

NUOVE TECNOLOGIE PER LA PRODUZIONE PRINCIPI TEORICI DI TELEVISIONE HIGH DYNAMIC RANGE

di ING. SERGIO BRIGHEL Technical Director, Trans Audio Video s.r.l.

Continuiamo ad occuparci di uno degli argomenti più attuali, l'HDR. Dopo il primo articolo introduttivo, proposto nel precedente numero di Digital Production, ecco un ulteriore-ma non ultimo-approfondimento focalizzato in particolare, sulla visualizzazione e sullo spazio colore.

Un sistema di ripresa composto dalla Amira della Arri, accessoriata con lo zoo 17-20 della Canon. Le nuove camere che montano sensori di largo formato, da 1" o Super-35 sono ora in grado di riprodurre il doppio del range dinamico di un display SDR.



Ing. Sergio Brighel
Technical Director - Trans
Audio Video



LE

telecamere moderne riescono a riprendere immagini in un range dinamico di 14-16 F-stop. Uno "stop" indica un raddoppio di luminosità; dunque si riferisce alla possibilità di sostenere un incremento della luminosità in progressione geometrica conservando la riproduzione dei dettagli, dai neri più profondi ai bianchi (o alle tinte cromatiche) più saturi. I monitor commerciali convenzionali riescono però a visualizzare soltanto 6 stop lineari non compressi.

Quando la Rec.709 HDTV (da ora indicata nel testo come SDR) fu ratificata, 6 stop sembravano sufficienti dal momento che tale intervallo contiene interamente il range dinamico delle comuni superfici non riflettenti: il bianco opaco è circa due stop più luminoso del grigio con riflettività al 18%, mentre il nero opaco si trova dai due ai tre stop più in basso. Tale compromesso fu dettato principalmente dalle limitazioni costruttive dei display, realizzati unicamente con la tecnologia dei raggi catodici in un tubo a vuoto, e che imponevano limiti invalicabili per le luminosità minima e massima ottenibili. Il bianco di picco raramente superava le 100-200 candele a metro quadro (unità di misura del S.I. per la luminanza di un'immagine, indicata con cd/m^2), mentre il nero, pur potendo virtualmente raggiungere un livello molto basso, risultava alterato dalle riflessioni sulla superficie lucida dello schermo.

In ogni caso, si riusciva ad ottenere un rapporto di contrasto dell'ordine di 1000:1. In un'epoca come quella attuale, dove l'innovazione tecnologica procede a ritmo incalzante, è sorprendente notare come questa tecnologia sia stata l'unica disponibile per la visualizzazione d'immagini televisive per più di cinquanta anni. In seguito LCD e Plasma hanno sostituito tali dispositivi, permettendo di realizzare schermi di dimensioni maggiori con ingombri e pesi inferiori; al principio con risultati qualitativi discutibili. Difatti, per una luminosità massima sostanzialmente invariata rispetto alla tecnologia precedente, si riscontrò un incremento del livello minimo del nero e conseguente deterioramento del contrasto. A peggiorare la situazione, la maggior parte dei dispositivi consumer erano (e tuttora sono) caratterizzati da uno schermo a superficie lucida. La scelta è dettata principalmente da ragioni di marketing, poiché da spento il monitor appare più gradevole alla vista (ma non solo: uno schermo opaco comporta inevitabilmente un leggero blurring delle immagini); ma ciò degrada la prestazione sui neri profondi per la riflessione della luce ambientale e per il riverbero degli stessi spettatori sullo schermo quando sono illuminati dalla luce emessa dalle porzioni chiare delle immagini visualizzate.

Progressi nella tecnologia di visualizzazione

I limiti tecnologici illustrati per un display SDR impediscono la riproduzione fedele delle sorgenti di luce da specularità, dei riflessi, delle luci incidenti in camera e, all'estremo opposto, dei dettagli nelle ombre più scure o delle situazioni di luce non controllata, come i contrasti da background

luminosi presenti in ambienti illuminati più debolmente (ad esempio, le finestre in una stanza). Quanto descritto è la causa primaria del "look televisivo" così riconoscibile e al quale gli spettatori sono stati abituati per decenni.

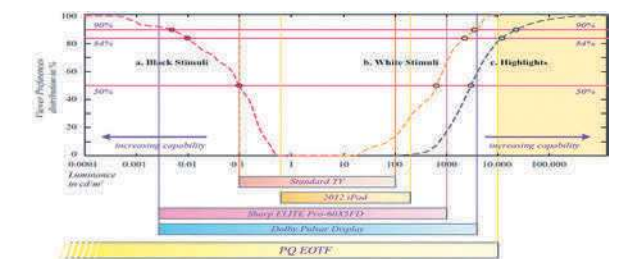
Ma, a distanza di circa venticinque anni dalla prima ratificazione della televisione SDR (ITU BT.709-1, 1993), le camere professionali meno performanti (specialmente con sensori di largo formato, da 1" al Super-35) sono ora in grado di riprodurre il doppio del range dinamico di un display SDR [Immagine 1].

Tuttavia, anche le tecnologie di visualizzazione si sono evolute nello stesso periodo e, nel caso di quelle più moderne (OLED oppure LED RGB local dimming), è possibile la riproduzione del range dinamico completo di una camera con una compressione tonale soltanto minima e con colori più saturi di quanto sia mai stato possibile in precedenza. Laddove la televisione SDR ha requisiti tecnici soddissfatti da monitor con soli 6 stop di dinamica, oggi un display che è classificabile High Dynamic Range (HDR) deve averne almeno 11-12, a partire da neri (e corrispondenti livelli di rumore) molto bassi, fino a livelli di luminosità massima certamente elevati.

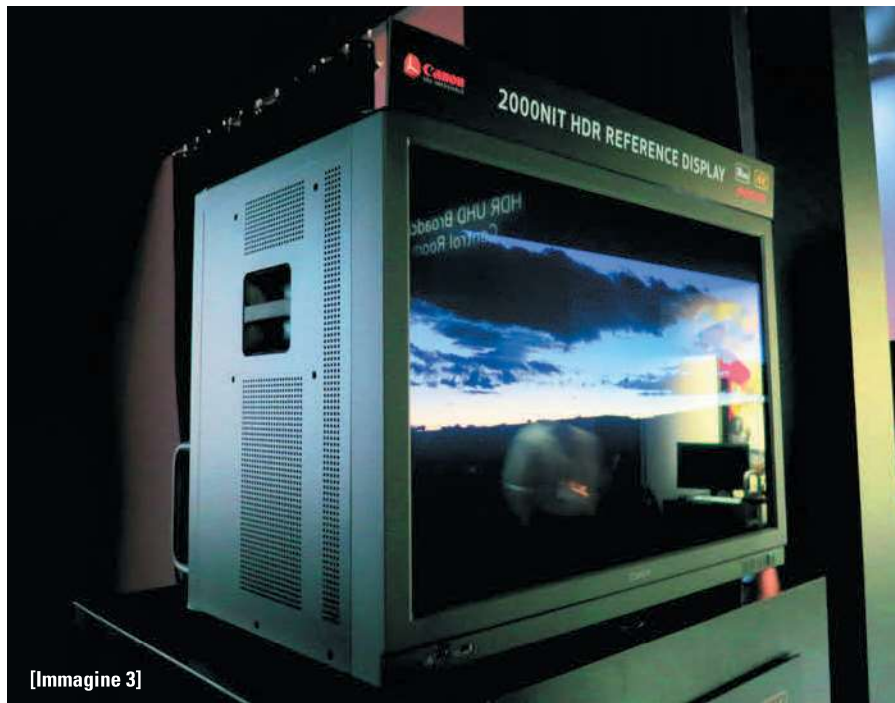
Come anticipato, la luminosità di un display è misurata in candele per metro quadro, anche dette "nits" (dal latino nitere: brillare). Un monitor SDR grado 1 calibrato emetterà al massimo 100 cd/m^2 per essere conforme alla raccomandazione internazionale ITU BT.709. Un monitor consumer SDR può arrivare a picchi di 400-500 cd/m^2 . Per la televisione HDR gli standard ratificati prevedono la riproduzione di neri profondi fino a 0.0001 cd/m^2 e picchi fino a 10.000 cd/m^2 [Immagine 2].

Se da un lato quest'aspetto desta meraviglia e aspettativa, dall'altro pone quesiti importanti per la sostenibilità energetica su larga scala (il consumo di monitor HDR consumer attuali, con luminosità fino a 1000-1500 cd/m^2 è in media del 50% maggiore al corrispondente in SDR di pari dimensioni), sulla produzione di calore e sulle dimensioni laterali dei dispositivi.

Uno degli aspetti più critici per display LCD e OLED, di fatto quello più limitante tecnologicamente, risiede nella necessità di dissipare rapidamente il calore accumulato tra i pixel, considerando che gli spessori dei pannelli sono davvero ridotti. Ciò impone sistemi molto raffinati per la conduzione e la successiva eliminazione del calore con dispositivi di ventilazione forzata. Il consumo energetico aumenta ulteriormente, così come dimensioni laterali e pesi. Al NAB 2017 era mostrato in anteprima un impressionante monitor Reference Canon 4K-HDR da 30" (tecnologia LCD IPS retroilluminato LED con microzone RGB), in grado di generare luminosità di picco di oltre 2000 cd/m^2 . Le sue dimensioni laterali erano altrettanto impressionanti! [Immagine 3]



Un monitor SDR grado 1 calibrato emetterà al massimo 100 cd/m^2 per essere conforme alla raccomandazione internazionale ITU BT.709. Mentre, un monitor consumer SDR può arrivare a picchi di 400-500 cd/m^2 . I test, hanno dimostrato, utilizzando immagini reali e in computer graphics, che il 50% del campione trova gradevoli le riproduzioni comprese in un range da 0.1 a 650 cd/m^2 ; per incontrare il 90% dello stesso campione, occorre estendere il range riprodotto nell'intervallo da 0.005 cd/m^2 a >3000 cd/m^2 8.



[Immagine 3]

La vera natura del High Dynamic Range

Il termine HDR resta molto ambiguo, difficile da comprendere nelle sue implicazioni, senza il supporto di un'esperienza visiva reale. E, una volta osservata una sequenza televisiva in HDR, si percepisce immediatamente il suo significato; soprattutto, una volta posti a confronto due monitor HDR identici (oppure un monitor che supporti il dual view, in altre parole con lo schermo diviso in due parti identiche in grado di mostrare impostazioni differenti per la stessa sequenza [Immagine 4]), regolati per visualizzare simultaneamente lo stesso contenuto in modalità SDR e HDR, si sperimenta un vero e proprio "imprinting" che istintivamente - e profondamente - concretizza a livello percettivo la vastità del salto qualitativo.

Il punto più sorprendente riguarda la rituale domanda sulla risoluzione delle immagini. Chiunque si trovi di fronte ad un contenuto HDR non la pone mai per prima; la percezione della risoluzione è infatti soggetta a molteplici fattori: distanza dallo schermo di visione e grandezza dello stesso; post-produzione, condizioni generali di visione, rumore medio delle sequenze.

E' probabilmente il fattore meno rilevante tra quelli che definiscono la qualità di un'immagine: risoluzione spaziale (numero pixel in un quadro), risoluzione temporale (numero di quadri/secondo), quantizzazione (numero bit per pixel), contrasto (range dinamico) e spazio colore (percentuale coperta del gamut cromatico del sistema di visione umana) [Immagine 5]. All'aumentare del range dinamico dell'immagine, gli aspetti qualitativi immediatamente prioritari diventano il gamut cromatico e la risoluzione temporale. Non è un caso che i broadcaster siano fortemente attratti da una televisione HD-HDR@50p piuttosto che da una UHD-HDR come prossimo obiettivo realistico! Uno studio condotto dalla BBC dimostra che ci sono forti deterrenti alla convenienza qualitativa del 4K, principalmente a causa delle dimensioni medie degli ambienti domestici e della dimensione media degli schermi.

Osservando una sequenza in HDR, l'attenzione è diretta-

ta su aspetti d'impatto psicologico talmente profondo da lasciare incantati: la brillantezza delle scene, il range dei colori e la loro saturazione: il loro "volume" (termine che ricorrerà nel seguito) arricchito dalla verosimiglianza dei riflessi e delle sorgenti speculari; parimenti, la profondità delle parti più scure dell'immagine che non nascondono più i dettagli delle forme contenute. Si pensi ai riflessi della carrozzeria di un'auto quando è illuminata dal sole. La loro intensità aiuta a definire le forme. Si pensi al ricordo di un falò, di un vulcano in attività ripreso di notte, o di fuochi artificiali vissuti in prima persona, delle insegne colorate al neon, delle luci posteriori led di un'auto.

Ebbene, quell'idea di "saturazione cromatica", è solo un pallido ricordo in ogni riproduzione televisiva in SDR (in cinematografia l'esperienza è differente per una diversa dinamica delle pellicole, per le particolari condizioni di visione a bassissima luminosità [48 cd/m² in DCI]), ma anche per il diverso modo di percepire il movimento delle immagini).

Tutte queste componenti della comune esperienza visiva sono assenti per costruzione in qualsiasi riproduzione elettronica convenzionale. E la stessa esperienza visiva si memorizza in echi che ricorrono per qualche ora dopo aver visto tali immagini. Il mondo reale diventa "HDR", rimanda a quanto appena visto, qualora si incontrino con lo sguardo specularità, riflessi, dettagli contestualmente apprezzabili in una scena che contenga contrasti molto ampi. C'è inoltre un altro effetto collaterale di un'esperienza di visione in HDR: l'apparente tridimensionalità delle immagini. La presenza di contrasti così prossimi all'esperienza reale (la visione fisiologica indica in circa 13-14 stop la dinamica in fase di adattamento avvenuto con un contrasto di circa 6000:1, sia essa scotopica [al buio], fotopica [in piena luce diurna], o mesopica [tipicamente in interni]) inganna in qualche modo il cervello che ricostruisce piani spaziali differenti (ed illusori) basandosi sull'intensità relativa dei riflessi (quelli più intensi sono riconducibili ad oggetti più vicini, quelli più deboli ad oggetti più lontani).



[Immagine 4]

[Immagine 3] Uno dei monitor HDR, presentati al recente NAB di Las Vegas. Si tratta di un monitor Reference Canon 4K- HDR da 30" (tecnologia LCD IPS retroilluminato LED con microzone RGB), in grado di generare luminosità di picco di oltre 2000 cd/m². Le sue dimensioni laterali erano altrettanto impressionanti.

[Immagine 4] E, una volta osservata una sequenza televisiva in HDR, si percepisce immediatamente il suo valore; soprattutto, una volta posti a confronto due monitor HDR identici (oppure un monitor che supporti il dual view, in altre parole con lo schermo diviso in due parti identiche in grado di mostrare impostazioni differenti per la stessa sequenza).

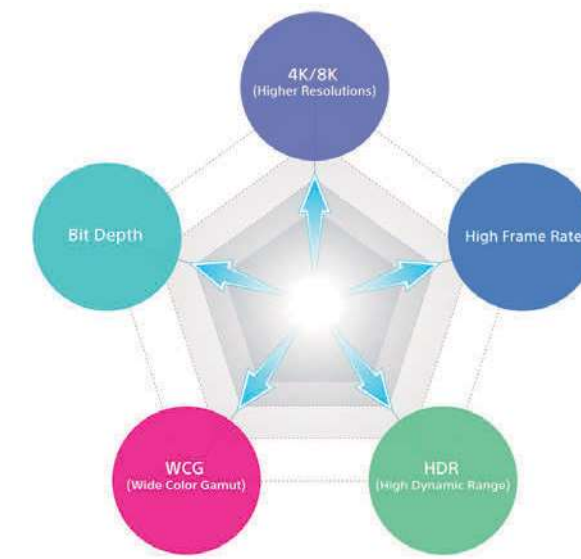
Codifica di segnali televisivi in High Dynamic Range

Ma HDR non significa soltanto, e semplicemente, immagini più luminose. Un display di questo tipo riproduce quasi tutto il range dinamico di una camera professionale con gli stessi contrasti presenti nella scena reale. Se da un lato ciò è interessante per le implicazioni qualitative, allo stesso modo impone una rinnovata attenzione alla composizione filmica della scena.

Laddove alcuni elementi indesiderati potevano prima essere nascosti nelle ombre o diluiti nelle alte luci, il range dinamico riprodotto - ora senza compressioni - li evidenzierà in maniera netta.

La disponibilità di tecnologie, in grado di realizzare monitor con range dinamico più esteso, ha indotto diversi ricercatori e gruppi di studio a condurre rilievi statistici sulle preferenze degli spettatori, al fine di identificare i requisiti minimi di una televisione ad ampio range dinamico. Nella ricerca è innanzitutto emerso che gli schemi di codifica digitale debbano tenere conto delle curve caratteristiche di sensibilità al contrasto dell'occhio umano. E' inoltre necessario considerare il comportamento adattativo dell'occhio al variare del livello d'illuminazione medio, in particolare nel passaggio dalla luce al buio (che richiede tempi variabili da 8 a 30 minuti) e nel passaggio inverso (molto più rapido: in genere valutabile nell'ordine dei secondi).

Il comportamento medio degli spettatori è stato studiato con display sperimentali in grado di supportare un range dinamico da 0.004 cd/m² a 20,000 cd/m², cioè un contrasto di 50000:1 e caratterizzato dal colour gamut DCI P3 impiegato in cinematografia elettronica. Gli esperimenti, condotti con immagini reali e in computer graphics, dimostrano che il 50% del campione trova gradevoli le riproduzioni comprese in un range da 0.1 a 650 cd/m²; per incontrare il 90% dello stesso campione, occorre estendere il range riprodotto nell'intervallo da 0.005 cd/m² a >3000 cd/m² [Immagine 2]. Questo incremento è proprio quello necessario per riprodurre più fedel-



La risoluzione, è probabilmente il fattore meno rilevante tra quelli che definiscono la qualità di un'immagine. All'aumentare del range dinamico dell'immagine, gli aspetti qualitativi immediatamente prioritari diventano il gamut cromatico e la risoluzione temporale.

mente le specularità, i riflessi e i dettagli nelle ombre. La televisione HDR è interamente contenuta in quest'ultima affermazione e la sua implementazione si risolve nella codifica ottima di un segnale digitale capace di visualizzare tali estremi.

In SDR i contenuti supportati hanno luminosità variabili tra 0.1 e 100 cd/m². Conseguentemente occorre riferirsi alla curva sensitometrica dell'occhio in questo range di livelli. La curva descrittiva adeguata è la Vries-Rose, che definisce la percezione psicofisica della luminosità in modo proporzionale alla radice quadrata dell'intensità energetica effettiva.

Tale curva (con la sola eccezione della parte iniziale) è quella indicata nella raccomandazione ITU BT.709 come "gamma" o "correzione gamma" ed è la relazione che mappa la luce lineare di una scena in una scala nella quale, ad ogni incremento, corrisponde un cambiamento percettivamente costante (il raddoppio dell'intensità luminosa corrisponde al raddoppio della sensazione percettiva). Questa funzione è conosciuta come Optical to Electrical Transfer Function (OETF).

Nella televisione analogica essa ha permesso di mantenere costante la percezione del rumore elettrico rispetto a tutti i livelli d'illuminazione riprodotta. In digitale, in modo utile, ha permesso di minimizzare il numero di bit richiesti per ogni campione rispetto a una riproduzione di tipo lineare. Si è infatti dimostrato che 8-10 bit per campione sono sufficienti per lo scopo senza introdurre effetti evidenti di posterizzazione nelle luci elevate o di rumorosità alle basse luci. Spesso i display SDR consumer hanno un processo interno a soli 8 bit. Nel caso un segnale sia codificato con 10bit/campione, gli ultimi due sono scartati. Tale semplificazione è possibile perché in BT.709 i due bit meno significativi sono responsabili solo del livello minimo del nero (al di sopra del rumore) e non alterano in modo significativo il livello della massima luminosità riproducibile, sebbene sia proprio il suo aumento ad essere auspicabilmente necessario.

Ma HDR non significa soltanto, e semplicemente immagini più luminose. Un display di questo tipo riproduce quasi tutto il range dinamico di una camera professionale con gli stessi contrasti presenti nella scena reale.

[Immagine 6]

La curva sensitometrica di Barten descrive la linea al di sopra della quale la maggior parte delle persone comincia a vedere gli effetti di step discreti di luminanza e che si manifestano con banding sullo schermo. E questo modello è stato quello impiegato dalla Dolby e definisce i livelli digitali in termini di luminosità assoluta sul display e non in termini di luminanza assoluta in fase di acquisizione

[Immagine 7]

La curva di Weber, descrive la sensibilità dell'occhio umano quando osserva particolari dissolti in un background uniformemente illuminato. Tra i due estremi esibisce pendenze differenti e nel range intermedio si ricolloca all'interno della curva specificata nella raccomandazione BT.1886.

[Immagine 8]

La curva PQ permette oltre 24 stop di dinamica, quantizzazione a 10/12 bit e livelli di luminosità minimi e massimi associati ai codici digitali rispettivamente di 0.005 cd/m² e 10.000 cd/m². Si osservi che solo il 7% dei codici digitali definisce il range da 5.000 a 10.000 cd/m²

[Immagine 9]

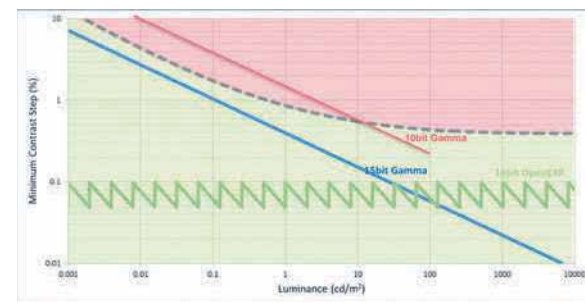
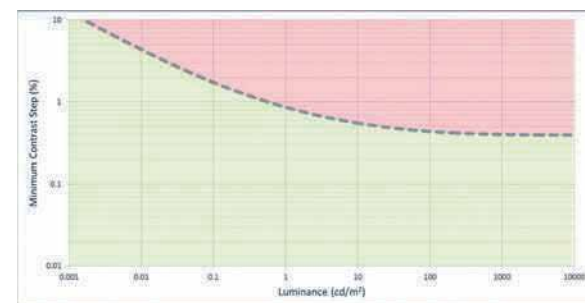
Gli studi condotti rivelano che 10 bit a campione, sono sufficienti ad evitare effetti importanti di banding. Confrontando il risultato rispetto alla curva BT.1886 scalata fino a 1000 cd/m², la rappresentazione in PQ delle aree scure richiede meno bit a campione della BT.1886, confermando il problema di uso inefficiente di bit/campione della curva SDR.

Anche la più recente raccomandazione BT.2020, pur definendo la televisione Ultra High Definition (UHDTV) non altera la curva OETF, e mantiene il contatto con la rappresentazione delle immagini in SDR non alterando la OETF della BT.709. Espande comunque lo spazio colore fino a contenere al 99.9% il luogo di Pointer (curva chiusa che include le tinte di tutte le superfici ed oggetti naturali) e definisce una migliore risoluzione spaziale e temporale (apre al 4K e 8K, aumenta il frame rate fino a 100 fps).

Si noti che, fin quando i CRT erano gli unici o i più diffusi dispositivi di visualizzazione, non c'è mai stato bisogno di specificare una curva inversa all'OETF, cioè una Electro-Optical Transfer Function (EOTF), in quanto l'inversione della correzione gamma era contenuta implicitamente nella stessa tecnologia del CRT. Come detta da Charles Poynton: una fortuita coincidenza! La risultante funzione di trasferimento Optical to Optical (OOTF, anche nota come Gamma end-to-end) consentiva in questo modo una sostanziale identità percettiva tra scena reale ripresa e scena elettronica riprodotta (fatte salve le limitazioni già enunciate, la considerazione delle condizioni di visione medie e gli intenti artistici).

E' solo nel 2011, dunque quando il mercato dei CRT è stato del tutto annullato dal sopraggiungere delle nuove tecnologie, che ITU ha standardizzato un'EOTF, ripresentando in essa la caratteristica tipica di un CRT.

Nonostante il clamoroso paradosso, si evidenzia il nodo centrale di questo tentativo. La base fondamentale per interpretare un segnale in digitale risiede nella descrizione del metodo per convertirlo in energia luminosa su un dispositivo trasduttore. E' dunque la EOTF e non la OETF che definisce interamente lo scopo di una codifica digitale. La EOTF quando applicata ad un segnale con campioni ad 8,10,12 bit dovrebbe idealmente evitare la visione delle discontinuità di riproduzione tonale tra livelli digitali adiacenti. Pertanto EOTF e profondità di quantizzazione dovrebbero armonizzarsi il più possibile con la curva tipica di sensibilità al contrasto dell'occhio umano.



[Immagine 6]

Perceptual Quantiser PQ

Quanto ricercato è condensato nella curva sensitometrica di Barten [Immagine 6]. Essa descrive la linea al di sopra della quale la maggior parte delle persone comincia a vedere gli effetti di step discreti di luminanza e che si manifestano con banding sullo schermo. E questo modello è stato quello impiegato dalla Dolby per computare un'EOTF ottimizzata dal punto di vista percettivo. Essa definisce dunque i livelli digitali in termini di luminosità assoluta sul display e non in termini di luminanza assoluta in fase di acquisizione (seguendo dunque il campione statistico sul gradimento descritto in precedenza).

Tale curva, definita in modo appropriato "Quantizzatore Percettivo" (Perceptual Quantiser: PQ nel seguito) ha un comportamento che ben si adatta alla curva Vrien-Rose per i livelli più bassi d'intensità luminosa e mostra una compressione con pendenza praticamente nulla per le intensità più elevate, in modo consistente con il comportamento logaritmico della curva di Weber [Immagine 7]. Quest'ultima descrive la sensibilità dell'occhio umano quando osserva particolari dissolti in un background uniformemente illuminato.

Tra i due estremi esibisce pendenze differenti e nel range intermedio si ricolloca all'interno della curva specificata nella raccomandazione BT.1886 (in grado di invertire la codifica che descrive efficacemente tutte le superfici opache). Complessivamente, la curva PQ permette oltre 24 stop di dinamica, quantizzazione a 10/12 bit e livelli di luminosità minimi e massimi associati ai codici digitali rispettivamente di 0.005 cd/m² e 10.000 cd/m². Si osservi che solo il 7% dei codici digitali definisce il range da 5.000 a 10.000 cd/m² [Immagine 8].

Gli studi condotti rivelano inoltre che 10 bit a campione sono sufficienti ad evitare effetti importanti di banding. Confrontando il risultato rispetto alla curva BT.1886 scalata fino a 1000 cd/m², la rappresentazione in PQ delle aree scure richiede meno bit a campione della BT.1886, confermando il problema di uso inefficiente di bit/campione della curva SDR (una volta accomodata per ospitare un range dinamico più esteso) [Immagine 9]. A parità di risultati occorrerebbe un numero di bit >12, rendendo difficile l'implementazione tecnologica di ogni aspetto della catena produttiva [Immagine 10-11].

Il metodo proposto da Dolby Vision è dunque duale di quello convenzionale. In termini di livelli assoluti di luminosità, non si ammette più la relatività tra scena reale e riprodotta, bensì si legano i livelli d'intensità luminosa assoluti della scena reale con quelli riprodotti dal monitor. Il particolare monitor di visione impiegherà di riprodurre solo il range dinamico di cui è capace. Tutte le parole codice associate a livelli di luminosità al di fuori di quell'intervallo saranno semplicemente ignorate. Monitor di riferimento permettono invece di utilizzare il loro range dinamico per rimappare le parole codice in un "HDR Range" selezionabile dall'utente.

Evidentemente, se un'EOTF è fissata, sarà la OETF ad essere dipendente dal dispositivo di acquisizione (la camera). In questo caso si è guardato a quanto è consuetudine nel cinema, dove l'impiego di curve OETF di tipo logaritmico è realtà da molti anni. Si pensi, ad esempio, alla curva caratteristica dei sistemi di scansione di negativi

analogici "Cineon" introdotta da Kodak all'inizio degli anni '90, molto simile alla caratteristica dell'occhio umano in una fase stabile di adattamento. Essa permette una codifica notevolmente più efficiente di quanto stabilito nella raccomandazione BT.709. Il formato Cineon consente di mappare 14 stop di luce lineare, cioè un range dinamico di circa 16.000:1, in un segnale digitale quantizzato a 10 bit con eccellenti risultati qualitativi.

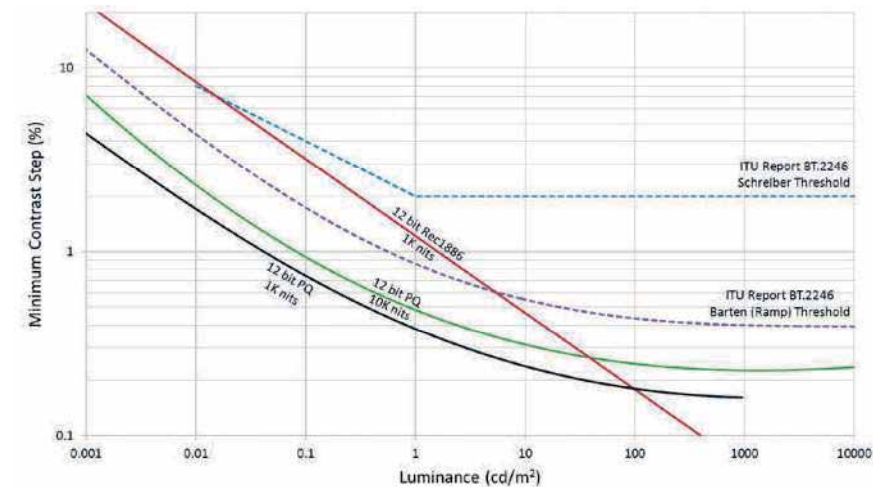
Pertanto, un sistema di produzione in HDR PQ sarà costituito da una camera che impieghi una curva di "correzione gamma" proprietaria a profilo logaritmico (Canon Log, Sony Log, Arri Log etc.), necessaria a sfruttare l'intero range dinamico del dato sensore, seguita da una LUT di conversione al PQ⁻¹ [Immagine 12].

Il risultato di quest'approccio è condensato in uno standard SMPTE: ST 2084 (contenuto a sua volta nel ITU BT.2100) che istituisce una vera e propria discontinuità con il passato e, in particolare, necessita l'uso di metadata per una corretta visualizzazione.

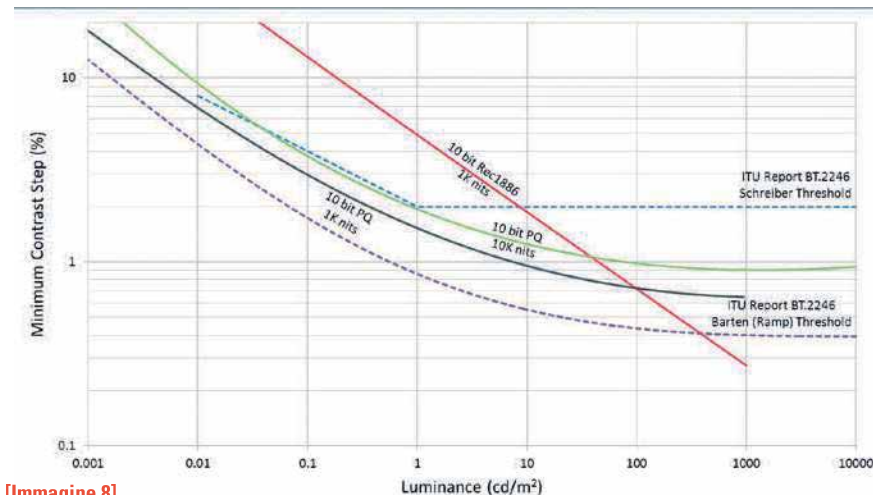
Nella raccomandazione ST.2086 si definiscono i metadata statici in grado di descrivere l'andamento della luminosità media e di picco del contenuto HDR in funzione delle capacità del monitor per la visualizzazione (in pratica danno indicazioni sul modo in cui il contenuto HDR è stato masterizzato). Essi includono le coordinate CIE (x,y) dei primari RGB, il punto di bianco e la luminanza massima e minima del display. Nello stream digitale del contenuto è allegato il MaxFALL (Maximum Frame/Average Light Level) che corrisponde al frame con la luminosità media più elevata nell'intero flusso dati (espresso in cd/m²). Un ulteriore metadata allegato è il MaxCLL (Maximum Content Light Level) che corrisponde al pixel più luminoso dell'intero flusso dati (anche quest'ultimo è rappresentato in cd/m²). I metadata MaxFALL/MaxCLL sono calcolati solo per la parte attiva dell'immagine. I valori per i neri non sono invece processati. Quando il display riceve i metadata regola la EOTF per realizzare una OOTF ottima che consente di tenere conto dei diversi display usati per masterizzazione e visualizzazione, nonché le diverse condizioni di visione.

Ciò permette di conservare al meglio gli intenti artistici del contenuto.

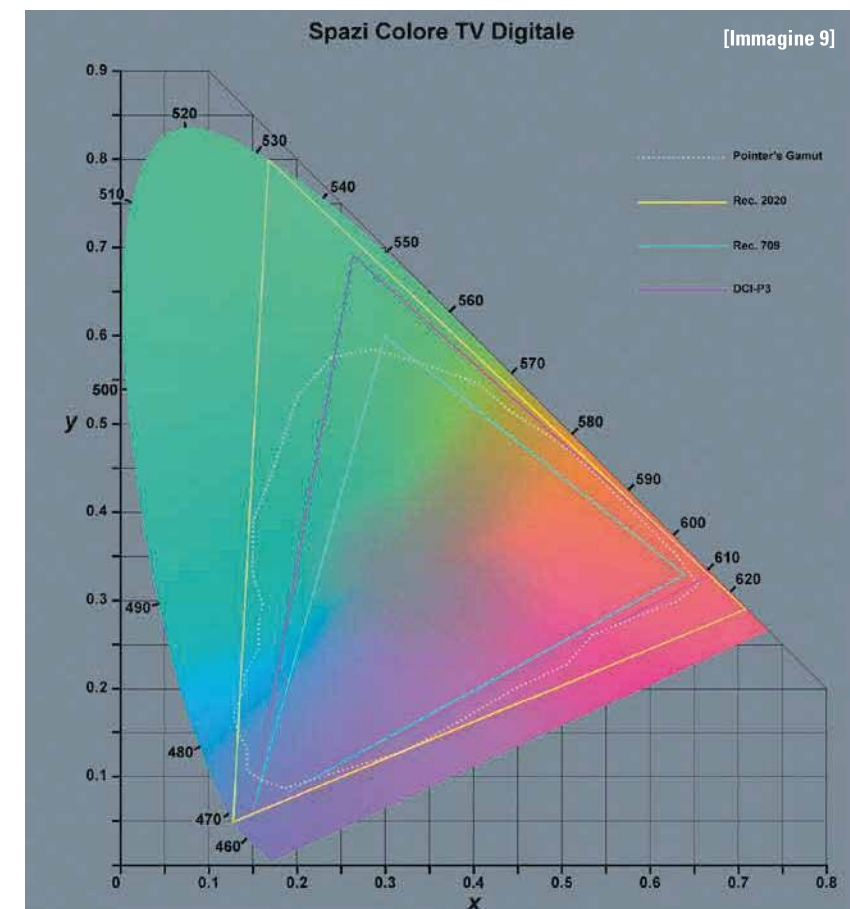
Quando il display riceve i metadata regola la EOTF per realizzare una OOTF ottima che consente di tenere conto dei diversi display usati per masterizzazione e visualizzazione, nonché le diverse condizioni di visione. Ciò permette di conservare al meglio gli intenti artistici del contenuto.



[Immagine 7]



[Immagine 8]



[Immagine 9]

Hybrid Log Gamma

Il metodo Dolby Vision non ha trovato consenso unanime; BBC e NHK hanno commentato infatti che la televisione ha sempre lavorato con livelli di luminanza relativi tra scena reale e rappresentazione su schermo. Legare la luminanza reale della scena con la sua rappresentazione su schermo porta inevitabilmente a una rivoluzione completa dei modelli produttivi, distributivi e percettivi: primo tra tutti l'assenza di qualunque retro-compatibilità!

Non è inoltre strettamente necessario disporre simultaneamente di un range dinamico di 10^7 (oltre 24 stop!), considerando che l'occhio umano al più arriva a 10^4 (13-14 Stop) in una fase stabile di adattamento. Da tale critica nasce l'approccio definito Hybrid Log Gamma: esso preserva il metodo classico "display-independent", fissando un'OETF che sia parallela a quella stabilita nella raccomandazione BT.709 fino al codice corrispondente al 50% dell'intensità massima (messo in corrispondenza del bianco di riferimento a 100 cd/m^2), e poi prosegue estendendosi fino al codice relativo al 100% (messo in relazione con un valore limite di 1200 cd/m^2).

Anche in questo caso, la curva HLG impiega un segmento conforme alla curva di Weber per la parte di curva dal 50% in poi. Nella parte inferiore non è invece presente il primo tratto lineare contenuto nella definizione del "gamma" BT.709 (introdotto originariamente per limitare la visibilità del rumore per bassi livelli di luminosità). Tale accorgimento non è più richiesto nelle camere HDR, le quali si ritengono sicuramente meno rumorose.

A differenza del PQ, HLG non richiede l'impiego di metadata dato che l'EOTF risultante sarebbe compatibile automaticamente con monitor SDR e HDR (solo quelli certificati sarebbero in grado di riprodurre la seconda parte della OETF HLG). Essendo inoltre indipendente dal display (che, appunto, offre una rappresentazione "relativa" rispetto alla realtà della scena fisica), si adatta a dispositivi con diverse luminosità massime in un ampio range di ambienti di visione. I risultati delle sperimentazioni effettuate sono sicuramente incoraggianti (sebbene non estremi come nel caso del PQ). Quando un segnale HLG è visualizzato su un monitor a 10 bit reali con 2000 cd/m^2 di luminosità massima si riesce a visualizzare un range dinamico di $2 \cdot 10^5 / 17.6$ stop. In HLG sarà la curva inversa EOTF a dover essere definita dal dato display e, coerentemente, modifiche end-to-end sono possibili (in modo simile a quanto avviene correntemente) per ottimizzare il risultato sulla base delle condizioni di visione e intenti artistici.

Entrambi i metodi sono, di fatto, standardizzati dalla raccomandazione ITU BT.2100.

I display attuali con certificazione HDR possono emettere un massimo variabile tra 400 cd/m^2 e 4.000 cd/m^2 e le specifiche per il Dolby Vision indicano che un'immagine HDR può contenere elementi con luminosità massima fino a 10.000 cd/m^2 (al momento la maggior parte dei contenuti in HDR sono masterizzati per evidenziare un picco di luminanza tra i 1000 cd/m^2 e i 4000 cd/m^2).

La profondità delle ombre è altrettanto importante dell'intensità delle luci più intense. Laddove un monitor SDR può riprodurre neri minimi a un livello di $1,2 \text{ cd/m}^2$, il Canon DP-V2420, ad esempio, arriva a riprodurre un nero minimo di 0.025 cd/m^2 . Si noti che i 6 stop di range intermedio, che vanno dal bianco opaco al nero opaco avranno esattamente la stessa resa di un monitor SDR.

La differenza è che in HDR si va al di fuori di questo intervallo di riproduzione tonale, con gradazioni di bianco e nero che sono più luminose e più scure di qualsiasi cosa che si possa apprezzare in SDR. E poiché non si devono costringere 14-15 stop di dinamica in un contenitore dove è possibile visualizzarne solo 6, ogni stop è visualizzato approssimativamente con lo stesso contrasto. Il risultato è un incremento dei dettagli visualizzati in questi estremi, analogamente a una vera e propria esplosione del "volume" dell'immagine e della sua cromaticità.

Spazio Colore in High Dynamic Range

Il sistema televisivo attuale non supporta uno spazio molto ampio di colori riproducibili. Sicuramente molto ridotto rispetto a quello percepibile dall'occhio umano. Una televisione di tipo HDR non può prescindere da quest'aspetto poiché il suo scopo è di avvicinare l'esperienza dello spettatore alla visione reale. Ciò che manca nella televisione SDR è il volume della sensazione cromatica. Saturazione e brillantezza non riescono a trovare una coniugazione con le limitazioni di sole 100 cd/m^2 e spazio colore come definito dalla raccomandazione BT.709.

Lo spazio colore BT.709 - definito dal triangolo con vertici nei tre primari R,G,B e inscritto nello spazio colore CIE 1931 [Immagine] - è valido solo per piccoli valori di luminanza. Al crescere dell'intensità della luce, il range di colori riproducibili si restringe in modo coerente con il fatto che all'intensità relativa di 1.0 lo spazio colore si riduce ad un singolo punto: il bianco D65 definito dallo standard. Lo spazio colore parametrato rispetto alla luminanza è dunque un solido tridimensionale. Assumendo di disegnare tale solido in un sistema di assi coordinati aventi coordinate CIE X,Y nel piano e la luminanza sull'asse Z, ogni sezione ottenuta ad un certo valore di intensità luminosa rappresenterà l'insieme delle tinte riproducibili in quella condizione (il vantaggio della CIE XYZ è che le tinte sono definite in termini di radianza spettrale e tutti i colori visibili possono essere rappresentati usando solo valori positivi delle tre componenti XYZ).

La migliore descrizione per quest'ultima è attraverso una terna di valori in virgola mobile, tipicamente con parole di 32 bit, ma ciò renderebbe molto complessa una codifica digitale. Si passa dunque per una conversione da valori in virgola mobile a interi, tenendo conto delle limitazioni dell'occhio umano e del fatto che esso può percepire solo un numero limitato di livelli di luminanza e di tinte [Immagine]. Il triangolo comunemente mostrato è relativo ad un valore proporzionale di luminanza di 0.0722, dal quale punto in poi il primario del canale del blue non può più aggiungere un contributo

al valore della luminanza complessiva (la spiegazione di questo aspetto risiede nel fatto che, con il sistema additivo RGB, ad un certo punto occorrerebbe un contributo negativo di uno o più primari, il che elettricamente non è possibile). Al fine di incrementare la luminanza è allora necessario de-saturarlo per aggiungere il rosso o il verde o una miscela dei due. Dal punto 0.00722 in poi la sezione non è più triangolare ma quadrilatera. Alla luminanza relativa di 0.2126 il rosso, in modo simile, non può più aggiungere un contributo netto alla luminanza dell'immagine e quindi è necessario aggiungere blue e verde o una miscela dei due. La sezione diventa pentagonale, ma fino al valore relativo di $0.0722 + 0.2126 = 0.2848$, dopo il quale la luminanza può crescere solo aggiungendo il verde; la sezione ritorna ad essere quadrilatera.

Alla luminanza relativa di 0.7152 si ritorna a una sezione pentagonale, poi ancora quadrilatera, e infine triangolare che tende a convergere in un punto (bianco D65) quando la luminanza approssima il valore relativo di 1,0. Considerando inoltre che il metodo di codifica di un segnale televisivo è quello a componenti Y'Cb'Cr', detto "a luminosità non costante", si intuisce come molte delle tinte naturali non possano essere contenute in questo schema. La rappresentazione di quanto acquisito dalla telecamera è rimappata in tinte che non sono quelle reali né per tonalità, né per brillantezza e saturazione. La raccomandazione ITU BT.2020 e la BT.2100 stabiliscono allo scopo un nuovo spazio colore con primari monocromatici che allargano notevolmente lo spazio dei colori riproducibili.

Al momento, non esiste alcun monitor o proiettore in grado di coprire interamente tale spazio (e analogo discorso potrebbe essere fatto per lo spazio colore ACES, non trattato in questa sede per brevità). Le percentuali migliori (monitor Canon, Sony, Dolby) sono in un range dal 75% all'86%. Per la prima volta, gli enti ratificatori hanno progettato uno standard "future-proof" la cui completa definizione è possibilmente realizzabile in toto con dispositivi di generazione futura.

Analogamente, resta da risolvere il problema dell'attuale codifica Y'Cb'Cr' a luminanza non costante. I principi della scienza del colore indicano che combinando linearmente i primari RGB si ottiene la vera luminanza della scena ripresa, denotata con Y. Se un sistema video operasse in questo modo, sarebbe aderente al principio della "luminanza costante". Per ragioni di semplicità circuitale, invece, si realizza una combinazione lineare di segnali R' G' B' già sottoposti al processo del Gamma (OETF) [Immagine]. In questo modo si ottiene una quantità differente, Y' definita Luma, che è responsabile di alcuni artefatti, tra cui quello di generare tinte con brillantezza non costante, in quanto occorrerebbe decodificare la luminanza da tutte e tre le componenti Y' Cb' Cr' e non solo da Y' (una parte della vera luminanza associata si trova contenuta nei segnali di cromaticità per effetto del processo video descritto).

Un nuovo metodo di codifica cromatica, inserito nello standard BT.2100, potrebbe risolvere definitivamente questo problema. ICtCp è stato progettato tenendo conto del sistema di visione umana; descrive dunque le differenze appena visibili e percettivamente uniformi nello spazio colore.

A differenza del volume CIE XYZ non è normalizzato alla luminanza di picco 1.0 o ad una determinata fase di adattamento visivo. E' invece una rappresentazione assoluta dei colori, utilizzabile per descrivere il numero di colori visivamente distinguibili rappresentabili da un dispositivo trasduttore.

La rappresentazione in coordinate ICtCp permette di comparare rapidamente i "volumi" cromatici tra dispositivi diversi. Dopo opportune trasformazioni matematiche si ha in particolare che:

- La luce incidente è rappresentata analogamente a come è catturata dall'occhio umano: mediante tre diversi tipi di recettori (coni) con sensibilità di picco sintonizzata sulle lunghezze d'onda lunghe (L), medie (M) e corte (S).
- La luce lineare è convertita non linearmente in un segnale composto secondo la risposta fisiologica dell'occhio stesso.
- Il segnale non lineare finisce in un sistema di differenziazione dei colori che lo separa in tre diversi percorsi: il percorso chiaro-scuro, il percorso giallo-blu isoluminoso "Tritan", il percorso verde rosso isoluminoso "Protan".

Si approfondirà l'argomento in una prossima pubblicazione.

Il sistema televisivo attuale non supporta uno spazio molto ampio di colori riproducibili. Sicuramente molto ridotto rispetto a quello percepibile dall'occhio umano. Una televisione di tipo HDR non può prescindere da quest'aspetto poiché il suo scopo è di avvicinare l'esperienza dello spettatore alla visione reale. Ciò che manca nella televisione SDR è il volume della sensazione cromatica.