 11 million to 1. Again, no one can see code 1.

Digging deeper into the visibility of video signal codes, we can ask, what is the ratio of luminances produced by codes 1 and 2?

Eq 11.4

$$\frac{\left(\frac{2}{219}\right)^{2.4}}{\left(\frac{1}{219}\right)^{2.4}} = 5.278; \quad \frac{\left(\frac{22}{219}\right)^{2.4}}{\left(\frac{21}{219}\right)^{2.4}} = 1.118; \quad \frac{\left(\frac{219}{219}\right)^{2.4}}{\left(\frac{218}{219}\right)^{2.4}} = 1.011$$

At the left is the answer for the 8-bit consumer video situation. The luminance ratio between the first two codes is about five to one: Going from code 1 to code 2 multiplies the luminance by five. At first glance, that factor is surprisingly large; however, it turns out to be entirely consistent with the behaviour of vision at very low luminances. As code value increases, the ratio diminishes; at code 22 (about 10% video level), the luminance ratio between adjacent video codes is about 1.1, and at reference white, the ratio has fallen to about 1.01 (the nominal one percent *Weber contrast* of vision science).

Poynton, Charles (2009), "Perceptual uniformity in Digital Imaging," in *Proc. Gjøvik Color Imaging Symposium* (GCIS 2009): 102–109.

The traditional power law of video is well matched to perception, even at eight bits. But, home theatre calibrators know that 8-bit plasma panels – the very early ones – performed very poorly, exhibiting severe banding artifacts. Why? They were 8-bit *linear light* devices. The luminance ratios between adjacent panel driving codes didn't match visual perception. The solution for PDPs was to increase the bit depth driving the panel – and to add spatial and/or temporal dither.

For AMOLED displays, one open question is this: What is the native "law" by which the OLED driving codes are translated to luminance? I could tell from examining the BVM-E250 images that its characteristic wasn't linear. However, it's highly unlikely that AMOLEDs exhibit the power-function behaviour of CRTs. When I find out, and piece together the implications for image quality, I'll let you know!

POYNTON'S VECTOR

12 Gamma estimation

7.5-unit black level ("setup") is a relic of 525-line SD analog video. It has no place in 625, no place in digital video, and no place in HD.

Many video engineers and textbooks mistakenly use the word *luminance* to refer to the video signal *Y'*, which I argue should be called *luma.* Luminance can't be computed from luma alone.

The ITU-R's 2.4 value is appropriate for a 100 nt display in a dark viewing condition. In *Poynton's Vector* issue 3, I discussed how that value should change for brighter displays or brighter viewing conditions.

Gamma in video systems is a numerical parameter related to the mapping from video signal values to light at the display screen. Video systems use various signal coding schemes, from analog voltage to 8-bit or 10-bit digital codes, with or without provision for footroom ("blacker-than-black") and headroom ("whiter-than-white"). We'll simplify by assigning reference black the value 0 and reference white the value 1; you can call these 0% and 100%.

Luminance properly refers to a specific spectral weighting associated with the lightness sensation of vision; it can be measured by an instrument such as a photometer or a colorimeter. *R'G'B'* signals are presented to a colour display, and red, green, and blue light is produced. We are interested in the luminance of the red, green, and blue components. The colour science term for the linear-light *RGB* quantities, relative to their reference values, is *tristimulus values.*

A display's mapping of signal to light is described by an *electro-optical conversion function* (EOCF) that characterizes light produced by the display at the display surface. Displays are typically viewed in environments having a small amount of ambient illuminance which diffusely reflects at the display surface and contributes to luminance (as veiling glare). We exclude that light from the EOCF.

It's a sad story of studio practice that for the last 5 decades we haven't had a proper standard for gamma. However, a few months ago the ITU-R recommended, in Rec. BT.1886, gamma of 2.4 for HD studio reference displays. In high-end video (such as home theatre), we seek to mimic the image appearance at mastering. The desired EOCF is very close to a pure power function of the form $L(V) = V^{2.4}$.

Now, some math. A pure power function, when plotted in log-log coordinates, plots a straight line whose slope is the value of the exponent. In our case, log of tristimuli versus log of video signal plots a straight line whose slope ($^{\text{rise}}/_{\text{run}}$) is the value of gamma. However, many displays exhibit departures from pure power functions, and if the departures are large enough, visual disturbances might result. We'd like to characterize the overall gamma.

We can estimate gamma with just two measurements. As the "run" along the *x*-axis, use video signal values 0.1 and 1, a decade – that is, factor of ten – apart. As the "rise" on the *y*-axis, take the logs of the corresponding luminances L(0.1) and L(1). We would expect these two luminance values to be about 2.4 decades apart – that is, we expect L(0.1) to be about $10^{-2.4} \cdot L(1)$, that is, about 0.004 of reference white.

You can estimate gamma as the log of the ratio of those luminance values, or these additional two ratios that I will describe in a moment:

Eq 12.1

$$\gamma \approx \log_{10}\frac{L\left(1\right)}{L\left(0.1\right)} \approx \log_{10}\frac{L\left(0.8\right)}{L\left(0.08\right)} \approx \mathbf{\frac{5}{3}}\log_{10}\frac{L\left(0.8\right)}{L\left(0.2\right)}$$

Even in home theatre displays, reference white luminance is liable to be depressed a little owing to saturation. Effects of saturation on the gamma estimate can be reduced by using video signal values 0.08 and 0.8, still a decade apart: Gamma can be estimated as the log of the ratio L(0.8) to L(0.08), the second case in Equation 12.1. However, video code 0.08 is getting uncomfortably close to black level. The third case in Equation 12.1 takes a $^{3}/_{5}$-decade range of video signals (0.6 = log $^{0.8}/_{0.2}$) and estimates gamma as $^{5}/_{3}$ times the log of the corresponding luminance ratio.

Home theatre calibration software has historically calculated gamma not just from two measurements but ten, taking this average:

Eq 12.2

$$\gamma = \frac{1}{9}\sum_{i=1}^{9}\frac{\log_{10}\dfrac{L\left(\frac{i}{10}\right)}{L\left(1\right)}}{\log_{10}\frac{i}{10}}$$

The log of zero is minus infinity: We omit video signal value 0. Caution is necessary if any L($^{i}/_{N}$) – particularly L(0.1) or L(0.2) – is close to zero, perhaps owing to poor black level setting.

This equation is not as scary as it seems at first glance. We take ten display luminance measurements, at video signals 0.1, 0.2, …, 1.0. The summation is over nine samples; the average gamma is ultimately just the average of quantities related to those nine samples.

Within the average are nine ratios; each numerator is a log, and each denominator is a log. The ratio is $^{rise}/_{run}$, that is, the slope. The fact that logarithms are taken in both the numerator and denominator reflects computation of the slope in log-log coordinates.

In the ratio within the summation, the numerator reflects luminance; we take the log of the ratio of each luminance step L($^{i}/_{N}$) with the luminance of reference white L(1). The denominator reflects the video signal; we take the log of each step $^{i}/_{N}$ of the video signal. The reference white video signal has the value 1, we need not divide $^{i}/_{N}$ by that. The quantities being "logged" are all less than one; the logs are negative, and each ratio is positive.

In brief, we are averaging the slopes of nine lines in log-log space. Denoting each measurement point [x, y], each line joins [log $^{i}/_{N}$, log $L(^{i}/_{N})$] with the reference white point [log 1, log L(1)].

A problem is now evident: The reference white luminance L(1) occurs nine times in the sum. That single luminance value has nine times the influence on the average gamma than each of the other nine values. If you seek to obtain a meaningful display gamma estimate using this formulation, you must avoid saturation at 100% video.

Researchers working toward the adoption of the ITU-R studio EOCF standard didn't establish gamma using the approach of Equation 12.2. Instead, they used a numerical optimization technique to estimate the parameters of an equation modelling the EOCF. Such an optimization technique should, in my view, be adopted by home theatre practitioners. Such a proposal is likely to be the subject of a future issue.

POYNTON'S VECTOR 13 Black level

The subject of black has been raised before in previous issues of *Poynton's Vector.* In Issue 1, I discussed adjustment of black level. In Issue 4, I discussed "legal" and "valid" codes and the footroom codes below reference black. In Issue 11, I computed the relative luminance levels to be expected around reference black. Several additional aspects of black deserve further discussion.

The term *setup* implies insertion of an offset such that the range of video levels from black to white is reduced so as to maintain the white level. The term *pedestal* implies insertion or alteration of an offset without this correction.

Let's start with "setup." In ancient times – the 1940s, when vacuum tubes ruled electronics – raster scanning on a CRT was prone to cause retrace lines to be visible. Video signal levels drifted owing to temperature dependencies and component aging. Video engineers established *blanking level* of the video signal, intended to cause the beam to fully extinguish, and *black level,* taken as the lowest level of meaningful picture information on the way up to white. The blanking to white range was quantified in 100 units, later standardized by the Institute of Radio Engineers (IRE, the predecessor of the IEEE). Black level was established at 7.5 units, leaving 92.5 units for the picture range. That was 1941; colour television was still a decade away.

By 1964, at the introduction of 625-line video, European engineers decided that circuitry was sufficiently stable that separating blanking and black was no longer necessary. Setup was abolished for 625/50.

In the U.S., 7.5-unit setup was retained for compatibility reasons. Japan adopted the NTSC system, with its 525-line scanning and its 7.5-unit setup; however, Japanese television engineers soon realized that the 7.5-unit setup was useless. In about 1985, they undertook a country-wide effort to reconfigure the entire broadcasting infrastructure to eliminate setup. I suppose that many consumers adjusted brightness to compensate; I suppose that many didn't. But Japanese analog SD material from that era into the present day has zero setup.

The +16 interface offset of BT.601 provides signal processing "footroom." It was not intended to mimic or provide for setup.

Upon establishment of the first studio digital interface standard – BT.601, in 1984 – setup was abolished from the digital domain. Eight-bit coding was agreed to have reference black at interface code 16 and reference black at interface code 235. The codes below 16 were provided for filter undershoot and accommodation of analog tolerances, but no provision was made – nor was it necessary – for setup.

When it became feasible to do substantial arithmetic processing of video in the digital domain, in order for black to stay black – say, upon multiplication by a gain factor – the interface offset had to be removed. The footroom region must be represented in digital arithmetic using negative numbers. I came to denote the reference black and white levels in terms of *processing levels* after removal of the interface offset. For me, eight-bit digital video has *processing* reference levels 0 and 219. If footroom codes are clipped – as is common in PC video – quality suffers. Reference black remains at processing code 0 even with increased bit depth such as 10 bits. The math of video is easiest using reference levels 0 and 1, again using negative numbers for the footroom region (and numbers up to about 1.09 for headroom).

In 1984, at the introduction of digital video in the studio, it was clear that analog interfaces would remain in use for quite some time, both in studios and in consumers' premises. In studios the situation was complicated because Sony and Panasonic introduced different ill-considered analog component interfaces that involved setup or vestigial setup elements. In the consumer domain, the S-video analog interface was introduced in 1987 for S-VHS videocassette recorders. It had 3 versions: an American version with 714 mV video and 7.5-unit setup, a Japanese version with 714 mV video and zero setup (NTSC-J), and a European version with 700 mV video and zero setup. When DVD was introduced in 1995, its SD component analog interface needed the same three versions; however, many DVD player manufacturers got the details wrong; consumer image quality suffered.

Standardization of HDTV followed the European lead: HD analog interfaces never had setup; digital HD followed the BT.601 example.

Computers are digital! Application software should not, in my opinion, mimic setup in the SD analog domain. If you have old-fashioned users, teach them that setup is obsolete.

Upon the introduction of DV consumer recording in about 1995, many camcorders incorrectly recorded black at 8-bit interface codes below 16, apparently in a misguided attempt to "maximize dynamic range." I'll have more to say about dynamic range in a future issue.

If wide-gamut colour is ever to be deployed in content creation, it will almost certainly use the xvYCC (x.v.Color) scheme whereby wide-gamut colours are coded into values that reach below reference black or above reference white – that is, lie outside the range 0 to 1. In Issue 4, I commented that content creators – or more specifically, the "quality control" houses that they hire – should disable their clippers. Clipping was once necessary to avoid overmodulation in analog NTSC transmitters, but those transmitters have now been decommissioned.

The story so far has many false starts and premature optimizations. The correct way forward clearly has no setup! Reference black should be called 0 no matter what bit depth is in use, even in the abstract math domain where reference white is represented as the value 1. It is entirely reasonable to use percentages – refer to video levels as 0 percent and 100 percent, if you like – but with the sole exception of analog 525-line SD, black should be taken as zero, not 7.5 percent.

I'll leave you with one subtle aspect of studio mastering practice. BLACK LEVEL is set in the studio using PLUGE; the resulting luminance is about 0.03 nt. Reference black has nonzero luminance! Modelling the display EOCF as a pure 2.4-power function, reducing video level below reference black drives display luminance down to theoretical zero luminance. On the 10-bit interface code scale, with reference white and black at codes 64 and 940 respectively, theoretical zero luminance is obtained at about code 32. Accommodating filtering undershoot is one reason to retain footroom codes in video content. The capability to represent theoretical zero luminance is another reason.

Charles Poynton
tel +1 416 535 7187
charles @ poynton.com
www.poynton.com

SpectraCal

Poynton's Vector

14 Colour management systems (CMS) in video

I've discussed *colour management systems* tangentially in previous issues, but it seems timely to dedicate an issue to the topic.

Virtually all commercially deployed imaging systems are based upon $R'G'B'$ colour components. Several "flavours" of $R'G'B'$ are in use. By *flavours* I mean that the colours of the pure primaries – technically, the *chromaticities* of red, green, and blue – differ. SD and HD today both use the chromaticities of ITU-R BT.709, adopted 20 years ago. Fifteen years ago, the BT.709 primaries were adopted in the sRGB standard; that coding is now ubiquitous in computing and digital still photography. In graphics arts, sRGB is common, but Adobe RGB 1998 is sometimes used, and Pro Photo RGB is occasionally used.

$R'G'B'$ components may be transformed into different colour representations. $Y'C_BC_R$ is used for digital video and HD; $X'Y'Z'$ is used for digital cinema distribution; and the CMYK – cyan, magenta, yellow, and black – system is used for graphics arts.

About 15 years ago, several companies active in graphics arts, computer graphics, and digital photography collaborated to devise an architecture and a standard to enable predictable mathematical colour transforms among devices. This scheme evolved into the International Color Consortium (ICC) standard for *colour management system* (CMS). Device-specific colour data is stored in *profiles.* The transform for a display ("monitor") is a fairly simple combination of 1-D lookup tables and a set of coefficients for a 3×3 matrix. The profile for a CMYK printer is a combination of 1-D lookup tables and one or more 3-D lookup tables, usually large and complex. Even without looking inside a profile, you can guess whether you have a monitor profile or a printer profile by simply examining the file size: A matrix profile for a display is typically a few KBytes; a 3-D LUT for a CMYK printer is typically a few hundred KBytes.

If image data encoded to one set of RGB primaries is naïvely "plugged-in" to a display having different primaries, colour is not displayed correctly. In the ICC approach, an image data file can contain an embedded profile that specifies the colour space of the image data. An ICC-savvy receiving application can potentially process the image data and transform it appropriately for the display primaries. In some cases, out-of-gamut colours may result. A 3-D LUT profile can incorporate gamut mapping to accommodate such situations.

Video has historically used no such technology. Apart from troublesome but fairly minor variations (SMPTE, EBU, BT.709), video has historically “locked-down” its primary chromaticities in paper standards. Consistent colour display has been enabled with no cumbersome processing at the receiver and no requirement for metadata.

During the last decade, the ease of incorporating VLSI into television receivers has led many CE manufacturers to wonder what additional colour adjustments could be offered to entice consumers. A “colour management system” (CMS) offers such additional adjustments. Perhaps the term CMS was adopted from graphics arts colour management, perhaps it was independently contrived; either way, the CMS scheme now fairly common in video displays and video processors bears no resemblance to the ICC’s CMS scheme.

In video, a *colour management system* (CMS) enables independent control of hue and saturation – and sometimes also luma (or luminance) – of red, green, blue, yellow, cyan, and magenta. As far as I can tell, CMS circuits are always situated in the $Y'C_BC_R$ path, prior to $Y'C_BC_R$-to-$R'G'B'$ matrixing and prior to the display’s EOCF.

You might ask, why are black and white missing from the list of colours above? Well, they’re accounted for elsewhere. The luminance of white is already controlled through CONTRAST. Hue and saturation of white are already controlled through choice of colour temperature. The luminance of black is already controlled through BLACK LEVEL. Hue and saturation of black are already indirectly controlled through RGB-BIAS (sometimes called RGB-SCREEN, RGB-OFFSET, or RGB-LOW).

But back to red, green, blue, cyan, magenta, and yellow. You think, “Great! With a CMS, I can control everything!” Not so great, I argue.

The studio reference displays upon which professional content is mastered exhibit near-perfect additive behaviour. At content mastering, cyan is always exactly the sum of blue and green. Magenta is exactly the sum of blue and red, and yellow is exactly the sum of green and red.

As a home theatre calibrator, how can you recreate the director’s experience at mastering if cyan isn’t the sum of blue and green? Well, you can’t. You must seek additivity. You must bypass the CMS (if that is possible), set the CMS to pass $Y'C_BC_R$ unchanged (if the display implements that option), or make sure the CMS settings achieve additive mixing (in the absence of detailed documentation from the manufacturer). Clearly, you’re faced with a challenge. You may require individual control of *R, G,* and *B* chromaticities to attain the BT.709 primaries; however, is it reasonable that CE manufacturers force you to use CMS controls to achieve that goal?

In computer graphics and consumer digital photography, “colour management systems” provide a way for image receiving equipment to effect colour transforms to deliver the colour intended by the content originator. In video, it’s exactly the other way around: “colour management systems” features defeat the goal of additive mixture, and thereby defeat the content originator’s colour intent!

It’s no wonder that content creators shun proposals to originate wide gamut colour to CE equipment, proposals such as using xvYCC to deliver digital cinema’s P3 RGB gamut. Movie makers think, “The CE industry can’t even present today’s BT.709 colour content correctly; how can we trust them with theatrical-quality colour?”

Your comments are welcome!

POYNTON'S VECTOR

15 Order-of-operations

It's an uncomfortable, unspoken truth of home theatre calibration, even in high-end products, that three-quarters of the effort is expended in dialing-out poorly chosen factory presets. Television receiver manufacturers are motivated to sell receivers, not to deliver accurate imagery. A receivers that manufacturers thinks is set up to deliver the "wow" factor in a fluorescent-lit showroom will certainly produce garish, over-bright, over-colourful images in a living room or a den.

It's unfortunate that BT.709 does not establish gamma. I wrote about that issue in Issue 3. A new ITU-T standard is forthcoming. Until then, use 2.4 in a typical living room environment at about 100 $cd \cdot m^{-2}$.

Video content is mastered on displays that exhibit characteristics very close to those specified in BT.709. To deliver a visual experience comparable to the experience of the program creator, you seek a calibration as close as possible to BT.709. (It is a question for the next few years whether content creators will master content in wide-gamut colour. If that takes place, then the principle may change.)

Projectors are typically delivered with factory settings that produce images resembling BT.709 calibration. However, at least one manufacturer is promoting "personalization" as a strategic direction. In my view, this is a dead end. Consumer manufacturers are already inserting their own personalities, distorting the imagery from the creative intent. I see no advantage in high-end manufacturers doing so. Even home theatre calibrators have trouble understanding colour management systems on those devices in which CMS capability is provided. To my mind it is completely unrealistic to think that you could explain to a consumer how to adjust a CMS to personalize his video content. To me, the value proposition for home theatre remains faithful presentation, recreating the director's experience of his or her own movie.

What is the best order in which to make adjustments when calibrating a video display, projector, or and television receiver to BT.709?

You can get a good start by choosing an ISF or THX preset if one is available. (One disclaimer: Although I was intimately involved in the recent additions to ITU-T standards for studio reference displays, I am not entitled to read the ISF and THX "standards." We are left to wonder what's in those standards, and to what extent commercial considerations intrude into the certification process.)

To calibrate to BT.709, we hope that the receiver's signal path resembles the path implicit in BT.709. That path has these steps:

- NTSC/PAL/MPEG decode;
- $Y'C_BC_R$-to-$R'G'B'$ dematrix;
- gain and offset (CONTRAST and BRIGHTNESS);

- individual channel gain and offset, and GAMMA (EOCF) control;
- display (perhaps including BACKLIGHT control).

If we were able to able to access video data after processing, at the point where processed data is presented to the panel, it would be easy to establish signal path adjustments. For example, in all modern digital display technologies we expect the reference black video signal to present the all-zeros code to the display. But we don't have access to the panel data, so we have to measure the display's light output instead. My suggested order of adjustments is based upon the following principle: Measure the thinnest layer possible – the shortest signal path. Correct that, then add layers one-by-one until you reach the front end of the system and the longest possible path. Imagine the receiver block diagram with conventional left-to-right signal flow. At each step, insert your test signal into the signal flow as close to the display as possible, that is, as far to the right as possible.

Receiver design engineers are apparently instructed to provide lots of controls to users, but I don't think they deliberately set out to confound calibrators. In all likelihood each control's midscale setting is the "unity" setting where processing is effectively bypassed. So, unless otherwise indicated, set each adjustment to its midscale detent. In the absence of a detent, use the Poynton algorithm: Set each control to the numerical average of the lowest and highest presented values.

If you can inject computer $R'G'B'$ (ranging 0–255), do so, to eliminate errors in CE (16–235) to IT (0–255) level conversion. Use a test signal that matches the native panel pixel count, to bypass resampling.

If the receiver has a BACKLIGHT control, this will be effected right at the display. Set it first, so as to achieve maximum required luminance.

Set individual $R'G'B'$ BIAS if these controls are provided. (These may be called SCREEN, OFFSET, or RGB-LOW.) They are likely to be optimum at midscale.

Set individual $R'G'B'$ GAIN if these controls are provided. (These may be called DRIVE or RGB-HIGH.) They are likely to be optimum at midscale. These may interact with COLOUR TEMPERATURE (WHITE REFERENCE).

Set BLACK LEVEL. For any modern digital display, it's likely to be optimum at midscale.

Set the poorly-named CONTRAST control to achieve the desired display luminance (e.g., 100 $\text{cd}\cdot\text{m}^{-2}$, representative of HD mastering).

Set GAMMA, if gamma adjustment is available.

The next block back toward the input is $Y'C_BC_R$-to-$R'G'B'$ dematrix. This block effects the BT.601/BT.709 luma coefficient setting, and in all likelihood incorporates CHROMA and HUE. Set these to unity.

If the receiver or projector has a so-called colour management system (CMS), it is almost certainly implemented in the $Y'C_BC_R$ domain. Do your best to defeat it. If the native display primaries differ from BT.709, the correct approach to compensate is to apply 2.4-gamma, transform the primaries in a linear-light 3×3 matrix, then impose the inverse of the display's EOCF. If that sounds complicated, it is. Few manufacturers do it. Many manufacturers bury comparable correction in the CMS, but compensation there will never perfectly match proper BT.709 dematrixing, decoding, and display.

Good luck! Your comments are welcome!

POYNTON'S VECTOR 16 What's "home theater calibration"?

POYNTON, CHARLES (2003), *Digital video and HDTV algorithms and interfaces* (San Francisco: Morgan Kaufmann).

I have written at length in previous issues – and in my book – about highly technical aspects of video signal processing, display, and calibration. Most of my writing is directed to people having background in math, physics, and electronics. In this note I seek to address everyday issues associated with obtaining high quality video at home.

I've always had trouble with the word "reproduction" applied to presentation, in the home, of professionally created audiovisual content. Taken literally, "reproduction" suggests "production again." However, viewers don't want to do the production again, they simply wish to display it. We seek proper *presentation* of recorded (and in today's world, streamed) media.

High-quality viewing (and listening) is satisfying for many sources of programming, including sports, home video, and even YouTube clips. However, the most demanding video content available to consumers is represented by movies delivered on Blu-ray. Movies delivered by cable or satellite are promising, but these sources are liable to have impairments introduced by severe compression in the transmission channel, so for our purposes we'll take our source to be Blu-ray.

To many people involved in home theater, the main goal is to recreate the filmmaker's vision in your home. Some people phrase the goal as "preserving the creative intent"; however, it seems to me that truly determining the director's artistic intent would be a deep philosophical (or perhaps even legal!) quest. It is sufficient for our purposes to establish the director's *experience* upon mastering the content: Audio and video enthusiasts seek to recreate, in their living rooms, dens, or home theaters, the *experience* of the director when the movie was created. The nub is this: What was the state of the display used to master the content, and what environment was it in?

Home theatre enthusiasts traditionally use motion picture film as their reference point for quality. It is implicit that display standards are in place when film is projected in a theater. Theaters have technical standards concerning presentation; however, consumer electronics (CE) equipment does not. It is a goal of home theater calibrators to bring home equipment into conformance with presentation standards.

Movies transferred to Blu-ray are mastered on studio-grade reference displays. They used to be called "broadcast video monitors," but that term is falling out of fashion as "broadcast" becomes less important in content creation, and as mastering displays are used not just for video but also for film. (See Issue 5.)

A studio-grade reference display has BT.709 primary chromaticities, CIE D_{65} white reference, a 2.4-power electro-optical function (EOCF), and 100 $\text{cd}\cdot\text{m}^{-2}$ reference white luminance. If that all sound technical, it is. For faithful presentation, you'd like to mimic the technical parameters of the studio. A good home theater calibrator needs the technical knowledge and skill to understand these parameters and instruments suitable to measure them.

CE manufacturers do not necessarily – or even usually – have the goal of approximating a studio display. Instead, they seek "good-looking pictures," whatever that means to any particular company. Most companies seek to sell receivers and displays; they do not generally seek to recreate the movie maker's experience. CE equipment diverges from studio standards for several reasons. One reason is that consumer equipment is sold in showrooms, not in living rooms: CE displays are optimized for over-bright viewing environments. Another reason is product differentiation: If all CE displays approximated studio displays, all CE displays would look alike, and the image displayed on a Sony would then appear the same as the image on a Samsung! CE vendors seek to "enhance" or "improve" the picture with technologies such as Sony's *X-Reality PRO Engine* or Samsung's *HyperReal Engine.* Such processing can be advantageous on really poor video sources – think 1980s vintage VHS tapes – but when as a consumer you have access to the exact bitstream that the movie-maker placed on a Blu-ray disc, at least as far as the movie-maker is concerned, "improvement" is not possible. It is a sad fact of calibration that half or two-thirds of the job involves "dialing-out" poorly chosen factory presets. You rely upon your calibrator's knowledge of particular manufacturers, makes, and models to accomplish that task.

I have been describing what you could call the scientific aspects of calibration. If you are running a small business producing computer-generated imagery (CGI) or visual effects (VFX) serving the movie business, then your own viewing preferences aren't relevant; what matters is bringing your equipment into conformance with the appropriate standards, and straight science and engineering suffices.

However, in a typical consumer home theatre, there are aspects that require judgement. Movies transferred to Blu-ray are mastered in very dark conditions (around 1 lux) that are almost certainly darker than your viewing preference. Your family members and guests would run the risk of tripping! If you prefer a somewhat brighter viewing environment, subtle changes in the technical parameters are necessary to achieve the intended appearance. You rely upon your calibrator to understand the issues, to assess your viewing environment, and to make the correct technical choices. If a calibrator has experience in post-production, that's a plus: A good understanding of the mastering environment helps to establish the home viewing environment.

Charles Poynton
tel +1 416 535 7187
charles @ poynton.com
www.poynton.com

POYNTON'S VECTOR 17 Linearity and the *-nesses*

Two recent incidents prompt me to write this month about terminology. Many people find discussion of terminology to be boring. However, according to a Zen master that I once met, the path to wisdom begins with calling things by their proper names.

The first incident concerns my scanning the glossary of a newly issued book about VFX and CGI – the second edition, nonetheless! The book shall remain nameless. I encountered this definition:

> **Linear color space:** A color space in which the relationship between a pixel's digital value and its visual brightness remains constant (linear) across the full gamut of black to white.

This passage is confused. First, the definition is circular: The definition of *linear* includes the word *linear.* Second, the concept of *linear* (as used in imaging) has only a tangential connection to *gamut.* Third, if it's a *pixel,* it must be digital: Pixels are by definition discrete; there are no analog pixels. Most seriously, "visual brightness" is a dangerously imprecise component of this definition. What characterizes *linear* in digital imaging is that a pixel component value (typically *R, G,* or *B*), over the active range, is proportional to *physical light power.* Vision has nothing whatsoever to do with it.

ENGELDRUM, PETER G. (2000), *Psychometric Scaling* (Winchester, Mass.: Imcotek Press).

The word *brightness* in the definition above is an example of what my colleague Peter Engeldrum calls a *-ness word.* According to Peter, *-ness* words are invariably related to perceptual quantities, none of which can be directly measured. Seeing the suffix *-ness* is a tip-off that the quantity being discussed isn't physical. But the whole *raison d'être* of a linear colour space is to form a direct connection to the physics of the imaging situation. For example, a scene-linear workflow has been introduced by Florian Kainz and his colleagues at ILM, and is being further developed by the Academy as the Image Interchange Framework (IIF) for digital cinema production.

www.oscars.org/science-technology/council/projects/iif.html

My definition has twelve more characters, but one fewer word:

> **Linear colour space:** Colour data wherein each component value, over most of its range, is proportional to light power – that is, proportional to radiance, intensity, luminance, or tristimulus value.

The qualifier "over most of its range" is necessary because linearity fails when component values below zero or above maximum are clipped.

To be complete, *linearity* accommodates an additive offset (bias) term as well as the (multiplicative) proportionality factor.

The second incident relates to my attending a recent seminar at which video calibration was discussed. A graph similar to Figure 17.1 was presented:

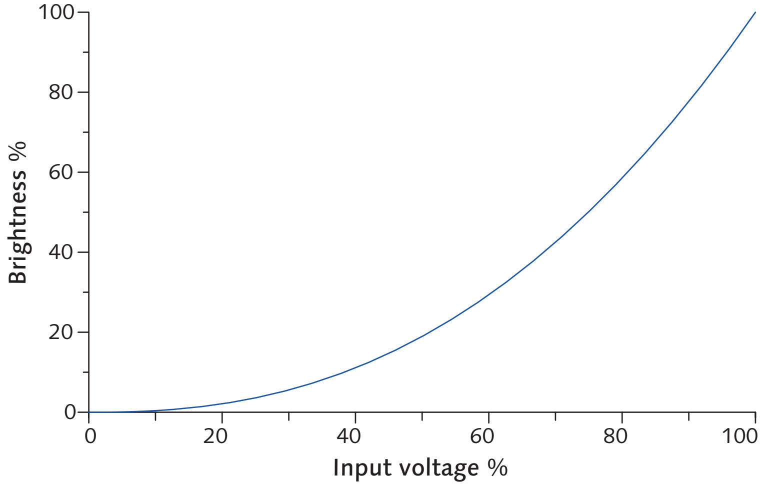


Figure 17.1 **Putative relationship** in avideo display takes an input voltage and produces "brightness."

In another seminar that I attended, the presenter called the video signal *stimulus*. This is wrong. *Stimulus* is properly the stimulus to perception, not the input to a display system.

What's wrong? The *x*-axis is labelled *Input voltage*, but modern displays are driven by digital signal values. More seriously, the *y*-axis is labelled *brightness*, but the display produces a physical quantity, tristimulus or luminance: It is the viewer's visual system – not graphed here! – that converts the physical light stimulus to a percept, a *-ness.*

The graph is corrected simply by relabelling the axes (Figure 17.2):

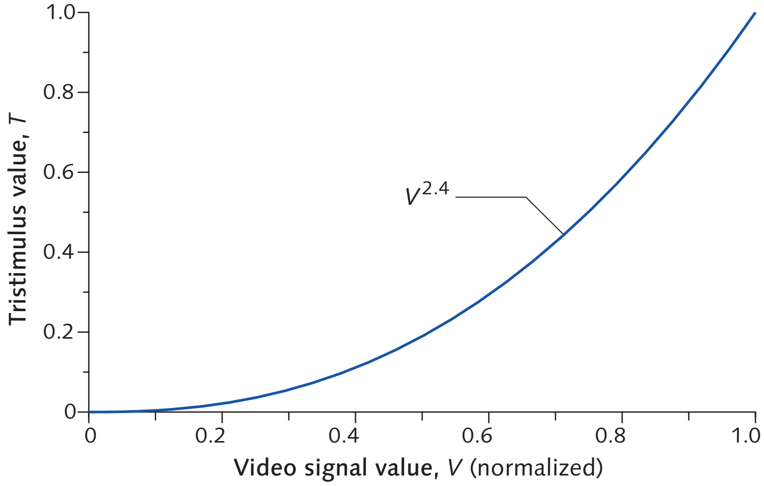


Figure 17.2 **Display EOCF** takes an *R'*, *G'*, or *B'* video signal component value – here in the range 0 to 1 – and produces the corresponding *R*, *G*, or *B* tristimulus value (relative, by definition). The luminance of each tristimulus value can be measured individually. The recently adopted ITU-R BT.1886 calls for a 2.4-power relationship for studio HD.

Figure 17.2 presents the *electro-optical conversion function* (EOCF) of a display. A historical CRT's light output is approximately the 2.4-power of voltage. (Modern displays incorporate signal processing to mimic CRTs.) Human vision perceives tristimuli or luminance nonlinearly: The perceptual response is approximately the tristimulus value raised to an exponent of 0.42. It is an amazing coincidence that a CRT exhibits the inverse of the characteristic of human vision! Video decoding with a CRT is a near-perfect match to vision. A video signal of 0.5 produces half the lightness. This fact is central to the design of video systems. This vital insight only emerges if you use the correct terminology.

Wisdom arises from referring to things using their proper names! ▪

Eq 17.1

$$0.42 \approx \frac{1}{2.4}$$

$$0.5^{2.4} \approx 0.189;\ \ 0.189^{0.42} \approx 0.5$$

Charles Poynton
tel +1 416 535 7187
charles @ poynton.com
www.poynton.com

POYNTON'S VECTOR

18 LUTs in 1000 words

"LUTs" – that is, *lookup tables* – are widely used in digital cinema (D-cinema), professional HD, and consumer HD. LUTs are found within cameras, within boxes used on-set, within post-production equipment, in external boxes used with display systems, and within certain displays. The contents of LUTs are established using a variety of techniques. In this article, we'll explore LUTs in various applications.

We start with displays. Practical D-cinema and HD displays produce additive mixtures of red, green, and blue primary components. Many display systems exhibit near-perfect independence of the three components (*channels*); however, even with channel-independence, most displays have a nonlinear relationship between signal value input and light output from each channel. For example, studio CRTs historically produced light power proportional to the 2.4-power of signal input. For high-quality display, the goal is to match the mastering display's conversion of signal to coloured light, and so the viewer's display must incorporate the 2.4-power or something close to it.

The 2.4 "gamma" is standardized in ITU-R BT.1886. See Poynton's Vector Issue 12, *Gamma estimation*.

Some display technologies, including PDP and DLP, use pulse-duration modulation. These displays exhibit near-perfect linear behaviour in each channel. In order to approximate the BT.1886 display "gamma," a suitable power function must be implemented in the signal processing chain. The required *degamma* function is typically implemented in three independent 1-D LUTs. Each LUT contains a mapping that implements the same power function. Signal values in the three channels are processed independently; the contents of three 1-D LUTs can be scaled independently to alter white balance.

By *linear,* I refer to digital image coding where across a large range of light power levels, halving of light power is associated with halving pixel value. In *logarithmic* coding, across a large range of light power levels, halving light power is associated with decreasing pixel value by a certain increment. (For example, in 10-bit Cineon coding, halving of scene exposure decreases pixel value by 90).

A PDP or DLP display may have primaries and white point that differ from the interchange standard (BT.709/BT.1886 for HD). Owing to the near-perfect additive colour mixing behaviour of PDP and DLP displays, mapping from the BT.709 interchange primary set such to the native primaries of a particular display device can be accomplished by a 3×3 matrix multiplication operation specified by 9 numbers: Small, predetermined fractions of each channel's value are added to or subtracted from the other channels. Three independent 1-D LUTs can't do the job of mixing channels together. (If the display primaries match the interchange primaries, the matrix isn't necessary.)

LCD displays typically exhibit a small degree of unwanted "coupling" between the channels. The required correction is a nonlinear function of the three signal component values. Three 1-D LUTs alone can't do the job, because that wouldn't allow channels to combine. A matrix can't do the job, because the required compensation isn't linear. Several other approaches could potentially work, but the practical solution is to implement a lookup table that is accessed by a combination of all three signal component values, and that produces all three values required by the display. If input and output values were 8-bits each, then we could use a table having 256×256×256 (that is, about 16 million) entries, where each entry contains three bytes. Such a table occupies about 48 megabytes. For professional applications, and for still images, that size isn't necessarily out of the realm of possibility; but for HD at about 60 megapixels per second, even at the high end of consumer space, it isn't practical. The solution is to rely upon the mapping being relatively smooth, and to *interpolate* – in simple terms, average – across a few thousand entries instead of a few million. You might get away with a table 10×10×10 – that is, 10^3 or 1000 words – but LUTs used in practice have perhaps 17^3 (that is, 4913) entries, and can easily have 10, 12, or even 16 bits per component at the output. Interpolation is typically either *trilinear* or *tetrahedral.* This *3-D interpolated LUT* solution allows compensating the subtle nonlinear mixtures exhibited by LCDs, but that treatment is necessary for only the highest quality studio displays.

When I say *3-D* here, I'm not referring to stereoscopy!

Other applications of 3-D LUTs are found in cinema. Camera film exhibits nonlinear mixture among colour components; so does print film. In film, the coupling ("crosstalk") components aren't linear, and so can't be removed just with a matrix. If you seek to remove these components, the most practical way is to use a suitable 3-D LUT. On the other hand, if your application seeks the film "look," you may need to *insert* such crosstalk components. You can use a 3-D LUT to do so. But by the time such imagery is made available for consumer viewing, the material has been mastered in a perfectly additive colourspace such as BT.709, and approved in that colourspace, with or without the "film look." If the display device is additive, a 3-D LUT isn't necessary downstream. For digital cinema exhibition, material is mastered in the DCI P3 *RGB* colourspace, also perfectly additive.

We've discussed 1-D LUTs and 3-D LUTs, but you may be asking, "What about 2-D LUTs?" In principle, an acquisition technology or a display technology could have two colour components that interacted nonlinearly, but a third that was independent of the other two. In that case, a 2-D LUT could be used to combine or compensate the two interacting channels, leaving the third alone. In practice, colour devices nearly always either exhibit channel independence or they exhibit interaction among all three channels, not just two of them. In principle, you use a 2-D LUT to alter chroma components – such as x and y, or C_B and C_R – leaving the third (luminance or luma) channel untouched. However, in practice, manipulation of all three channels is usually required; so, 2-D LUTs are very rarely found.

To calibrate a display – whether on-set, in the post-production suite, in a review room, or in a home theatre – you are tasked with bringing the display system into conformance with standards such as BT.709 and BT.1886. Generally, a 3×3 matrix and a set of three 1-D LUTs will suffice. To calibrate an LCD for studio use, a 3-D LUT may be necessary. If you are brought into the cinema production or computer-generated imaging/visual effects (CGI/VFX) world and asked to remove or insert the "film look," a 3-D LUT will be needed.

This article has 1000 words, excluding words in marginal notes!