

HDR & Dolby Vision-Mastering Prozess mit dynamischen Metadaten

Bachelorarbeit

im Studiengang
Audiovisuelle Medien

vorgelegt von

Daniel Jordan

Matr.-Nr.: 32453

am 25. Februar 2020

an der Hochschule der Medien Stuttgart

Erstprüfer/in: Prof. Dr. Jan Fröhlich

Zweitprüfer/in: Prof. Stefan Grandinetti

Ehrenwörtliche Erklärung

„Hiermit versichere ich, Daniel Jordan, ehrenwörtlich, dass ich die vorliegende Bachelorarbeit (bzw. Masterarbeit) mit dem Titel: „HDR & Dolby Vision-Mastering Prozess mit dynamischen Metadaten“ selbstständig und ohne fremde Hilfe verfasst und keine anderen als die angegebenen Hilfsmittel benutzt habe. Die Stellen der Arbeit, die dem Wortlaut oder dem Sinn nach anderen Werken entnommen wurden, sind in jedem Fall unter Angabe der Quelle kenntlich gemacht. Die Arbeit ist noch nicht veröffentlicht oder in anderer Form als Prüfungsleistung vorgelegt worden.

Ich habe die Bedeutung der ehrenwörtlichen Versicherung und die prüfungsrechtlichen Folgen (§26 Abs. 2 Bachelor-SPO (6 Semester), § 24 Abs. 2 Bachelor-SPO (7 Semester), § 23 Abs. 2 Master-SPO (3 Semester) bzw. § 19 Abs. 2 Master-SPO (4 Semester und berufsbegleitend) der HdM) einer unrichtigen oder unvollständigen ehrenwörtlichen Versicherung zur Kenntnis genommen.“

A handwritten signature in black ink, appearing to read 'Daniel Jordan', with a stylized, cursive script.

Stuttgart, 25.02.2020, Daniel Jordan

Kurzfassung

In dieser Arbeit werden die Entwicklungen von High Dynamic Range (HDR) sowie die technischen Merkmale wie Dynamikumfang, Farbraum, Bittiefe und Transferfunktionen erläutert. Gerade in Hinblick auf den Erhalt des Creative Intent und infolge dessen die Entwicklung von Dynamischen Metadaten und ihrer Bedeutsamkeit. Darüber wird der Mastering Prozess von Dolby Vision HDR in Davinci Resolve, die technischen Grundlagen für den HDR-Grading Prozess sowie die Besonderheiten im Dolby Vision Mastering im Hinblick auf SDR-Trims beschrieben.

Schlagwörter: High Dynamic Range, HDR, Dolby Vision, Davinci Resolve, Dynamische Metadaten, Wide Color Gamut, EOTF, Mastering, Color Grading

Abstract

This work describes the development of High Dynamic Range and its technical characteristics like Dynamic Range, Color Gamut, Bit Depth and Electro Optical Transfer Functions. Respectively on dynamic metadata and why they are relevant for the preservation of the creative intent. In addition the process of Dolby Vision Mastering in Davinci Resolve and the technical aspects of it – especially SDR-Trims – will be described.

Keywords: High Dynamic Range, HDR, Dolby Vision, Davinci Resolve, Dynamic Metadata, Wide Color Gamut, EOTF, Mastering, Color Grading

Inhaltsverzeichnis

Ehrenwörtliche Erklärung	1
Kurzfassung	2
Abstract	2
Inhaltsverzeichnis	3
Abbildungsverzeichnis	5
Tabellenverzeichnis	8
Abkürzungsverzeichnis	9
Vorwort	10
1 Die menschliche Wahrnehmung	11
1.1 Dynamikwahrnehmung.....	12
1.2 Farbwahrnehmung.....	13
2 HDR – High Dynamic Range	14
2.1 Was ist High Dynamic Range?	14
2.2 Marktübersicht HDR-TVs.....	16
2.3 HDR Produktionen und Streaming Anbieter	17
2.4 Übersicht HDR-Standards.....	18
3 HDR10 & Statische Metadaten	19
3.1 Wide Color Gamut und Volume Mapping.....	19
3.2 EOTF, OETF	23
3.3 Metadaten.....	30
3.4 Probleme	32
4 Dolby Vision & Dynamische Metadaten	35
4.1 Unterschiede zwischen HDR10 und Dolby Vision.....	35
4.2 Dynamische Metadaten.....	38
4.3 Trim Controls	40
4.4 Sehumgebung.....	41
5 Mastering in Dolby Vision	44
5.1 Hardware und Referenz-Monitore	44
5.1.1 HDR Studio Referenz Monitore	44
5.1.2 Consumer HDR-TVs	45
5.1.3 SDR-Referenzmonitor	45

5.1.4	Wiedergabekarten	45
5.1.5	Dolby Content Mapping Unit (CMU) und eCMU	46
5.1.6	Dolby Vision Lizenz.....	46
5.2	HDR-Grading in Resolve	47
5.2.1	HDR Projekt anlegen.....	47
5.2.2	Anpassung des Helligkeitsbereichs / HDR Mode	50
5.2.3	Grading und MaxFALL.....	50
5.2.4	HDR Schatten und Lichter.....	51
5.2.5	Clipping	52
5.3	SDR-Trim Pass.....	54
5.3.1	L1-Metadaten-Analyse	54
5.3.2	L3 Mid Tone Offset	56
5.3.3	Primary Trim Controls.....	56
5.3.4	Secondary Trim Controls.....	60
5.4	Export.....	62
5.4.1	Dolby Vision Master.....	62
5.4.2	Dolby Vision XML.....	63
5.4.3	Dolby Vision IMF.....	65
5.4.4	SDR-Trim Version Export.....	67
6	Schlusswort	68
7	Literatur	69

Abbildungsverzeichnis

Abbildung 1-1: Das Farbspektrum und die Empfindlichkeit von Zapfen (cones) und Stäbchen (rods) [1].....	11
Abbildung 1-2: Stäbchen und Zapfen [1]	11
Abbildung 1-3: Die menschliche Wahrnehmung im Vergleich mit LDR und HDR [2].....	12
Abbildung 1-4: CIExy 1931 Farbtafel mit sRGB-Farbraum	13
Abbildung 2-1: Belichtungsreihe eines HDR-Bracketing und das fertige kombinierte Bild. Überbelichtetes, unterbelichtetes und kombiniertes Bild (von links) [4].....	14
Abbildung 2-2: Vergleich von dem Verlust von Dynamic Range bei SDR und HDR Produktionen. Der menschliche Dynamikumfang (oben) im Vergleich zu SDR (Mitte) und HDR (unten) [5].....	15
Abbildung 2-3: HDR TV Lieferungen nach Displaytechnologie [8, S. 33]	17
Abbildung 2-4: UHD on-demand Titel nach Abodiensten in den UK (2017-2018) [8, S. 36].....	17
Abbildung 2-5: Anzahl der UHD-Titel sowie deren Distribution [8, S. 35]	17
Abbildung 3-1: 2D- und 3D-Farbraum mit RGB-Primaries [12]	19
Abbildung 3-2: Pointer Gamut enthält die Farben von 4000 diffusen Oberflächen die in der Natur vorkommen [6]	20
Abbildung 3-3: Color Gamut Vergleich von REC 709, DCI P3 und REC 2020 [13, S. 6]	20
Abbildung 3-4: Vergleich von Farbverläufen zwischen 8-Bit und 10-Bit [18].....	22
Abbildung 3-5: Vergleich von Bittiefen-Verlaufswerten und deren Präzision [19]	22
Abbildung 3-6: OOTF und EOTF Workflow mit PQ Kurve [20]	23
Abbildung 3-7: 12 Bit Rec.1886 Gamma Kurven [22]	23
Abbildung 3-8: Contrast sensitivity function (CSF) des Barten-Modells [22].....	24
Abbildung 3-9: Darstellung des Barten Modells [24]	25
Abbildung 3-10: Versuchsaufbau von Dolby: Mit einem NEC Projektor (links) und dem finalen Aufbau mit einem Christie Projektor (rechts) [21, S. 564].....	25
Abbildung 3-11: Bevorzugte Helligkeitslevel geordnet nach Erfahrungsgruppen [21, S. 565].....	25
Abbildung 3-12: Ergebnisse der Versuchsreihe von Dolby für die Festlegung von Dynamikumfang von Displayeinheiten [21, S. 566].....	26
Abbildung 3-13: Referenz PQ EOTF Formel [20]	26
Abbildung 3-14: 10 Bit PQ and Rec.1886 im Vergleich [22]	27
Abbildung 3-15: 12 Bit PQ und Rec.1886 Gammakurven [22]	27
Abbildung 3-16: Darstellung von Log, OpenEXR und 15bit Gamma im Vergleich zum Barten Threshold [24]	28
Abbildung 3-17: PQ Code Words Verteilung [24].....	29
Abbildung 3-18: Gamma (Rec.1886) Code Words Verteilung [24]	29
Abbildung 3-19: Berechnung der RGB Primaries von CIE XYZ 1931 [26].....	30
Abbildung 3-20: Metadaten für P3D65 / 1000 Nits Mastering Display.....	30

Abbildung 3-21: Durch eine nicht ganz optimale Tonwertkorrektur kann Zeichnung in den Schatten verloren gehen (links), im Master dagegen sind diese korrekt belichtet (rechts) [29]	32
Abbildung 3-22: Durch Helligkeitsänderung kann ein Hue Shift entstehen (links) im Vergleich das Master (rechts) [29]	33
Abbildung 3-23: Banding kann durch eine falsche Tonemapping Kurve entstehen [29].....	33
Abbildung 3-24: Gerade in den Highlights kann durch Tone Mapping Clipping entstehen [29].....	34
Abbildung 4-1: HDR10 Baselayer in Dolby Vision [5]	35
Abbildung 4-2: Beziehung zwischen HVS und ICtCp [31]	37
Abbildung 4-3: Vergleich PQ basierte BT.2020 NCL Y'C'bC'r und ICtCp [31]	37
Abbildung 4-4: Vergleich der maximalen Farbtonfehler von ICtCp und PQ basierte BT.2020 NCL Y'C'bC'r, bei dem Datensatz von Hung und Berns [31]	38
Abbildung 4-5: Ergebnisse der bevorzugten Helligkeitseinstellung in Relation zum Umgebungslicht [36].....	41
Abbildung 4-6: Ergebnisse der Studie. Y-Axe: cd/m ² [37]	42
Abbildung 4-7: Simulierte Visualisierung von HDR ohne Dolby Vision IQ.....	43
Abbildung 4-8: Simulierte Visualisierung von HDR mit Dolby Vision IQ	43
Abbildung 5-1: Aktivierung der DeckLink Karte in Resolve unter Preferences > Video and Audio I/O > System.....	46
Abbildung 5-2: Einstellung der Color Science in DaVinci Resolve	47
Abbildung 5-3: Einstellung der Farbräume in DaVinci Resolve.....	48
Abbildung 5-4: Einstellung um die Scopes in Resolve in Nits darzustellen	48
Abbildung 5-5: HDR-Scopes in DaVinci Resolve 16 in Nits	49
Abbildung 5-6: Aktivieren der Metadaten über HDMI	49
Abbildung 5-7: Anpassung durch eine modifizierte Kurve.....	50
Abbildung 5-8: Festlegung der Weiß- und Schwarzwerte durch Low Range und High Range mit den Log Farbrädern	51
Abbildung 5-9: Soft-Clipping LUT erstellen in Resolve.....	52
Abbildung 5-10: Aktivieren der Sof-Clipping LUT in Resolve.....	53
Abbildung 5-11: Durch die Soft Clipping LUT ist ein graden über 1000 Nits nicht möglich. Außerdem besitzt die LUT ein weiches Clipping.....	53
Abbildung 5-12: Globale Trim Controls in Davinci Resolve [34]	54
Abbildung 5-13: Multi-Track Workflow [34]	55
Abbildung 5-14: Resolve Output Blanking Einstellungen [34].....	55
Abbildung 5-15: Korrektur (rechts) durch Mid Tone Offset.....	56
Abbildung 5-16: Dolby Vision Primary Trim Controls in Resolve.....	56
Abbildung 5-17: Beispiel für eine Korrektur (rechts) mit Lift / Gamma / Gain eines zu flachen Bildes (links)	57
Abbildung 5-18: Beispiel für eine Korrektur (rechts) mit Saturation Gain.....	57
Abbildung 5-19: Beispiel für eine Korrektur (rechts) mit Chroma Weight	58
Abbildung 5-20: Beispiel für eine Korrektur (rechts) mit Tone Detail	58
Abbildung 5-21: Beispiel für eine Korrektur (rechts) mit Mid Contrast Bias.....	59

Abbildung 5-22: Beispiel für eine Korrektur mit Highlight Clipping	59
Abbildung 5-23: Dolby Vision Secondary Trim Controls [34]	60
Abbildung 5-24: Beispiel für eine Blau Korrektur (rechts) mit Saturation Trim Controls	60
Abbildung 5-25: Korrektur (rechts) von zu saturierten Rot (links) nach L1 Analyse	61
Abbildung 5-26: Beispiel für eine Korrektur (rechts) von falschem Magenta Farbton (links) mit Hue Trim Controls	61
Abbildung 5-27: Dolby Vision XML export findet sich unter Dateityp	62
Abbildung 5-28: Dolby Vision XML Auszug zu Mastering Display	63
Abbildung 5-29: Dolby Vision XML Auszug zu Target Display	63
Abbildung 5-30: Dolby Vision XML Auszug zu Shot	64
Abbildung 5-31: Dolby Vision XML Wrapper. Details befinden sich innerhalb von Video	64
Abbildung 5-32: Video Export Einstellungen in Resolve	65
Abbildung 5-33: IMF-Ordner Inhalt nach Export aus Resolve	66
Abbildung 5-34: Tone Mapping Export in Davinci Resolve	67

Tabellenverzeichnis

Tabelle 2.4-1: Übersicht der gängigen HDR-Standards	18
Tabelle 4.1-1 Vergleich von HDR10 und Dolby Vision	36
Tabelle 4.2-1 Dolby Vision Metadaten Version 4.0.2 XML für CMv4 [34]	38
Tabelle 5.1-1: HDR Referenz Monitore im Überblick	45
Tabelle 5.4-1: Netflix Spezifikationen für UHD Produktionen [39]	65
Tabelle 5.4-2: Netflix Spezifikationen für HD Produktionen [39]	66

Abkürzungsverzeichnis

CIE	Commission Internationale de l'Éclairage. Internationale Beleuchtungskommission
CIE 1931 XYZ	Normfarbsystem das von der CIE 1931 entwickelt worden ist
DR	Dynamic Range, Dynamik Umfang
DV	Dolby Vision
EOTF	Electro-Optical Transfer Function
HDR	High Dynamic Range
HD	High Definition
HDMI	High-Definition Multimedia Interface
HLG	Hybrid Log Gamma
HVS	Human Visual System
ICtCp	HDR und WCG Farbdifferenz Modell
ITU	International Telecommunication Union
JND	Just Noticeable Difference, geringster wahrnehmbarer Unterschied
LMS	Farbraum der durch die Menschliche Wahrnehmung von langen (long), mittleren (medium) und kurzen (short) Wellenlängen entsteht
LUT	Lookup Table
MaxFALL	Maximum Frame Average Light Level
MaxCLL	Maximum Content Light Level
Nit	SI Einheit für Helligkeit, wird angegeben in cd/m ²
OETF	Opto-Electronic Transfer Function
OpenEXR	Fließkomma Bildformat
PQ	Perceptual Quantizer Transfer Funktion
Rec.2020	Farbraum für HDR-TV
Rec.709	Farbraum für HD-TV
RGB	Farbraum durch die Mischung von Roten, Grünen and Blauem Licht
SDR	Standard Dynamic Range (Gegensatz zu HDR)
SMPTE	Society of Motion Picture and Television Engineers
WCG	Wide Color Gamut
XML	Extensible Markup Language
YCBCR	Farbdifferenzmodell mit nicht linearer Helligkeit
Y'C'bC'r	Farbdifferentmodell mit linearer Helligkeit und Farbkanälen

Vorwort

Vor Ihnen liegt die Bachelorarbeit „HDR-Mastering Prozess mit dynamischen Metadaten“. Diese Bachelorarbeit habe ich als Abschlussarbeit meines Studiums der Audiovisuellen Medien an der Hochschule der Medien Stuttgart verfasst. Ziel war es, sich intensiv mit dem Thema High Dynamic Range auseinanderzusetzen und vor allem die technischen Aspekte zu beleuchten sowie eine Anleitung für den Mastering-Prozess von Dolby Vision HDR zu schreiben.

Mein Ansporn für diese Arbeit war vor allem die doch technische Neuheit von HDR. Ich wollte mich intensiv mit diesem Thema auseinandersetzen da ich schon immer Spaß daran hatte, technische Aspekte von künstlerischen Prozessen zu verstehen. Zudem wollte ich eine Arbeit schreiben die anderen Personen, die sich neu mit dem Thema beschäftigen, eine solide Grundlage bietet und die erste Hürde nimmt, eine doch immer noch neue Technik zu nutzen.

Herzlichen Dank an Prof. Dr. Jan Fröhlich für die großartige Unterstützung während dieser Arbeit, Michael Hackl für die Bereitstellung einer Dolby Lizenz für Davinci Resolve ohne die vor allem der praktische Teil dieser Arbeit nicht möglich gewesen wäre, sowie Marika und Hubert Grosse-Onnebrink für das Korrekturlesen meiner Arbeit.

Ich wünsche Ihnen viel Freude beim Lesen dieser Bachelorarbeit.

Daniel Jordan

Stuttgart, 25. Februar 2020

1 Die menschliche Wahrnehmung

Licht ist eine elektromagnetische Strahlung, die bestimmte Wellenlängen besitzt. Als Licht wird meistens nur das bezeichnet, was das menschliche Auge wahrnehmen kann. Dieses Spektrum liegt im Wellenlängenbereich von 380nm bis 750nm und reicht von violett-blauem Licht bis zu rotem Licht.

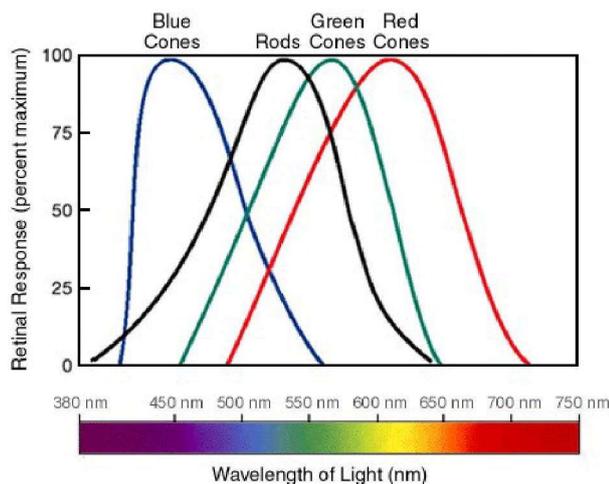


Abbildung 1-1: Das Farbspektrum und die Empfindlichkeit von Zapfen (cones) und Stäbchen (rods) [1]

Das menschliche Auge hat Stäbchen und Zapfen, die sich auf der Retina befinden. Diese reagieren auf verschiedene Wellenlängen unterschiedlich stark. Zapfen sind für die Farbwahrnehmung zuständig und die Stäbchen haben eine unbunte Empfindlichkeit. Sie sind vor allem für unseren Nachtsehsinn zuständig. Deswegen sind nachts auch alle Katzen grau, da die Stäbchen keine Farben wahrnehmen können, sondern nur auf Helligkeit reagieren.

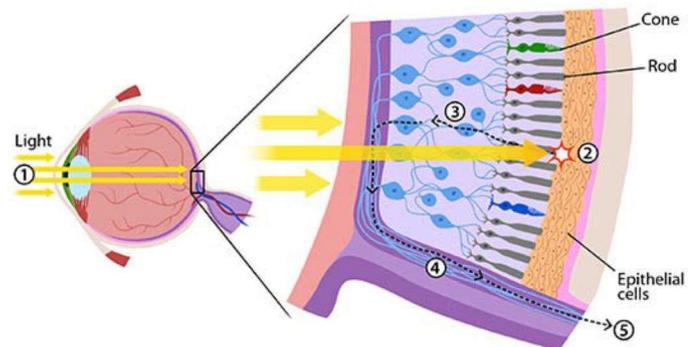


Abbildung 1-2: Stäbchen und Zapfen [1]

Die Zapfen sind wiederum in drei Gruppen unterteilt, die unterschiedlich empfindlich auf verschiedene Wellenlängen reagieren (s. Abb.1-1). Deshalb gibt es rote, blaue und grüne Zapfen.

1.1 Dynamikwahrnehmung

Die menschliche Wahrnehmung (Human Visual System, kurz HVS) war schon immer das Maßgebende, an der sich die Filmindustrie bezüglich Dynamikumfang und Farbe orientiert hat. Die Helligkeit wird in Nits oder cd/m^2 , einer Kerze auf einem Quadratmeter, angegeben. Durch die technischen Neuerungen, die mit HDR einhergekommen sind, ist jetzt eine noch bessere Annäherung an das HVS möglich. Das Auge passt bei einfallendem Licht die Größe der Pupille an und schränkt dadurch das eintreffende Licht ein. Somit wird der Dynamikumfang entsprechend geregelt. Denn das menschliche Auge kann auf einmal nur einen gewissen Dynamikumfang aufnehmen, dieser liegt ungefähr zwischen $0,1$ und $1000 \text{ cd}/\text{m}^2$ und passt sich je nach Umgebung mit einer Verschiebung nach Links oder Rechts an (s. Abb.1-1: Steady-State HVS). Die Stäbchen werden vor allem in dunklen Umgebungen aktiv.

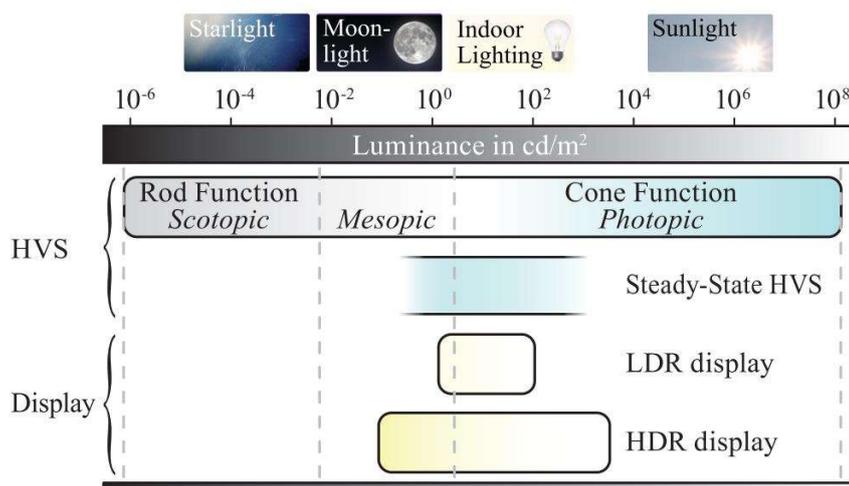


Abbildung 1-3: Die menschliche Wahrnehmung im Vergleich mit LDR und HDR [2]

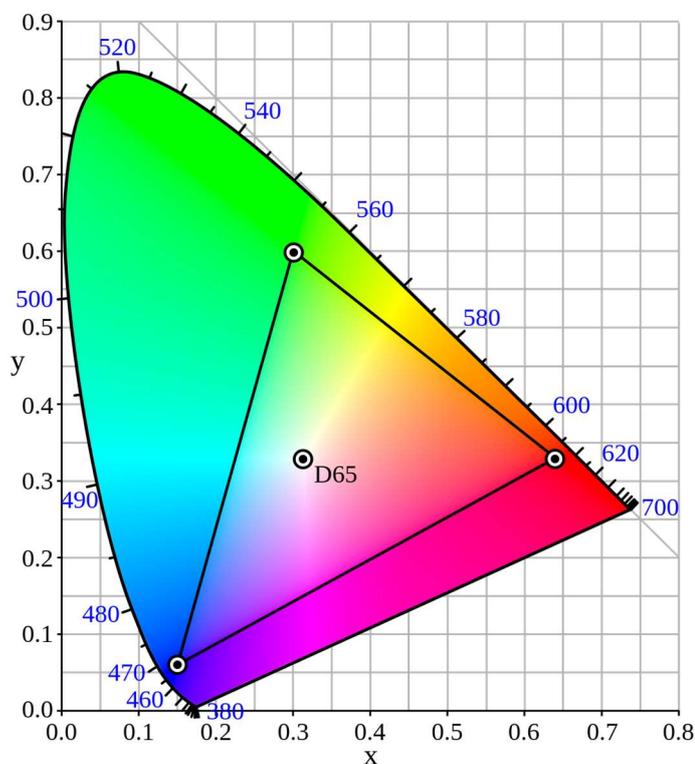
Wie in Abb. 1-3 zu erkennen, hat ein HDR-Display einen größeren Dynamikumfang als das Steady-State HVS und schafft somit eine höhere Immersion als Standard Dynamik Range (SDR), auch LDR genannt. Zum Beispiel durch die Folge einer hellen auf eine dunkle Szene muss sich das Auge erst kurz anpassen bevor es die Szene richtig wahrnimmt. So braucht das Auge erst einmal einen kurzen Moment, um in den Schatten Details zu erkennen. Oder der Zuschauer wird z.B. bei einer Tunnelfahrt erst einmal geblendet und seine Augen adaptieren sich dann nach einer gewissen Zeit an die neue Helligkeit. Dies kann inszenatorisch ein interessanter Aspekt sein. Denn der Zuschauer erlebt den Film in einer weiteren Ebene, der physiologischen Anpassung seiner Augen.

1.2 Farbwahrnehmung

Für die Farbwahrnehmung des Menschen sind hauptsächlich die Zapfen zuständig. Diese können in drei Gruppen unterteilt werden [3]:

- S-Zapfen (S für Short) sind empfindlich für kurzwelligeres blaues Licht im Wellenlängenbereich von etwa 400-500 nm (auch Blau-Zapfen genannt)
- M-Zapfen (M für Medium) sind empfindlich für mittlere Wellenlängen im Bereich von etwa 450-630 nm (auch Grün-Zapfen genannt)
- L-Zapfen (L für Long) sind empfindlich für langwelliges rotes Licht im Wellenlängenbereich von etwa 500-700 nm (auch Rot-Zapfen genannt)

Die Internationale Beleuchtungskommission erstellte 1931 ein Farbsystem um alle wahrnehmbaren Farben darzustellen, das CIE-Normfarbsystem oder auch CIE-XYZ. Für eine bessere Darstellung des dreidimensionalen Farbraums wurde die CIE-Normfarbtafel entwickelt (s. Abb. 1-4).



nehmbarer Farben darzustellen, das CIE-Normfarbsystem oder auch CIE-XYZ. Für eine bessere Darstellung des dreidimensionalen Farbraums wurde die CIE-Normfarbtafel entwickelt (s. Abb. 1-4).

Abbildung 1-4: CIExy 1931 Farbtafel mit sRGB-Farbraum

Dieses Farbsystem enthält auch theoretische Farben, die so nicht reproduzierbar sind. In Abbildung 1-4 ist zudem ein sRGB-Farbraum eingezeichnet, den die meisten Computerdisplays zur Darstellung verwenden. Dieser Farbraum deckt nur einen kleinen Teil der menschlichen Wahrnehmung ab. Mit HDR ist es das Ziel durch ein Wider Color Gamut, WCG, einem größeren Farbraum der menschlichen Farbwahrnehmung näher zu kommen.

2 HDR – High Dynamic Range

2.1 Was ist High Dynamic Range?

Fotografen kennen das Phänomen, wenn eine Landschaft mit dem Handy fotografiert wird, dann ist entweder der Himmel korrekt belichtet aber die Landschaft zu dunkel, oder aber die Landschaft sieht richtig aus und dafür ist der blaue Himmel dagegen nur weiß. Dies liegt am Dynamikumfang der Handykamera. Dieser reicht nicht aus, um beide Objekte richtig zu belichten. Das ist auch das Problem der Standard Dynamic Range, kurz SDR. Diese kann nur einen Bruchteil von dem, was der Mensch sieht, gleichzeitig darstellen (s. Abb. 2-2). So wirbt HDR mit dunklerem Schwarz und hellerem Weiß. Anfangs wurde noch geschertzt, dass eine Sonnenbrille beim HDR schauen gebraucht wird, was völliger Unsinn ist, denn der menschliche Dynamikumfang ist weitaus größer als der von HDR-TVs. Doch was ist mit der HDR-Funktion in z.B. Handys oder Kameras? Dieses HDR hat mit HDR-Video wenig zu tun. Denn beim HDR-Bracketing werden verschiedene Belichtungsreihen aufgenommen, meist drei an der Zahl, die die einzelnen Bereiche wie Schatten, Mitteltöne und Lichter jeweils richtig belichteten und im Nachhinein zusammenfügen. Dadurch wird aber nicht der komplette Dynamikumfang dargestellt, da dieser von der SDR begrenzt ist. So wird ein höherer Dynamikumfang in einen kleineren Container komprimiert und dadurch wirkt das Bild dann auch meist unnatürlich, da sich die Helligkeitsunterschiede nicht so stark wie in der Wirklichkeit unterscheiden. Und, der Mensch würde diese so auch nie, ohne seine Augen anzupassen zu müssen, gleichzeitig wahrnehmen.



Abbildung 2-1: Belichtungsreihe eines HDR-Bracketing und das fertige kombinierte Bild. Überbelichtetes, unterbelichtetes und kombiniertes Bild (von links) [4]

HDR bezeichnet die Möglichkeit ein breites Spektrum an Helligkeitsunterschieden in einem Videosignal darstellen zu können. So können gleichzeitig hellere Lichter und tiefere Schwärzen dargestellt werden und somit der Kontrastumfang in Bezug auf die Standard Dynamic Range (SDR) erheblich vergrößert werden. Zusätzlich zum vergrößerten Dynamikumfang beinhaltet HDR auch ein größeres Farbspektrum für brillantere Farben, mehr Bits pro Pixel für feinere Abstufungen, sowie eine Transferfunktion, angelehnt an die menschliche Wahrnehmung, die das elektronische Signal auf die Bits des Bildes überträgt.

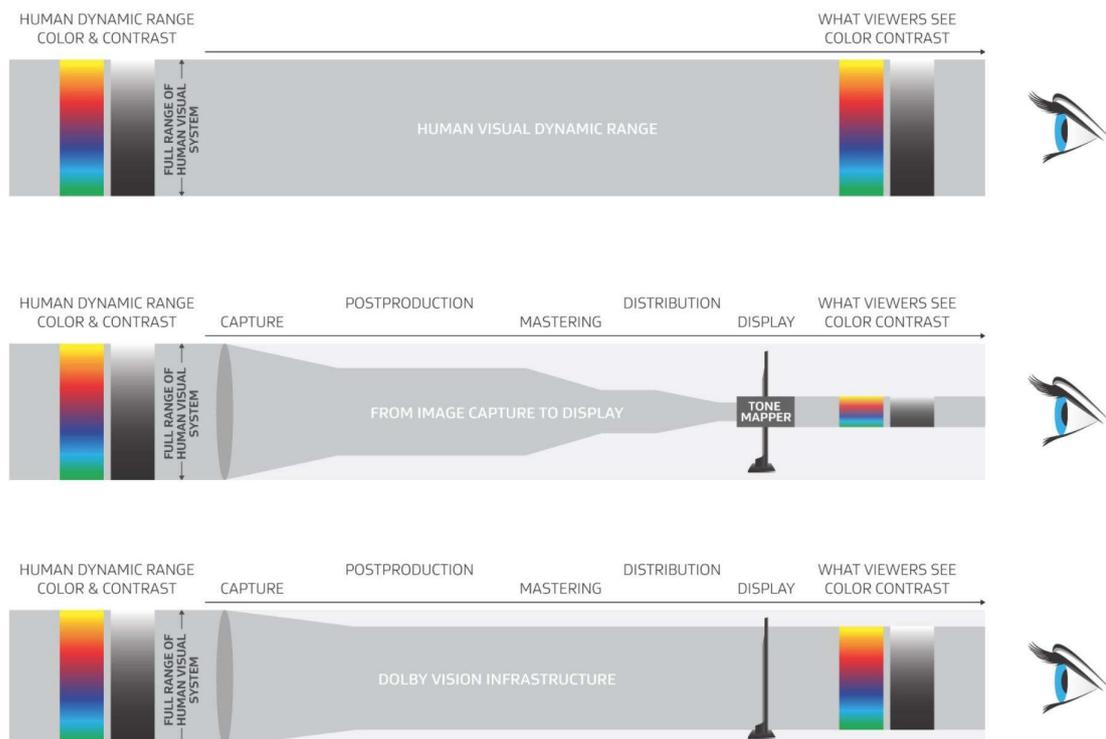


Abbildung 2-2: Vergleich von dem Verlust von Dynamic Range bei SDR und HDR Produktionen. Der menschliche Dynamikumfang (oben) im Vergleich zu SDR (Mitte) und HDR (unten) [5]

In der Spezifikation BT.709 von der ITU für High Definition TV (HDTV) wird ein Dynamikumfang von 0,0002 bis 100 cd/m² (Nits) festgelegt. HDR spreizt diesen von 0,00005 bis auf über 1000 cd/m² an [6]. In Abb. 2-2 ist der Verlust von Dynamik- und Farbumfang von Aufnahme bis Ausgabe auf einem Monitor nochmal grafisch dargestellt. Die meisten Kameras für Kinoproduktionen, wie von Arri, Red oder Panasonic, können bereits einen Dynamikumfang von 12-16 Blenden [7, S. 1-2] (ein Stop ist die Dopplung bzw. die Halbierung von Licht) aufnehmen. Jedoch wurde dieser Umfang bis vor der

Entwicklung von HDR nur bei qualitativ hochwertigen Filmprints genutzt. Denn die Limitierung von HD und somit Rec.709 hat lediglich 5 Blenden zugelassen hat. So wurden bisher diese 12 Blenden in den kleinen Container von Rec.709 gepackt und versucht, das Bild bestmöglich auszubalancieren. Mit HDR kann jetzt nahezu der gesamte Dynamikumfang moderner Kameras genutzt und auch auf einem Monitor dargestellt werden. Was natürlich viele Neuerungen an narrativen sowie auch technischen Aspekten mit sich bringt. Dies könnte mit dem Sprung von SD zu 4K verglichen werden.

Die Durchsetzung neuer Standards ist auch immer von der Adaption des Marktes sowie der Endverbraucher abhängig. Gerade in den letzten Jahren, mit dem Aufkommen von HD sowie UHD, begaben sich Endverbraucher auf der Suche nach einem neuen TV in einen wahren Dschungel aus verschiedenen Standards. Dieser wird sich in den nächsten Jahren konkretisieren, wenn mehr HDR-Inhalte produziert werden.

2.2 Marktübersicht HDR-TVs

Laut einer Prognose des Marktforschungsunternehmens IHS Markit (Abb. 4), wird es in den nächsten Jahren einen spürbaren Zuwachs an HDR kompatiblen Geräten geben. Was zudem voraussehbar ist, ist die Zunahme von Geräten, die HDR unterstützen. HDR kompatible Geräte bieten oft noch nicht die Brillanz und die technischen Eigenschaften für eine weitaus unverfälschte Darstellung von HDR-Video.

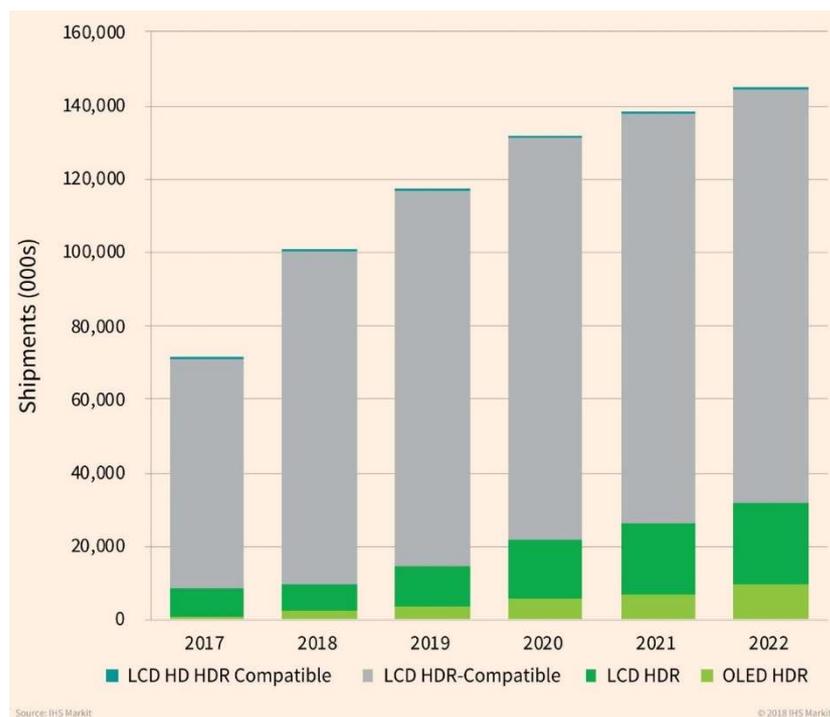


Abbildung 2-3: HDR TV Lieferungen nach Displaytechnologie [8, S. 33]

2.3 HDR Produktionen und Streaming Anbieter

Eine andere Blickrichtung bietet die Verbreitung von UHD Produktionen. Denn diese haben von 2017 bis Mitte 2018 erheblich zugenommen (Abb. 2).

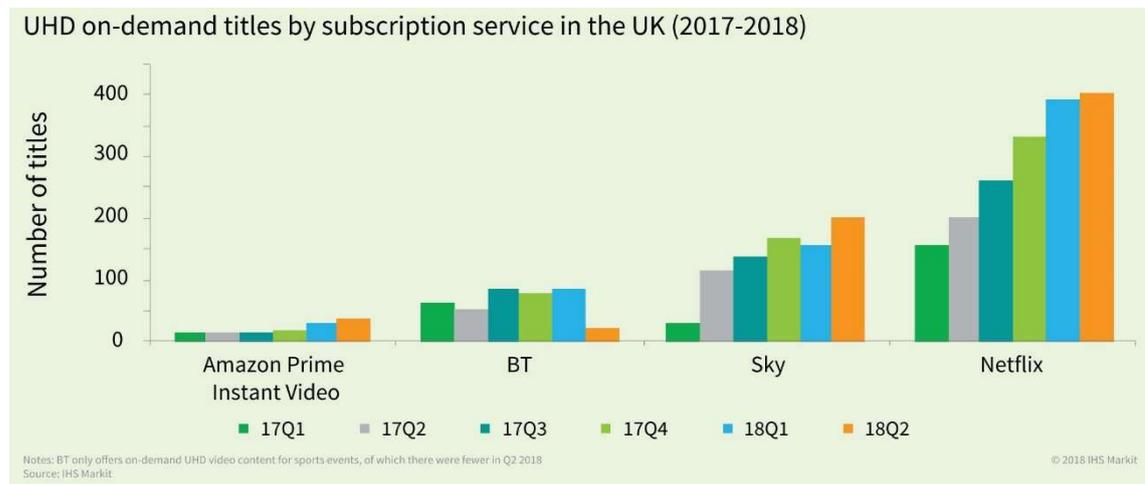


Abbildung 2-4: UHD on-demand Titel nach Abodiensten in den UK (2017-2018) [8, S. 36]

Die Streaminganbieter sind deutlich schneller mit der Adaption von HDR und UHD Inhalten, dies liegt womöglich auch an ihrem Unternehmensmodell und der unkomplizierteren Implementierung dieser Standards.

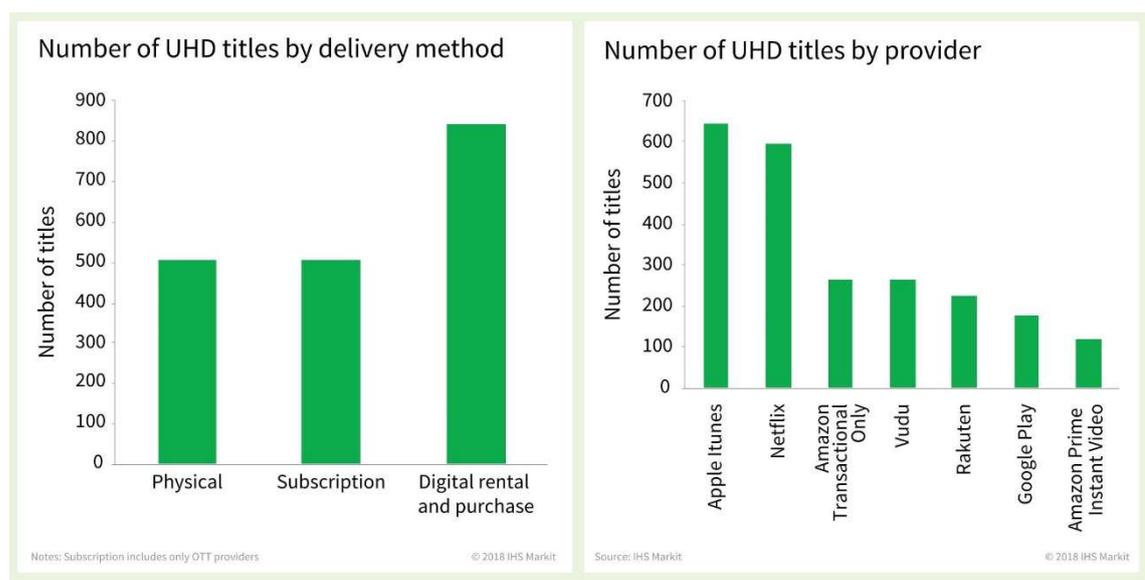


Abbildung 2-5: Anzahl der UHD-Titel sowie deren Distribution [8, S. 35]

2.4 Übersicht HDR-Standards

Momentan gibt es mehrere HDR-Standards. Die Formate ähneln sich größtenteils und haben eine etwas andere Herangehensweise. HDR10 ist ein freies Format, das für alle Hersteller zugänglich gemacht wurde. Die meisten HDR-TVs verstehen HDR10 und können dieses darstellen. Dolby Vision verspricht durch sogenannte Dynamische Metadaten, eine bessere Darstellung auf verschiedensten TVs. Dieser Standard ist jedoch proprietär, kostet Lizenzgebühren und schränkt TV Hersteller in ihrer Freiheit ein. Als Kontrahent zu Dolby Vision hat Samsung, HDR10+ als freies Format entwickelt, das wie Dolby Vision, dynamische Metadaten unterstützt. Da es immer wieder Fehden zwischen verschiedenen TV-Herstellern gibt und dieses Format von Samsung kommt, muss eine Adaption des Marktes erst einmal abgewartet werden. HLG ist das letzte größere Format, das von der NHK, japanisches TV und der BBC,ritisches TV, in Hinblick auf die terrestrische Übertragung entwickelt worden ist. HLG, Hybrid Log Gamma, ist rückwärtskompatibel zu UHD SDR-TVs die Rec.2020 unterstützen, in dem es bis 100 Nits, die klassische Gamma-Kurve enthält und danach in ein logarithmisches Signal wechselt. Technicolor entwickelt auch an mehreren eigenen Standards für HDR.

Tabelle 2.4-1: Übersicht der gängigen HDR-Standards

Standard	Einsatzbereich	Helligkeit	Hersteller
HDR10	B-Disc / Streaming / Gaming	1000 Nits	CTA
HDR10+	B-Disc / Streaming	4000 Nits	Samsung
Dolby Vision	B-Disc / Streaming	4000/10000 Nits	Dolby
HLG	TV-Übertragung	1000 Nits	BBC / NHK

3 HDR10 & Statische Metadaten

3.1 Wide Color Gamut und Volume Mapping

HDR bietet dunkleres Schwarz und helleres Weiß, aber auch kräftigere Farben, die dem HVS und der in der Natur vorkommenden Farben um einiges näherkommen, als mit dem alten Farbmodell Rec.709, das von der ITU für HDTV empfohlen wurde [9]. Außerdem beschreibt der Helmholtz-Kohlrausch Effekt, dass tief gesättigte Farben deutlich heller erscheinen als weniger gesättigte. Dies kann bis zu einer 30%-gen helleren Wahrnehmung beitragen [10], was wiederum die Immersion erhöht. 1931 wurde von der CIE der XYZ-Farbraum als Referenzfarbraum geschaffen. Die Farbprimaries X, Y und Z sind imaginär und dadurch stellt dieser Farbraum auch inexistenten Farben dar. Ein Farbraum besteht immer aus den Farbprimaries, den Eckpunkten des Farbraums, einem Weißpunkt, zum Beispiel D65 (Norm Licht), und einer Codierung [11]. Denjenigen, die schon mal etwas mit Farbräumen zu tun hatten, ist die „Schuhsohle“ (s. Abb. 3-1, links) sicher schon einmal begegnet. Diese stellt jedoch nur ein Teil des Farbspektrums dar. Color Volumes, oder Color Gamuts, sind dreidimensionale Räume, die verschiedene Helligkeitsabstufungen enthalten. So ist die Schuhsohle lediglich ein Ausschnitt einer bestimmten Helligkeitsebene.

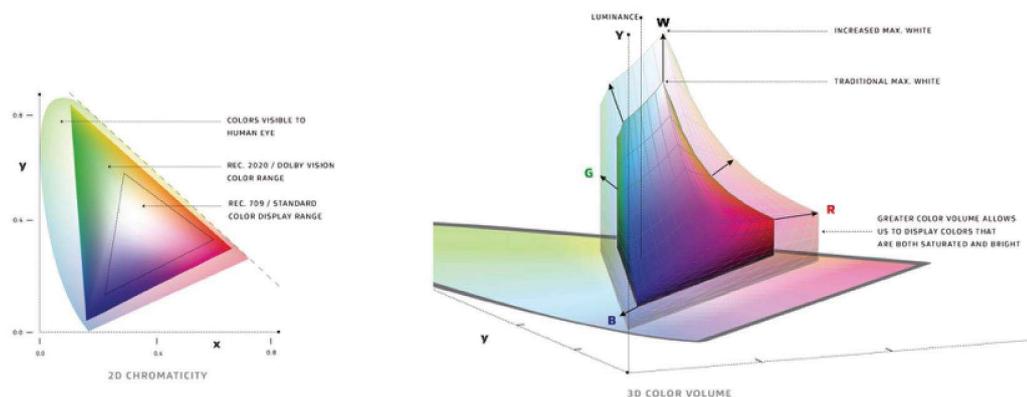


Abbildung 3-1: 2D- und 3D-Farbraum mit RGB-Primaries [12]

Doch wozu benötigt HDR einen neuen Farbraum? Rec.709 enthält nur 33,5% aller sichtbaren Farbwerte des CIE-Farbraums [6]. Das Ziel von HDR ist es, Farben und Dynamikumfang weitestgehend dem HVS und den in der Natur vorzufindenden Farben anzunähern, um dadurch ein realistischeres Bild zu schaffen. Deshalb musste ein neuer Farbraum

mit mehr Abdeckung als den in der Natur vorkommenden Farben geschaffen werden. Dr. Michael R. Pointer erstellte 1980 das Pointer Gamut, das von mehr als 4000 diffusen, real vorkommenden Oberflächen die Farbwerte enthält. Das Pointer Gamut wird als Referenz herangezogen das die in der Natur vorkommenden Farben repräsentiert. Der neue Farbraum BT.2020 deckt 99.9% des Pointer Gamuts ab und kann dadurch realistischere Farben abbilden [6].

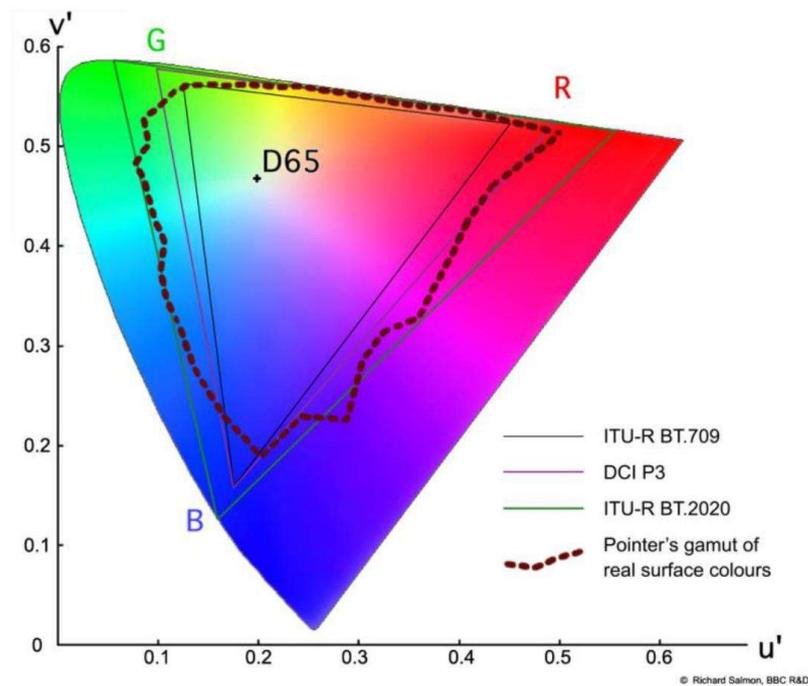
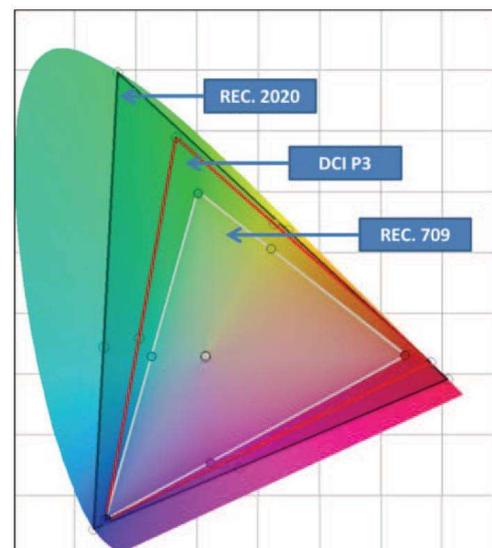


Abbildung 3-2: Pointer Gamut enthält die Farben von 4000 diffusen Oberflächen die in der Natur vorkommen [6]

Ein weiterer Vorteil eines so großen Farbraums ist die Möglichkeit, ebenfalls kleinere Farbräume wie Rec.709 und P3D65 darzustellen. So würde ein TV, der Rec.2020 abdeckt, auch Content der in P3 oder Rec.709 erstellt wurde, korrekt anzeigen.

Abbildung 3-3: Color Gamut Vergleich von REC 709, DCI P3 und REC 2020 [13, S. 6]



Doch da das Mastering Display, auf dem der Content gradated worden ist, meist ein größeres Colorvolume besitzt als das Display, auf dem der Endnutzer den Content sieht, muss ein

Color Volume Mapping durchgeführt werden. Dabei wird ein Farbraum in den anderen umgerechnet. Das Mapping wird auf zwei Ebenen durchgeführt. Zum einen muss die Helligkeit angepasst werden, da das Mastering Display meist eine höhere Spitzenhelligkeit besitzt als das Wiedergabegerät. In diesem Fall wird die Helligkeit auf die Helligkeit des Wiedergabegeräts getonemapped und ein Roll-Off, eine weiche Abstufung, auf die hellsten Bereiche angewendet, um hartes Clipping zu vermeiden. Zum anderen muss das Gamut angepasst werden, da das Mastering Display meist einen größeren Farbraum besitzt oder mehr davon abdecken kann als das Wiedergabegerät. Das Gamut Mapping muss so gestaltet sein, dass es die Sättigung der Farben wahrnehmungsspezifisch umrechnet und die Sättigungsunterschiede innerhalb des Bildes beibehält und nicht weniger gesättigte Farben entsättigt [6]. Die statischen Metadaten sind im ST.2086 Standard der SMPTE beschrieben und beinhalten die Farbprimaries (Eckpunkte des Farbraums), Weißpunkt, Spitzenhelligkeit und niedrigste Helligkeit des Mastering Displays [14]. Mit diesen Angaben kann das Wiedergabegerät den Inhalt umrechnen. Dadurch, dass nur ein Mapping über den gesamten Inhalt angewendet wird, können Probleme entstehen, auf die später noch eingegangen wird.

YCbCr und Y'CbCr

YCbCr ist ein Farbmodell, das für das Digitalfernsehen entwickelt und durch die ITU in ITU-R BT.601 standardisiert worden ist [15]. Bei diesem werden die Helligkeit Y und die zwei Farbkomponenten Cb (Blau-Gelb Kanal) und Cr (Rot-Grün Kanal) voneinander getrennt. Dadurch kann durch Chroma-Subsampling, also einer Farbunterabtastung, Bandbreite gespart werden. Durch die Verwendung eines nicht konstanten Helligkeitssignals breiten sich Farbfehler in das Helligkeitssignal aus. Dieser Fehler kann durch eine Gammakorrektur des Helligkeitssignals (Y'CbCr) behoben werden, doch diese Korrektur ist rechenaufwendig [16, S. 39]. Doch dieses Farbmodell kommt mit HDR und dem damit größeren Dynamikumfang und größerem Color Gamut an seine Grenzen. Die vorher kleinen Fehler werden in HDR um ein vielfaches größer und sichtbarer. So sind folgende Fehler zu beobachten [17]:

- Quantisierungsfehler wegen zu geringer Bittiefe
- Chroma subsampling Fehler wegen der nicht linearen Anordnung von Bits
- Verzerrungen beim Color Volume Mapping wegen falscher Annahmen des Farbtönen und Helligkeit
- Fehlerausbreitung von Farbkanälen in den Helligkeitskanal

Die ITU-R BT.2020 bietet eine Alternative um die Farbeindringung in das Helligkeitssignal zu unterbinden. Das konstante Helligkeitsformat $Y'C'bC'r$ löst aber nicht die Ausbreitung des Helligkeitssignals in die Farbkanäle $C'b$ und $C'r$ ab [17].

Bittiefe

Bittiefe ist die Anzahl der Bits für einen Farbkanal. Bei zu wenig Bittiefe werden Übergänge von einem Farbton zum anderen deutlich sichtbar. Wie in Abbildung 3-4 zu erkennen ist entstehen so harte Kanten beim Wechsel der Farbtöne und kein weicher Farbverlauf. Dies kommt mit dem Standard für HDTV häufiger vor, da Bildschirme, für die der Standard ursprünglich gedacht war, schon lange die 100 Nits Grenze übersprungen haben.

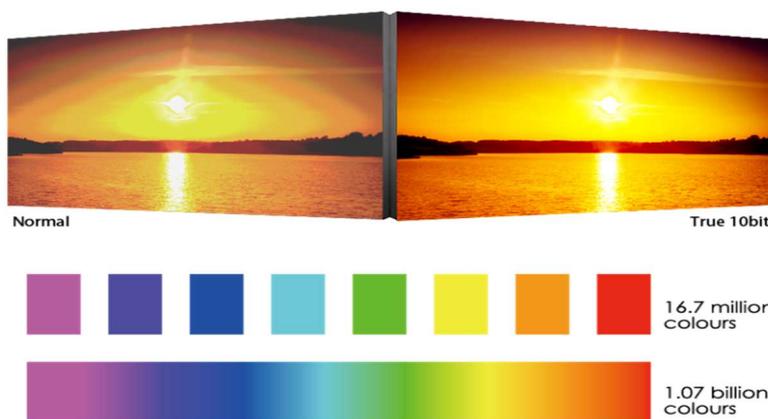


Abbildung 3-4: Vergleich von Farbverläufen zwischen 8-Bit und 10-Bit [18]

Für HDTV wurden in der ITU-R BT.709 8-Bit vorgesehen [6]. Das bedeutet 8-Bit pro Farbkanal, was wiederum 256 Farbabstufungen für eine Farbe ergibt. So können in einem 8-Bit Signal 16,8 Millionen Farben dargestellt werden. Durch das größere Color Volume in Rec.2020 werden für eine fehlerfreie Wiedergabe auch mehr Farbabstufungen benötigt. So benötigt HDR10, 10-Bit. Das sind 1024 Farbabstufungen pro Farbkanal und insgesamt 1 Milliarde Farben.

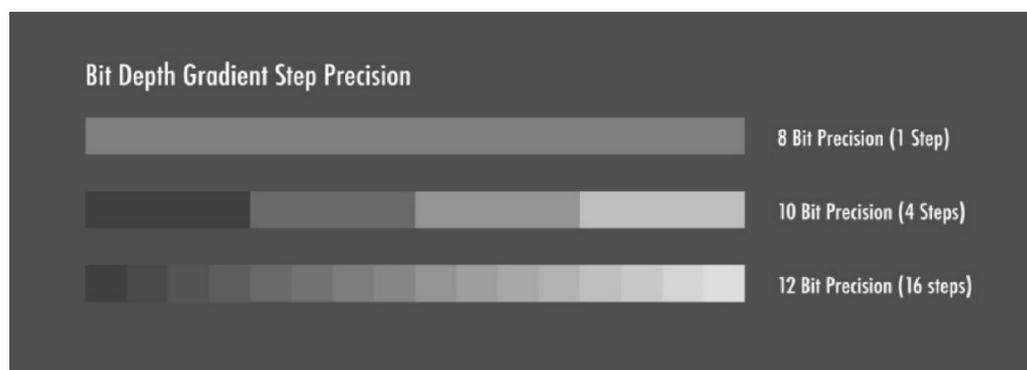
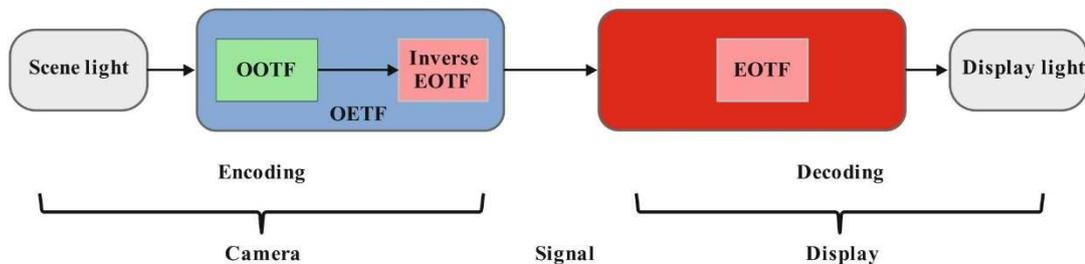


Abbildung 3-5: Vergleich von Bittiefen-Verlaufswerten und deren Präzision [19]

3.2 EOTF, OETF

Die Electro-Optical Transfer Function, kurz EOTF ist eine mathematische Funktion, um Spannungswerte oder Digitalwerte in Helligkeitswerte zu übersetzen. Dies findet im Wiedergabegerät statt und wandelt dort das Signal mit Hilfe der Funktion entsprechend um.



BT.2100-Ann1-04

Abbildung 3-6: OETF und EOTF Workflow mit PQ Kurve [20]

In SDR ist das Äquivalent hierzu das Gamma. Die Beziehung zwischen Helligkeits- und Digitalwerten bzw. Signal in HD-TV ist die Gamma 2.4 Kurve. Diese ist auf die Röhrenmonitore und ihren Kathoden bis in die 1930er Jahre zurückzuführen und wurde 2011, als fast schon alle Röhrenmonitore verschwunden waren, in der ITU-R BT.1886 niedergeschrieben. Röhrenmonitore hatten ungefähr einen Dynamikumfang von 0.1 cd/m² bis etwas über 100 cd/m². Diese Werte bestimmen auch heute noch HD-TV und andere Distributionssysteme [21, S. 563]. Doch das Gamma ist nur in ihrem limitierten Dynamikbereich der menschlichen Wahrnehmung ähnlich. Schon die meisten HDTV-Geräte haben eine Spitzenhelligkeit über 100 Nits. Gamma ist über diesen Bereich fehleranfällig und hat eigentlich bereits ausgedient. Denn übersteigt der Dynamische Umfang 100 Nits deutlich, zeigt Gamma seine Schwächen (Abb. 13).

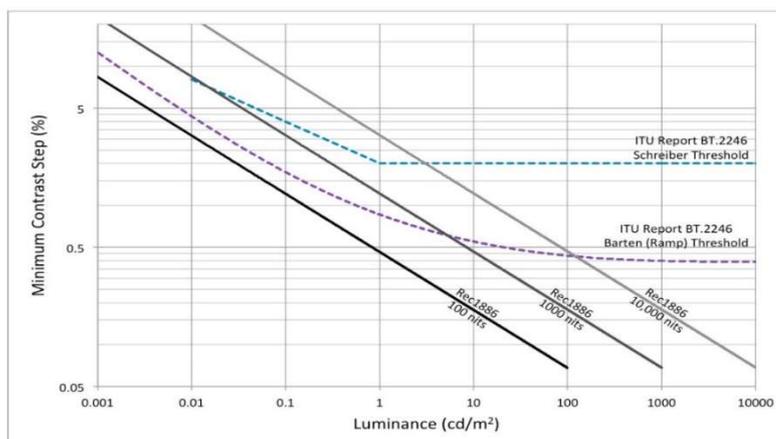


Abbildung 3-7: 12 Bit Rec.1886 Gamma Kurven [22]

Das Barten Modell dient schon seit mehreren Jahrzehnten als Grundlage zur Beurteilung der Kontrastwahrnehmung des HVS und basiert auf Physik, Optik und experimentell erfassten Daten [22].

$$CSF = \frac{1}{m_t} = \frac{M_{opt}(u)/k}{\sqrt{2 \left(\frac{1}{X_0^2} + \frac{1}{X_{max}^2} + \frac{u^2}{N_{max}^2} \right) \left(\frac{1}{\eta p E} + \frac{\Phi_0}{1 - e^{-(u/u_0)^2}} \right)}}$$

$$M_{opt}(u) = e^{-2\pi^2 \sigma^2 u^2}$$

$$\sigma = \sqrt{\sigma_0^2 + (C_{ab} d)^2}$$

$$d = 5 - 3 \tanh(0.4 \log(L X_0^2 / 40^2))$$

$$E = \frac{\pi d^2}{4} L \left(1 - (d/9.7)^2 + (d/12.4)^4 \right)$$

$k = 3.0$
 $\sigma_0 = 0.5 \text{ arc min}$
 $C_{ab} = 0.08 \text{ arc min/mm}$
 $T = 0.1 \text{ sec}$
 $X_{max} = 12^\circ$
 $N_{max} = 15 \text{ cycles}$
 $\eta = 0.03$
 $\Phi_0 = 3 \times 10^{-8} \text{ sec deg}^2$
 $u_0 = 7 \text{ cycles/deg}$
 $p = 1.2 \times 10^6 \text{ photons/sec/deg}^2/Td$

Abbildung 3-8: Contrast sensitivity function (CSF) des Barten-Modells [22]

Das Barten-Modell wurde in ITU-R BT.2246 zur Beurteilung der Kontrastwahrnehmung des HVS hinzugezogen. Die Kontrastwahrnehmung variiert mit der Helligkeit, dem Sehfeld und der Ortsfrequenz. Diese Variablen berücksichtigt das Barten-Modell indem es die minimal wahrnehmbaren Kontrastunterschiede zeigt [23]. Über dem Grenzbereich, dem Barten Threshold, sind Abstufungen in Verläufen möglich. Unter dem Grenzbereich sind für das HVS keine erkennbaren Fehler auszumachen (s. Abb. 3-9). Wie in Abbildung 3-7 ersichtlich, liegt Gamma mit 12-Bit unter dem Barten-Threshold bis 100 Nits. Doch wird der Dynamikumfang auf 1000 oder auf 10.000 Nits erhöht, dann liegt Gamma deutlich über dem Grenzbereich der Barten Ramp und ist daher anfällig für unsaubere Farbverläufe und Banding. Damit Gamma bei 10.000 Nits funktioniert, bräuchte das Signal eine Bittiefe von mindestens 15 Bit.

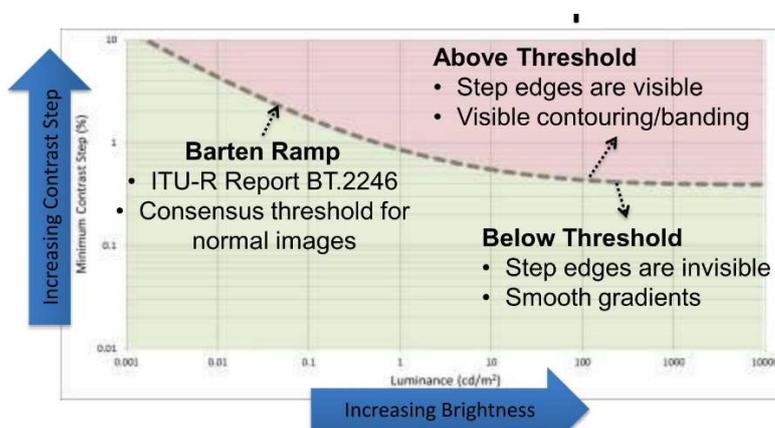


Abbildung 3-9: Darstellung des Barten Modells [24]

Hierdurch würden eine Menge Codewörter, vor allem in den hellen Bereichen, verschwendet werden. Deshalb musste eine neue EOTF für HDR entwickelt werden. Doch zu allererst sollte vorher die Präferenz von Helligkeit festgelegt werden, an die sich die neue EOTF orientiert. Dazu führte Dolby eine Studie zur Helligkeitswahrnehmung durch, um vor allem festzustellen, welche Werte für diffuses Weiß und Schwarz von Betrachtern bevorzugt werden, um daraufhin die EOTF an diesen Werten auszurichten. Das Versuchsssetup von Dolby, mit einem Christie CP2230 Projektor, erreichte einen Dynamikumfang von 0.004 bis 20.000 Nits. [21, S. 564] und deckt somit einen Bereich ab, der bisher so nicht reproduzierbar ist.



Abbildung 3-10: Versuchsaufbau von Dolby: Mit einem NEC Projektor (links) und dem finalen Aufbau mit einem Christie Projektor (rechts) [21, S. 564]

Die Studie fand heraus, dass bei 90% der Probanden die Schwarzwerte bei 0.005 cd/m^2 und diffuses Weiß bei ungefähr 3000 cd/m^2 liegen sollten. Doch ein Display sollte auch Specular Highlights, also z.B. das Glänzen einer Uhr oder spiegelnde Oberflächen, die um einiges heller sind als diffuses Weiß, darstellen können. 84% waren dort mit einer Helligkeit von 11.000 cd/m^2 zufrieden [21, S. 565].

Table 1: Luminance preference levels sorted by experience.

Luminance (cd/m^2)	50%		84%		90%	
	black	diffuse white	black	diffuse white	black	diffuse white
Technical	0.0398	691.8	0.0048	3162	0.0011	4677
Arts	0.1622	380.2	0.0657	1995	0.0500	3090
Naive	0.1820	631.0	0.0741	1445	0.0661	1995

Abbildung 3-11: Bevorzugte Helligkeitslevel geordnet nach Erfahrungsgruppen [21, S. 565]

Durch diese Versuchsreihe konnte Dolby nun einen Dynamikumfang festlegen, der die meisten Betrachter zufriedenstellen würde und auch für die Zukunft ausreichend weit angelegt war. In der Abbildung 3-12 sind die Studienergebnisse nochmals grafisch aufgearbeitet dargestellt und auch mit herkömmlichen Bildschirmtypen verglichen, die sich momentan auf dem Markt wiederfinden lassen.

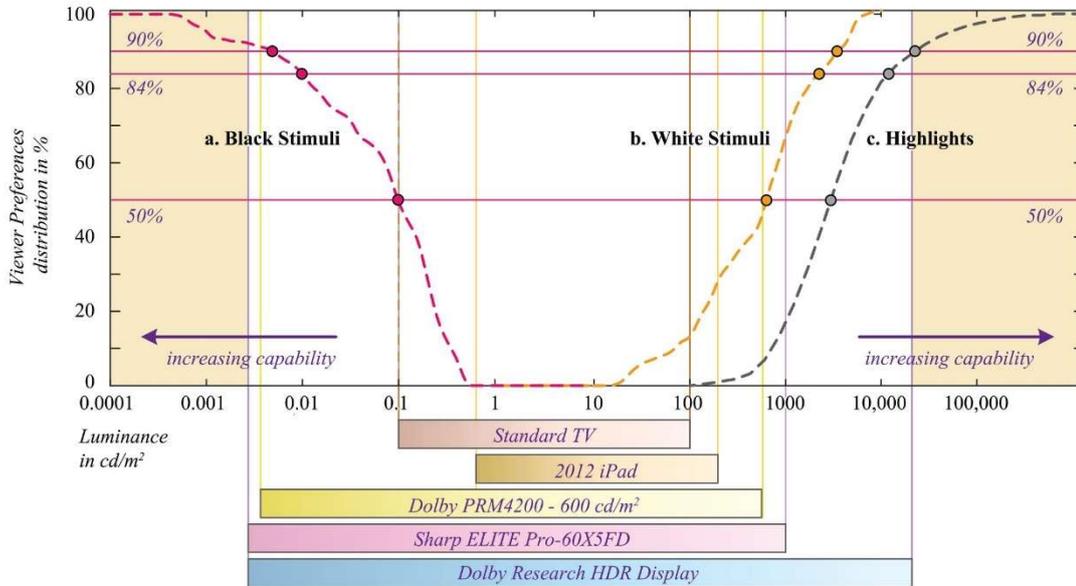


Abbildung 3-12: Ergebnisse der Versuchsreihe von Dolby für die Festlegung von Dynamikumfang von Displayeinheiten [21, S. 566]

Der Dynamikumfang, in dem die EOTF möglichst ohne Darstellungsfehler und effizient arbeiten sollte, wurde in der ST.2084 festgelegt. Der Dynamikumfang sollte zwischen 0 cd/m² und 10.000 cd/m² liegen [25]. Diese EOTF wird meist auch als PQ, Perceptual Quantization, bezeichnet.

$$F_D = \text{EOTF}[E'] = 10000 Y$$

$$Y = \left(\frac{\max\left[\left(E'^{1/m_2} - c_1\right), 0\right]}{c_2 - c_3 E'^{1/m_2}} \right)^{1/m_1}$$

Abbildung 3-13: Referenz PQ EOTF Formel [20]

where:

E' denotes a non-linear colour value $\{R', G', B'\}$ or $\{L', M', S'\}$ in PQ space in the range [0:1]

F_D is the luminance of a displayed linear component $\{R_D, G_D, B_D\}$ or Y_D or I_D , in cd/m². ^{4b}

So that when $R'=G'=B'$, the displayed pixel is achromatic.

Y denotes the normalized linear colour value, in the range [0:1]

$$m_1 = 2610/16384 = 0.1593017578125$$

$$m_2 = 2523/4096 \times 128 = 78.84375$$

$$c_1 = 3424/4096 = 0.8359375 = c_3 - c_2 + 1$$

$$c_2 = 2413/4096 \times 32 = 18.8515625$$

$$c_3 = 2392/4096 \times 32 = 18.6875$$

Die PQ ist so entwickelt, dass sie Bits so effizient wie möglich zuordnet und gleichzeitig die menschliche Wahrnehmung von Helligkeit berücksichtigt, daher Perceptual. Dadurch werden Darstellungsfehler minimiert. Die PQ ist so entwickelt worden, dass sie bei 12 Bits pro Farbkanal und einem Dynamikumfang von 0,0001 bis 10.000 Nits unter dem Barten Threshold liegt (s. Abb. 21). Bei 10 Bits pro Farbkanal, wie bei HDR10, liegt die PQ EOTF leicht über dem Barten Threshold und kann unter Umständen zu sichtbarem Stepping führen. Dieses ist jedoch in den meisten Fällen nicht bemerkbar, da dies im Bildrauschen meist untergehen.

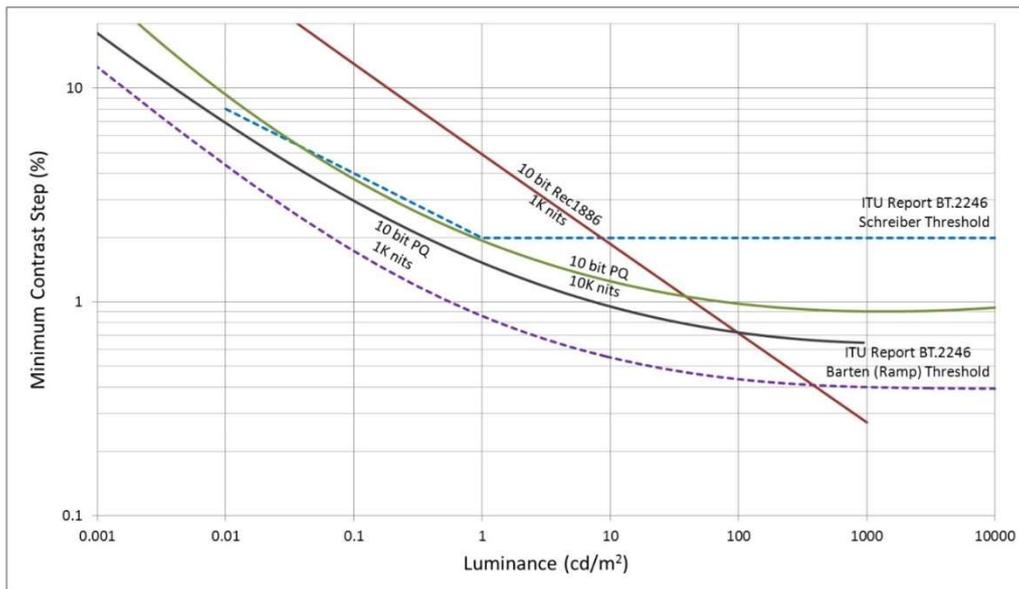


Abbildung 3-14: 10 Bit PQ and Rec.1886 im Vergleich [22]

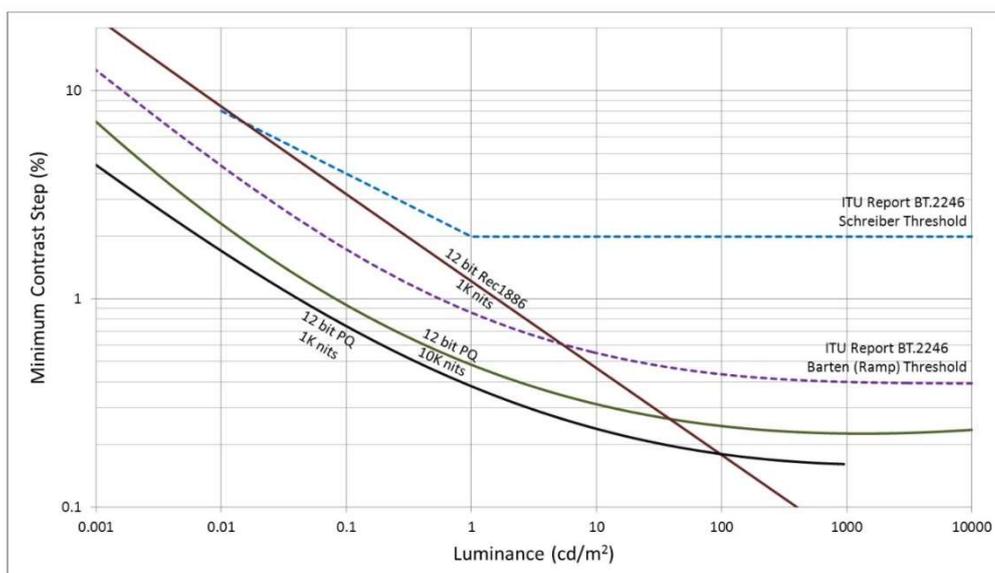


Abbildung 3-15: 12 Bit PQ und Rec.1886 Gammakurven [22]

Im Vergleich verschwendet die Gammafunktion vor allem in den hellen Bereichen Bits (s. Abb. 3-16) und Log-Formate verschwenden Bits in den dunklen Bereichen. Dadurch sind Log-Formate zwar für die Aufzeichnung geeignet, um Details in den dunklen Bereichen zu erhalten aber nicht für die Übertragung und das Abspielen, wo die Effizienz im Vordergrund steht.

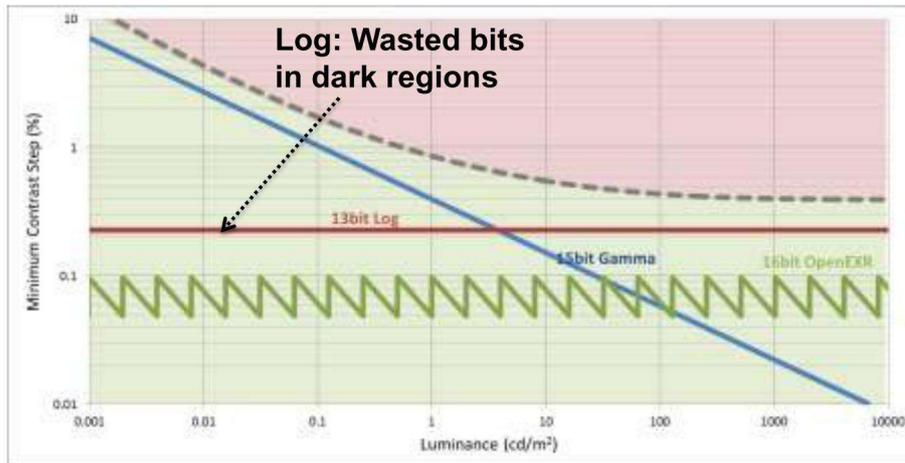


Abbildung 3-16: Darstellung von Log, OpenEXR und 15bit Gamma im Vergleich zum Barten Threshold [24]

Doch es könnte die Frage aufkommen, ob 10.000 Nits nicht etwas zu hoch gegriffen sind und dadurch Bits in den höheren Helligkeitsbereichen verschwendet werden. Obwohl dieser Wert anhand der Studie von Dolby [21, S. 565] und den Vorlieben der Betrachter ermittelt wurde, kann bisher kein käuflich erwerbbarer Monitor diese Helligkeit darstellen. Teure, kommerzielle Monitore, erreichen lediglich eine Spitzenhelligkeit von 4000 Nits und Endverbrauchergeräte kommen nicht mal über die 2000 Nits. Doch durch den logarithmischen Aufbau der PQ Kurve nehmen die Bitwerte im Bereich über 5000 Nits lediglich 7% des Speicherbedarfs ein [24]. Dadurch geht ungefähr 10% des Speicherbedarfs, der bis jetzt noch nicht dargestellt werden kann, verloren. So verwenden die in HDR hinzugekommenen Helligkeiten relativ wenig Platz im Hinblick auf die Verteilung der Bitwerte. So liegt der Bereich von 100 Nits zum Beispiel nahe dem Mittelpunkt

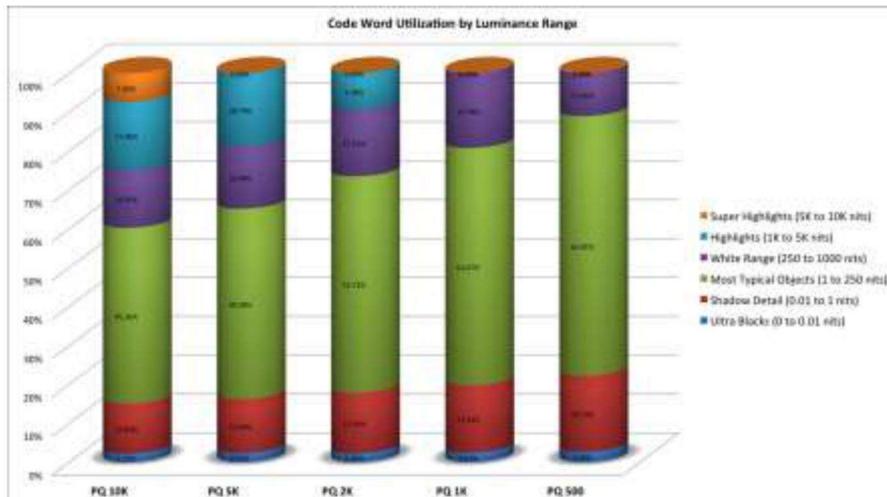


Abbildung 3-17: PQ Code Words Verteilung [24]

Im Vergleich dazu wäre die Verschwendung im traditionellen Gamma weitaus größer.

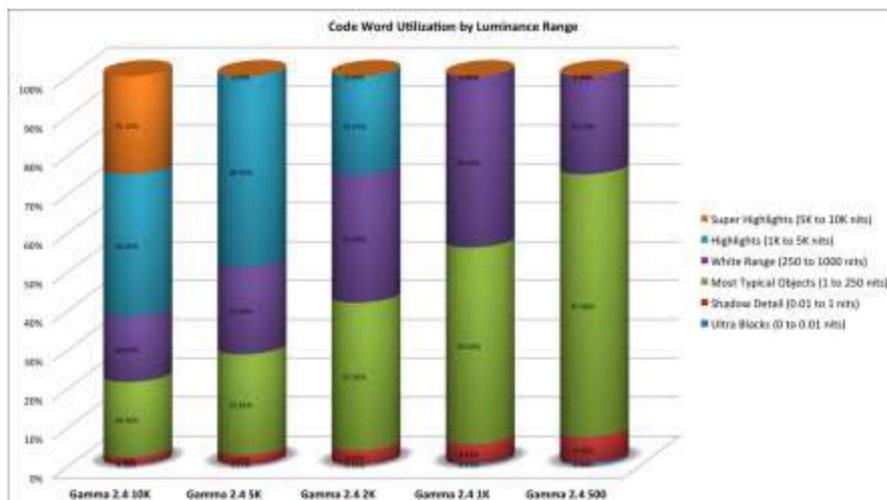


Abbildung 3-18: Gamma (Rec.1886) Code Words Verteilung [24]

Ein weiterer Aspekt, der zudem für den Maximalwert von 10.000 Nits spricht, ist die Zukunftssicherheit. Denn die technologische Entwicklung wird diesen Wert mit Sicherheit in Zukunft erreichen und anstatt wieder einen neuen Standard zu entwickeln, könnte dieser beibehalten werden.

3.3 Metadaten

Metadaten im Kontext von HDR, sind vor allem für die Anpassung des Bildes an das jeweilige Wiedergabegerät vorhanden. Die sogenannten Statischen Metadaten, die sich während der Wiedergabe des Videos nicht ändern, wurden in der ST.2086 von der SMPTE standardisiert und beinhalten Farbprimaries (RGB), Weißpunkt, Spitzenhelligkeit und niedrigste Helligkeit des Mastering Displays [14].

Die RGB Primaries und der Weißpunkt werden als $1/2 \cdot (X, Y)$ Werte gespeichert.

$$f(X_{Primary}) = 100,000 \times X_{Primary} \div 2$$

Abbildung 3-19: Berechnung der RGB Primaries von CIE XYZ 1931 [26]

$$f(Y_{Primary}) = 100,000 \times Y_{Primary} \div 2.$$

So wird ein Rot-Primarie von (0.68, 0.32) in CIE XYZ als R(34000,16000) gespeichert.

Die Maximalhelligkeit und die Minimalhelligkeit eines Mastering Displays wird in Nits*10.000 gespeichert. Der Dynamikumfang eines Displays von 0,0001 bis 1000 Nits wird als 10000000 und 1 in den Metadaten gespeichert.

Beispielhaft könnten die Metadaten eines DCI-P3 Displays mit einer Maximalhelligkeit von 1000 Nits, einer Minimalhelligkeit von 0,0001 Nits und einem Weißpunkt von D65

```
<r1:MasteringDisplayMinimumLuminance>1</r1:MasteringDisplayMinimumLuminance>
<r1:MasteringDisplayMaximumLuminance>10000000</r1:MasteringDisplayMaximumLuminance>
▼<r1:MasteringDisplayWhitePointChromaticity>
  <r2:X>15635</r2:X>
  <r2:Y>16450</r2:Y>
</r1:MasteringDisplayWhitePointChromaticity>
▼<r1:MasteringDisplayPrimaries>
  ▼<r2:ColorPrimary>
    <r2:X>34000</r2:X>
    <r2:Y>16000</r2:Y>
  </r2:ColorPrimary>
  ▼<r2:ColorPrimary>
    <r2:X>13250</r2:X>
    <r2:Y>34500</r2:Y>
  </r2:ColorPrimary>
  ▼<r2:ColorPrimary>
    <r2:X>7500</r2:X>
    <r2:Y>3000</r2:Y>
  </r2:ColorPrimary>
</r1:MasteringDisplayPrimaries>
```

wie folgt aussehen:

Abbildung 3-20: Metadaten für P3D65 / 1000 Nits Mastering Display

Mit diesen Angaben kann das Wiedergabedisplay sein Color Gamut und seinen Dynamikumfang mit dem des Mastering Displays vergleichen und bei der Wiedergabe das Video dementsprechend tonemappen.

Zwei weitere Metadaten, die zusätzlich zu den ST.2086 Metadaten von der BDA, Blu-ray Disc Association, verlangt werden, sind MaxFALL und MaxCLL [27].

MaxCLL, Maximum Content Light Level, beschreibt in Nits wie Hell der hellste Pixel im gesamten HDR-Video ist. Dieser Wert wird bei HDR10-Video bei 1000 Nits liegen, da dies der höchst mögliche Wert im Standard ist [5].

MaxFALL, Maximum Frame-Average Light Level, beschreibt den höchsten frame-durchschnittlichen Helligkeitswert in Nits. Dazu wird Frame für Frame die Durchschnittshelligkeit aller Pixel errechnet. MaxFALL ist ein erheblicher Wert, denn überschreitet der MaxFALL des HDR-Videos die des Displays, regelt dieses automatisch die Helligkeit herunter, auch wenn die Spitzenhelligkeit des Displays höher ist als der MaxFALL. Dies ist eine Schutzvorkehrung des Displayherstellers gegen Überhitzung, um damit eine Beschädigung des Displays zu verhindern. Durch diese statischen Metadaten wird der Videostream an das Wiedergabedisplay angepasst. Dadurch, dass nur die Extremwerte des gesamten Videos als Metadaten übergeben werden, können andere Szenen, die das Display wiederum ohne Anpassungen in ihrem vollen Dynamikumfang wiedergeben hätte können, womöglich in ihrer Dynamik beschränkt und der Creative Intent des Künstlers dadurch verfälscht wiedergegeben werden.

3.4 Probleme

Gerade wenn das Wiedergabedisplay in Helligkeit und Color Gamut niedrigere Werte aufweist, als in dem, in dem das Video gemastert wurde, muss das Bild in Farbe und Helligkeit entsprechend angepasst werden. Das Color Volume Mapping kann in zwei Bereiche unterteilt werden [28, S. 4]:

- Das Tone Mapping (Helligkeit) kann mit einem Roll-Off in den Schatten und den Highlights relativ einheitlich skaliert werden.
- Das Gamut Shaping (Farbigkeit) bekommt seine Parameter zur Umrechnung des Gamuts durch Entscheidungen, die während des Gradings getroffen werden. Das HDR Master wird durch SDR-Trims auf niedrigere Helligkeitswerte durch den Coloristen angepasst. Und diese Einstellungen werden bei der Umberechnung herangezogen (lediglich bei Dynamischen Metadaten).

Dadurch, dass Tone Mapping und Color Gamut Mapping bei HDR10 nur einmal über den ganzen Film angepasst werden, kann dies zu verschiedenen Problemen führen.

Schatten

Gerade in den dunklen Bildbereichen können Informationen schnell verloren gehen. Falls das Tone Mapping zu aggressiv in diesen Bereichen wirkt oder kein feines Roll-Off einsetzt, können Zeichnungen verloren gehen. Bedeutsame narrative Elemente könnten so im Film nicht dargestellt werden was zur Folge hätte, dass z.B. der Schurke, der sich im Schatten versteckt, nicht wahrgenommen werden kann.



Abbildung 3-21: Durch eine nicht ganz optimale Tonwertkorrektur kann Zeichnung in den Schatten verloren gehen (links), im Master dagegen sind diese korrekt belichtet (rechts) [29]

Color Hue Shift

Bei Wiedergabe auf einem Display, das lediglich eine niedrigere Helligkeit ausgeben kann als wie das Bild gemastert wurde, kann es zu Farbtonverschiebungen, dem Color Hue Shift kommen. So erscheint die Sonne deutlich mehr orange als ursprünglich gewollt.



Abbildung 3-22: Durch Helligkeitsänderung kann ein Hue Shift entstehen (links) im Vergleich das Master (rechts) [29]

Banding

Banding wird meistens mit zu wenig Bittiefe in Verbindung gebracht. Jedoch kann auch durch eine Tonwertkurve, die die Bit Werte nicht optimal anordnet, z.B. durch Knicke oder Ausschweifungen, Banding entstehen.



Abbildung 3-23: Banding kann durch eine falsche Tonemapping Kurve entstehen [29]

Clipping

Wenn Displays ab einem gewissen Grad an Helligkeit keinen Roll-Off anwenden, sondern diese Bereiche Clippen, gehen Informationen verloren, die vielleicht für die narrative Bedeutsam sein könnten. Ein Beispiel hierzu könnte eine „Erscheinung“ in den Wolken sein, die dadurch nicht mehr sichtbar wäre.



Abbildung 3-24: Gerade in den Highlights kann durch Tone Mapping Clipping entstehen [29]

4 Dolby Vision & Dynamische Metadaten

4.1 Unterschiede zwischen HDR10 und Dolby Vision

Dolby hat mit Dolby Vision ein proprietäres HDR Format entwickelt, das vor allem durch seine dynamischen Metadaten den Creative Intent des Regisseurs und des Coloristen weitest möglich erhalten soll. Diese dynamischen Metadaten werden vor allem für die Umrechnung des Color Volumes, also des Gamuts und der Tonalität, auf Wiedergabegeräten, die eine niedrigere Abdeckung des Farbraums und der Helligkeit besitzen, genutzt. Dolby Vision besitzt einen HDR10-Baselayer und baut dementsprechend auf HDR10 auf.

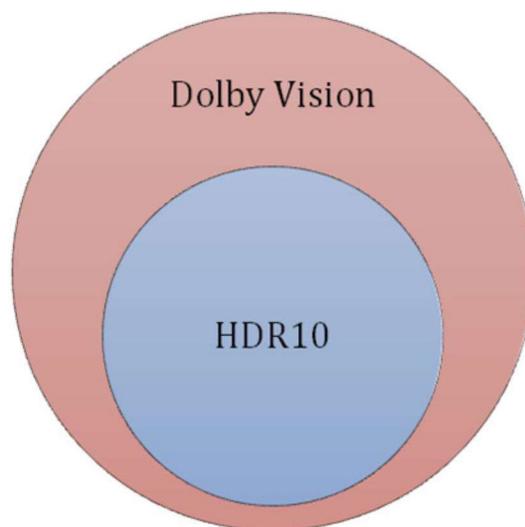


Abbildung 4-1: HDR10 Baselayer in Dolby Vision [5]

So sind HDR10 und Dolby Vision nicht komplett als zwei unterschiedliche HDR-Formate zu verstehen. Sondern Dolby Vision ist eher eine Erweiterung von HDR10, die das Erlebnis von HDR-Video ein-

heitlicher darstellt.

Im Kern besitzt Dolby Vision, DV, die gleichen Eckdaten wie HDR10. Beide haben ein Wider Color Gamut mit Rec.2020 und wenden die PQ ST.2084 EOTF an, um Bit Werte den Helligkeitswerten so anzuordnen, dass diese keine Code Values verschwenden und keine Bildfehler entstehen. Bei HDR10 wird auf 1000 Nits gemastert. Dolby Vision geht hier noch einen Schritt weiter und mastert mit einem Headroom (bis 10.000 nits) auf 4000 Nits. Bei der Bittiefe sind 10 oder 12 Bit bei DV möglich. Wie in Abb. 3-14 und 3-15 zu sehen ist, ist die 12-Bit Variante in Bezug auf das Barten-Ramp die sicherere Variante. Der entscheidene Unterschied liegt bei den Metadaten, die mit den beiden Formaten einherkommen. HDR10 bietet lediglich statische Metadaten, die sich nicht im Laufe des Videos ändern. Diese sind von der SMPTE in der ST.2084 standardisiert, hinzu kommen noch MaxFALL und MaxCLL, die aber auch beide statisch sind. Dolby Vision beinhaltet eine ganze weitere Palette an Metadaten, die sich dynamisch während des Films

verändern und jeweils für eine Szene bzw. einen Frame andere Werte annehmen können. Dadurch erzielt DV eine genauere Anpassung des Films an die unterschiedlichsten TV-Geräte. Diese Metadaten sind in der ST.2094-10 von der SMPTE standardisiert und näher beschrieben [30, S. 4].

Tabelle 4.1-1 Vergleich von HDR10 und Dolby Vision

Eigenschaften	HDR10	Dolby Vision
Wide Color Gamut	Rec.2020	Rec.2020
Helligkeit	1000 Nits	4000 Nits (bis 10.000 Nits)
Bittiefe	10 Bit	10 / 12 Bit
EOTF	PQ ST.2084	PQ ST.2084
Metadaten	ST.2086	ST.2086 & ST2094-10
Farbmodell	YCbCr	ICtCp / YCbCr
Codec	HEVC (Main 10)	HEVC / AVC (Main 8/10)

Dadurch, dass Dolby TV-Geräte lizenziert werden müssen, besitzen diese eine eigene konsistente Color Mapping Engine, die eine bessere Umrechnung gewährleistet und auf allen Dolby Geräten ein einheitliches Bild liefert [5]. Im Gegensatz zu HDR10-only Geräten, wo jeder Hersteller „seine eigene Suppe kocht“ und dadurch ein nicht gleichbleibendes Erlebnis zu Stande kommt.

Im Vergleich zu HDR10 verwendet Dolby Vision das Farbmodell ICtCp, das für HDR und WCG konzipiert worden ist. ICtCp hat die gleichen Operatoren wie Y'C'bC'r (NCL, Non Constant Luminance) und besitzt die gleichen Vorteile wie Y'cC'bcC'rc, aber verbessert dabei die Farbeinheitlichkeit [31]. ICtCp wurde wie die PQ EOTF in Bezug auf das HVS und seine Farbverarbeitung entwickelt.

Das Auge nimmt Farbe in drei Schritten wahr [32]:

- (1) Das eintreffende Licht wird mit den Zapfen erfasst. Dabei sind diese in langen, mittleren und kurzen Wellenlängen empfindlich.
- (2) Das lineare Licht wird in ein nicht lineares Signal umgewandelt
- (3) Das Signal wird dann separiert in Hell-Dunkel-Intensität, Gelb-Blau Tritan und Rot-Grün Tritan

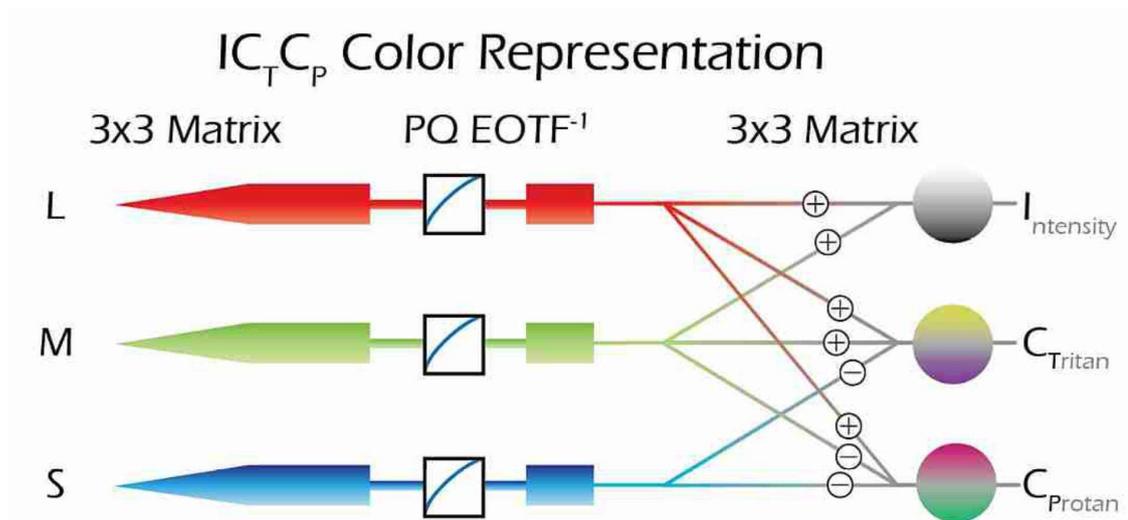


Abbildung 4-2: Beziehung zwischen HVS und ICtCp [31]

ICtCp entspricht in etwa der PQ Helligkeit Y und bietet eine bessere Separierung des Helligkeitskanals und der Farbkanäle. Die konstante Helligkeit von ICtCp verhindert vor allem bei Chroma-Subsampling oder Gamut-Mapping die Helligkeitsveränderung des Signals, wenn nur Änderungen an den Farbinformationen vorgenommen worden sind. Im Vergleich von Y'C'bC'r und ICtCp ist zu erkennen, dass es gerade in gesättigtem Blau und Rot zu Helligkeitsfehlern kommt. So geht eine Entsättigung mit dem Verlust von Helligkeit einher.

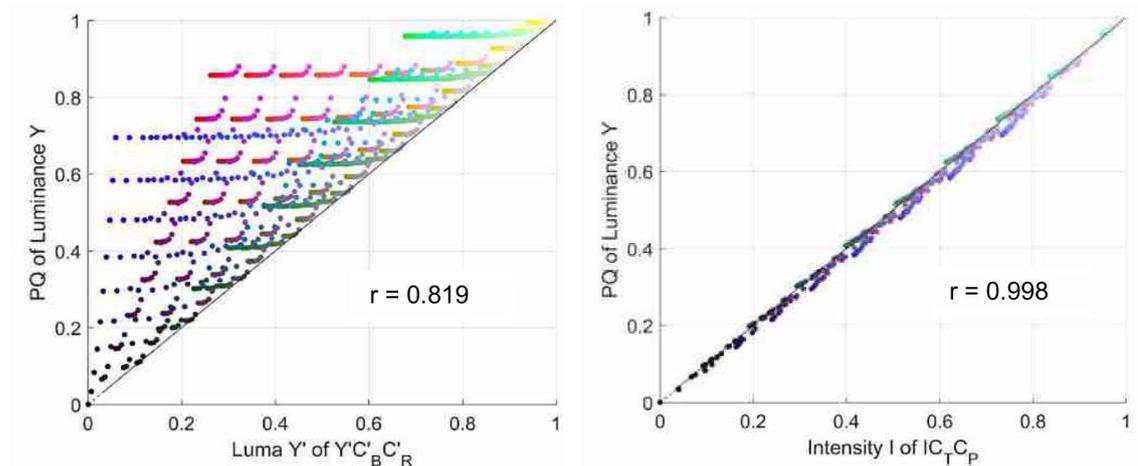


Abbildung 4-3: Vergleich PQ basierte BT.2020 NCL Y'C'bC'r und ICtCp [31]

Die Farbton-Linearität ist besonders dann bedeutend, wenn die Sättigung verändert wird. Dies kommt beim Gamut Volume Mapping oft vor. Durch die Linearität wird vor allem bei der Umrechnung von Color Volumes in andere Farbräume der Farbton beibehalten und nur die Sättigung geändert. Ist keine Linearität gegeben, sind Änderungen in der Sättigung auch im Farbton zu erkennen. Es sind komplexe Algorithmen nötig, um diesem Effekt entgegenzusteuern. Gerade bei Hauttönen, bei denen Hue-Shifts besonders schnell auffallen, ist eine Linearität des Modells beträchtlich. Werden die beiden Farbmodelle angesichts des Datensets von Hung und Berns [33] verglichen stellen sich im Blauton maximale Abweichungen, die bei ICtCp mehr als die Hälfte weniger betragen, dar. Dies wird auch in den Hauttönen widerspiegelt (s. Abb. 4-4).

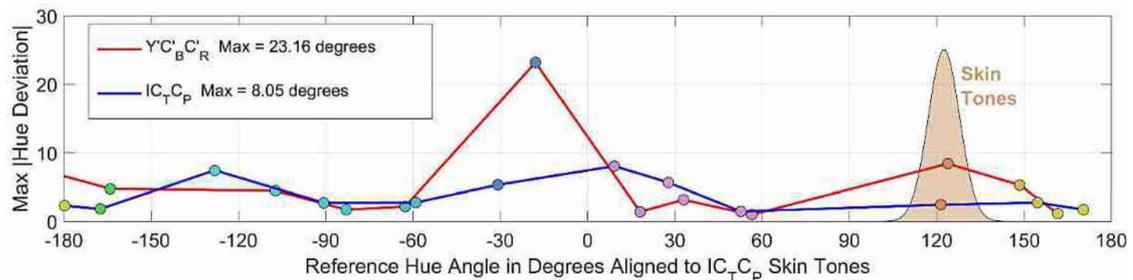


Abbildung 4-4: Vergleich der maximalen Farbtonfehler von ICtCp und PQ basierte BT.2020 NCL Y'CbCr, bei dem Datensatz von Hung und Berns [31]

4.2 Dynamische Metadaten

Im Gegensatz zu den statischen Metadaten, die in [14] standardisiert worden sind und für die gesamte Dauer des Films sich nicht ändern, sind die Dynamischen Metadaten in Dolby Vision Content Frame-By-Frame oder von Einstellung zu Einstellung unterschiedlich.

Dolby Vision Metadaten können in vier Gruppen unterteilt werden:

- (1) Statische Metadaten: ST.2084 (L0), MaxCLL & MaxFALL (L6)
- (2) Dynamische Metadaten zur Feststellung der Dynamic Range: L1, L3
- (3) Dynamische Metadaten zur Farbanpassung: L2, L8, L9
- (4) Dynamische Metadaten zur Festlegung des Aspect Ratios: L5

L1 Metadaten werden automatisch im Gradingssystem generiert und bestehen aus L1Min, L1Mid sowie L1Max und bilden somit den tonalen Umfang der Szene ab [34].

Tabelle 4.2-1 Dolby Vision Metadaten Version 4.0.2 XML für CMv4 [34]

Metadata Level/Field	Description
LEVEL 0	GLOBAL METADATA (STATIC)
Mastering Display	Describes the characteristics of the mastering display used for the project
Aspect Ratio	Ratio of canvas and image (active area)
Frame Rate	Frame Rate
Target Display	Describes the characteristics of each target display used for L2 trim metadata
Color Encoding	Describes the image container deliverable
Algorithm / Trim Version	CM algorithm version and Trim version
LEVEL 1	ANALYSIS METADATA (DYNAMIC)
L1 Min, Mid, Max	Three floating point values that characterizes the dynamic range of the shot or frame Shot-based L1 metadata is created by analyzing each frame contained in a shot in LMS color space and combined to describe the entire shot as L1Min, L1Mid, L1Max Stored as LMS(CMv2.9) and L3 Offsets
LEVEL 3	OFFSET TO L1 (DYNAMIC)
L3 Min, Mid, Max	Three floating point values that are offsets to L1 Analysis metadata as L3Min, L3Mid, L3Max L3Mid is a global user defined trim control L1 is stored as CMv2.9 computed values, CMv4 reconstructs RGB values with L1 + L3
LEVEL 2	BACKWARDS COMPATIBLE PER-TARGET TRIM METADATA (DYNAMIC)
Reserved1, Reserved2, Reserved3, Lift, Gain, Gamma, Saturation, Chroma and Tone Detail	Automatically computed from L1, L3 and L8 (lift, gain, gamma, saturation, chroma, tone detail) metadata for backwards compatibility with CMv2.9
LEVEL 5	PER-SHOT ASPECT RATIO (DYNAMIC)
Canvas,Image	Used for defining shots that have different aspect ratios than the global L0 aspect ratio
LEVEL 6	OPTIONAL HDR10 METADATA (STATIC)
MaxFALL, MaxCLL	Metadata for HDR10 MaxCLL - Maximum Content Light Level MaxFALL - Maximum Frame Average Light Level
LEVEL 8	PER-TARGET TRIM METADATA (DYNAMIC)
Lift, Gain, Gamma, Saturation, Chroma, Tone Detail, Mid Contrast Bias, Highlight Clipping 6-vector (R,Y,G,C,B,M) saturation and 6-vector (R,Y,G,C,B,M) hue trims	User defined image controls to adjust the CMv4 algorithms per target with secondary color controls
LEVEL 9	PER-SHOT SOURCE CONTENT PRIMARIES (DYNAMIC)
Rxy, Gxy, Bxy, WPxy	Stores the mastering display color primaries and white point as per-shot metadata

Die L1 Metadaten sind Fließkommazahlen, die den niedrigsten (L1Min) und höchsten (L1Max) Helligkeitswert beschreiben sowie den Durchschnittswert L1Mid. Zusätzlich zu den Level 1 Metadaten kann ein Offset im Grading-System festgelegt werden. Dieser Offset wird in den Level 3 Metadaten gespeichert und ist in der CMv4 hinzugekommen.

Die dynamischen Metadaten, die für die Farbanpassungen herangezogen werden, werden in den Level 2 Metadaten automatisch durch Level 1, 3 und Level 8 Daten berechnet und dienen der Abwärtskompatibilität zu dem älteren Content Mapping, kurz CMv2.9. Die CMv4 L8 Trim Controls können in zwei Gruppen unterteilt werden [34]:

- (1) Primary Trims: Lift, Gamma, Gain , Chroma Weight, Saturation, Tone Detail, Mid Contrast Bias, Highlight Clipping
- (2) Secondary Trims: Sättigungs- und Farbtonanpassungen der Farbkanäle Rot, Gelb, Grün, Cyan, Blau und Magenta

4.3 Trim Controls

Dolby Vision kann durch seine Metadaten auch SDR-Video rendern. Es ist verpflichtend, dass der Colorist das Mapping der gesamten Timeline auf SDR Rec.709 100 Nits überprüft, auch wenn dieser die SDR-Version nicht über diesen Weg erstellt [34]. Zudem können andere Trim-Ziele durch den Coloristen gesetzt werden, um das Mapping für diese Bereiche zu kontrollieren und anzupassen. Dies verbessert das Mapping über den gesamten Helligkeitsbereich, denn es kann passieren, dass das Mapping nicht ganz dem entspricht, was die Intention des Coloristen war. Bei einem auf 4000 Nits gemasterten Film wird zudem das Mapping für 100 Nits kontrolliert. Durch die breite Helligkeitsspanne kann es vorkommen, dass bei 1000 Nits das Mapping nicht ganz korrekt ist.

Deswegen empfiehlt Dolby zusätzlich zu dem 100-Nit-Trim [34] folgende Target-Trims [34]:

- Für ein 4000-Nit Dolby Vision HDR Master ein 1000-Nit-Trim
- Für ein 1000-Nit Dolby Vision HDR Master ein 600-Nit-Trim

Diese Trims können auch vom Kunden verlangt werden. Diese zusätzlichen Trims sind mit wenig Mehrarbeit verbunden, da die großen Anpassungen bereits beim 100-Nit Ziel gemacht worden sind und meist nur bei wenigen Einstellungen noch angepasst werden müssen. In den L3-Metadaten wird ein globales Offset durch den Coloristen festgelegt.

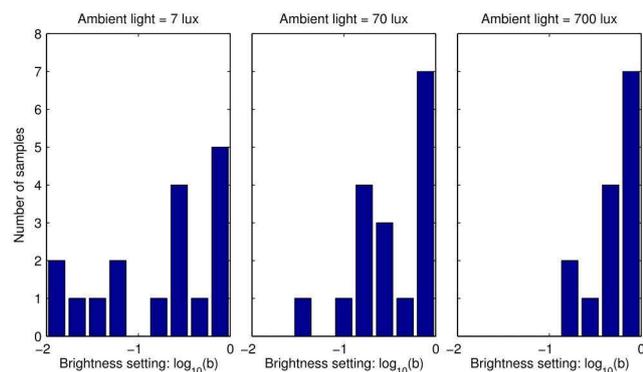
Dieser wird in Resolve, als Mid Tone Offset angezeigt und auf Framebasis oder Shotbasis festgelegt. In diesem Offset wird die Belichtung der Mitteltöne angepasst, ohne dabei die Schatten oder Lichter zu beeinflussen.

Die Level 8 Metadaten unterteilen sich in der Grading Software nochmal in Primary und Secondary Trim Controls. Die Primären Trim Einstellungen betreffen die Saturierung der einzelnen Farbkanäle. Diese können entweder einzeln oder insgesamt angesteuert werden. In den Sekundären Trim Einstellungen können Farbtonverschiebungen in den einzelnen Farbkanälen adressiert werden.

4.4 Sehumgebung

Umgebungslicht hat maßgebliche Auswirkungen auf die Sichtbarkeit von Details und den wahrgenommenen Kontrast in Bildern, die auf Displays gezeigt werden [35]. Ist die Umgebung heller, bevorzugten Probanden einer Studie die helleren Displayeinstellungen in dem TV-Gerät [36]. Die bevorzugten Einstellungen in geringem Umgebungslicht waren relativ breit gefächert. die Einstellungen konkretisierten sich jedoch, je höher das Umgebungslicht wurde.

Abbildung 4-5: Ergebnisse der bevorzugten Helligkeitseinstellung in Relation zum Umgebungslicht [36]



Gerade in den dunklen Bereichen können Zeichnungen durch eine zu helle Umgebung verloren gehen. In einer Studie wurde festgestellt, dass bei Helligkeitswerten von 0,01-1,0 cd/m² die Umgebungshelligkeit einen großen Einfluss auf die Sichtbarkeit dieser Zeichnungen hat und ab einer Helligkeit von 1 cd/m² das Umgebungslicht keinen großen Einfluss mehr spielt [37].

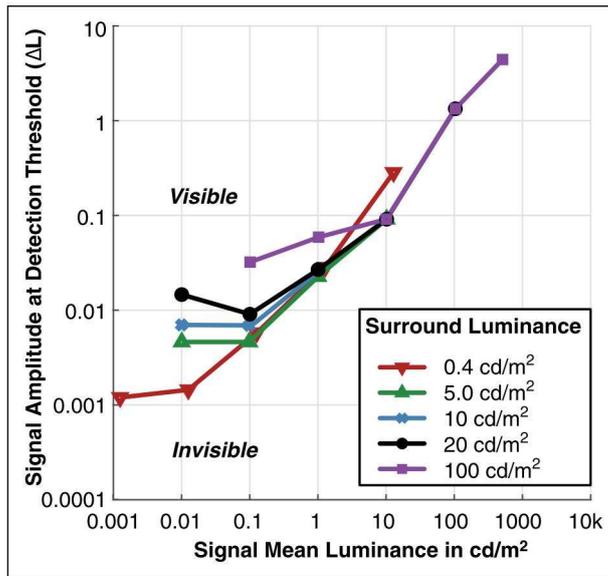


Abbildung 4-6: Ergebnisse der Studie.
Y-Axe: cd/m^2 [37]

Dolby Vision IQ wurde auf der CES 2020 mit dem neuen Panasonic HZ2000 OLED und der LG Gallery OLED Series vorgestellt [38]. Mit dieser Neuerung zu Dolby Vision geht Dolby noch einen Schritt weiter und bezieht die Umgebung, in der der Content geschaut wird, mit in die Um-

rechnung ein. Dolby Vision IQ erkennt durch einen eingebauten Lichtsensor die Umgebungshelligkeit und passt das Bild, zusammen mit den dynamischen Metadaten, entsprechend der Lichtverhältnisse an [38].

Dadurch, dass nicht jeder in einem verdunkelten Zimmer seine Lieblingsserie oder Lieblingsfilm in HDR anschaut, sondern dies auch oft in einem hell erleuchteten oder Tageslicht geflutetem Raum stattfindet, sieht verschwinden oft in den dunklen Bereichen die Zeichnung. Diesen Effekt steuert Dolby Vision IQ mit einer Anhebung der Helligkeit des Displays entgegen.



Abbildung 4-7: Simulierte Visualisierung von HDR ohne Dolby Vision IQ

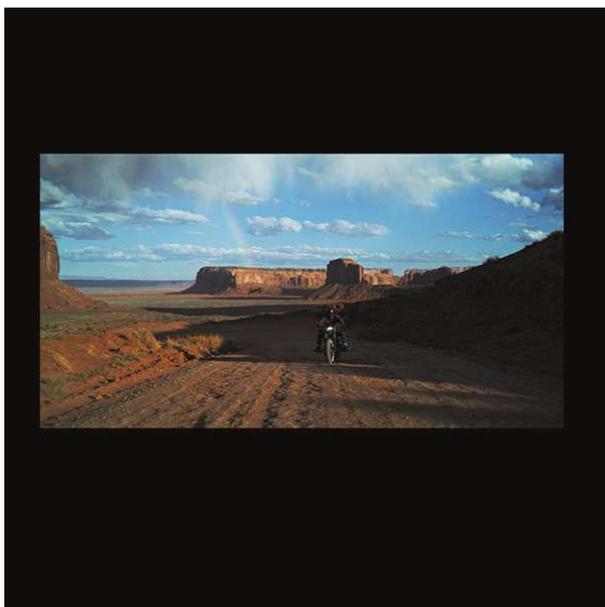


Abbildung 4-8: Simulierte Visualisierung von HDR mit Dolby Vision IQ

5 Mastering in Dolby Vision

Für das Mastering und Color Grading in Dolby Vision sind ein paar technische Voraussetzungen nötig, wie ein Display, das HDR möglichst akkurat wiedergibt und eine Grading Software, die HDR und Dolby Vision unterstützt. Gerade wenn auf Studio Monitore als professionelle Referenz gesetzt wird, kann dies schnell ins Geld gehen. Eine Möglichkeit, die Kosten in diesem Bereich überschaubar zu halten, ist auf einem Consumer HDR-TV zu graden. Diese geben kein so akkurates Bild wieder wie ein Referenzmonitor, können aber kalibriert ein annehmbares Bild liefern. Es führt kein Weg um eine professionelle Grading Software, wie Lustre, Film Master oder Baselight herum. Grading Tools in Schnittprogrammen, wie Premiere Pro oder FCP-X, sind in HDR bisher nicht zu gebrauchen. In diesem Versuchsaufbau wurde mit DaVinci Resolve 16, einem LG C9 HDR Display, einer Blackmagic Decklink 4K Mini Monitor und einer Dolby Lizenz gearbeitet. Als Material wurde auf Raw Sample Footage von Arri und Red zurückgegriffen.

Neben den technischen Voraussetzungen, die für das Mastering gegeben sein müssen, sind auch Grundlagenkenntnisse zu Farbräumen, EOTFs und HDR erforderlich. Diese sind in dieser Arbeit zuvor vermittelt worden. Auch in der Grading Software gibt es einige Besonderheiten, auf die geachtet werden müssen, ebenso wie gewisse Einstellungen, die notwendig sind.

5.1 Hardware und Referenz-Monitore

Die technischen Voraussetzungen, die erfüllt sein müssen, sind gerade im HDR-Grading relativ hoch, aber noch überschaubar. So muss eine Workstation vorhanden sein, die problemlos 4K Material mit 12-Bit verarbeiten kann und mit RAW-Material keine Probleme hat. Da die meisten Workstations in Grading Suiten bereits UHD Produktionen gestemmt haben, wird dies kein Thema sein.

5.1.1 HDR Studio Referenz Monitore

Leider gibt es auf dem Markt bislang keine große und vor allem erschwingliche Auswahl an Studiomonitoren für HDR. Der Dolby Pulsar und der Dolby PRM-32FHD Monitor sind beide HDR tauglich, jedoch nicht käuflich erwerblich sondern können lediglich gemietet werden. Die einzigen Monitore, die häufig für HDR-Grading genutzt werden und käuflich zu erwerben sind, sind der Sony BVM-X300, Canon DP-V2420 und der Dolby

PRM-4220. Allerdings liegen diese Referenz Monitore im Preissegment von 30.000 Euro.

Tabelle 5.1-1: HDR Referenz Monitore im Überblick

Monitor	Auflösung	Farbraum	Helligkeit	Panel
Sony BVM-X300	4K	BT.2020	1000 Nits	10-Bit OLED
Canon DP-V2420	4K	BT.2020	1200 Nits	10-Bit IPS LCD
Dolby PRM-4220	2K	DCI-P3	600 Nits	12-Bit
FSI XM310K	4K	P3	3000 Nits	10-Bit OLED

5.1.2 Consumer HDR-TVs

Eine alternative zu den teuren Referenz Displays können Consumer TVs sein, die nur einen Bruchteil derer kosten. Vor allem die OLED-Serie von LG (C8 /C9) wird hierfür gerne genutzt, da diese, korrekt kalibriert, ein gutes Bild abgeben und auf eine Spitzenhelligkeit von 1000 Nits kommen. Das Bild ist nicht zu 100% referenzgetreu und diese Monitore haben auch keine SDI-Inputs. Doch mit Resolve ist es möglich, über HDMI die korrekten Flags mitzuliefern, sodass der TV in den HDR-Modus umschaltet. Der TV wird per HDMI direkt an eine DeckLink oder UltraStudio Karte angeschlossen, danach wird das HDMI-Tunneling in Resolve aktiviert.

5.1.3 SDR-Referenzmonitor

Für die SDR-Trim braucht es auch einen SDR-Referenzmonitor, der über SDI und einer SDI-Wiedergabekarte mit der Workstation verbunden ist. Hier eignen sich zum Beispiel die Monitore von Flanders Scientific, wie der FSI DM 250 OLED.

5.1.4 Wiedergabekarten

Um eine direkte Ausgabe auf den Monitor aus Resolve zu bekommen, und um das Color Management des Betriebssystems zu umgehen, wird eine professionelle Wiedergabekarte benötigt. Häufig verwendet sind die Karten von Black Magic. Da diese Firma auch die Videosoftware Resolve vertreibt, kann von einem reibungslosen Miteinander von Gra-

ding Software und Karte ausgehen. Das günstigste Modell ist die „DeckLink Mini Monitor 4K“ Karte. Diese kann über HDMI 2.0a HDR Metadaten aus Resolve mit ausgeben. Diese Karte wurde auch in diesem Versuchsaufbau verwendet.

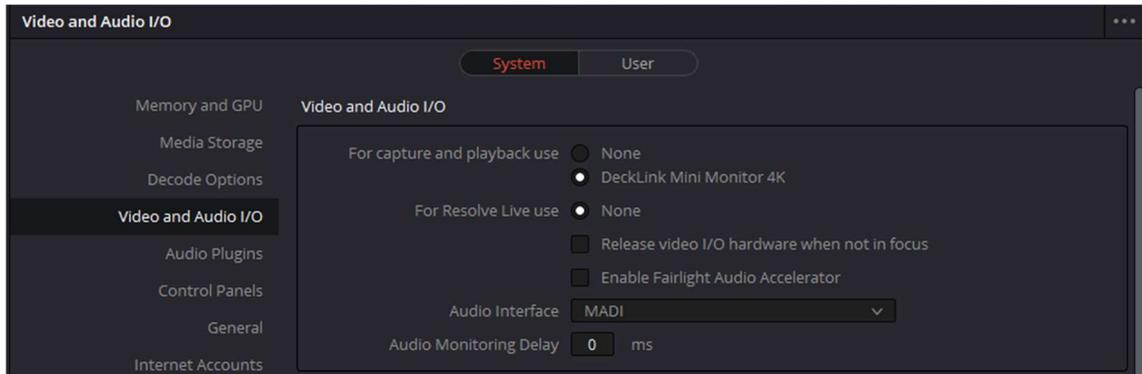


Abbildung 5-1: Aktivierung der DeckLink Karte in Resolve unter Preferences > Video and Audio I/O > System

5.1.5 Dolby Content Mapping Unit (CMU) und eCMU

Die Content Mapping Unit ist mit den entsprechenden Metadaten aus der Grading Software, für die Konvertierung von HDR zu SDR verantwortlich. Hierzu gibt es zwei Möglichkeiten:

- Externe Hardware CMU (eCMU), die von Dolby bezogen werden muss
- Seit CMv4 gibt es ein in die Grading Software integrierte iCMU

5.1.6 Dolby Vision Lizenz

Um die Trim-Controls in Resolve zu aktivieren, benötigt es eine Lizenz von Dolby, die in den Einstellungen (*File > Dolby Vision*) geladen werden kann. Die Dolby Vision Lizenz besteht aus zwei Teilen, der Lizenz und einer Konfigurationsdatei.

 cm_config_v1.3.bin	28.05.2019 19:23	BIN-Datei	53 KB
 HDM_Stuttgart_iCMU_EXP_02292020_temp_license.bin	10.02.2020 17:04	BIN-Datei	1 KB
 READ ME FIRST!! License Key Upload Instructions_rev6	04.08.2019 22:04	Adobe Acrobat D...	38 KB
 Using Resolve v15 with the external CMU	16.08.2018 12:17	Office Open XML-...	579 KB

Damit wären die technischen Bedingungen für ein HDR Dolby Vision Mastering erfüllt. Wie der Color Grading Prozess von HDR aussieht und was SDR-Trims sind, wird im nächsten Kapitel detaillierter erklärt.

5.2 HDR-Grading in Resolve

Color Grading in HDR unterscheidet sich nicht maßgeblich von dem Grading für SDR. Es sind die gleichen Kontrollelemente wie Lift, Gamma und Gain zu verwenden. Der große Unterschied zu SDR ist der größere zusätzliche Headroom für Lichter, bei denen SDR ab 100 Nits geclippt hat. Zudem sind deutlich brillantere Farben möglich, auch in Bereichen weit über 100 Nits, bei dem in SDR alles nur noch Weiß war.

5.2.1 HDR Projekt anlegen

Die Grundeinstellungen eines HDR Projekts unterscheiden sich nicht groß von dem eines SDR Projekts. Dazu sollten zuerst einmal alle Grundeinstellungen vorgenommen werden wie Auflösung, Frame Rate und danach die Einstellungen für HDR abgeändert werden.

Project Settings > Master Project Settings > Timeline Format > Color science > DaVinci YRGB Color Managed

Als erstes sollte der Farbraum des Projekts eingestellt werden. Dazu sollte in den Master Project Settings die Color Science auf DaVinci YRGB Color Managed gestellt werden. Auch ein ACES Workflow würde funktionieren, jedoch hat ACES immer seine Eigenheiten die nicht jedem liegen. Von Netflix wird ACEScct mit einer P3-D65 ST2084 ODT (Output Device Transform) empfohlen.

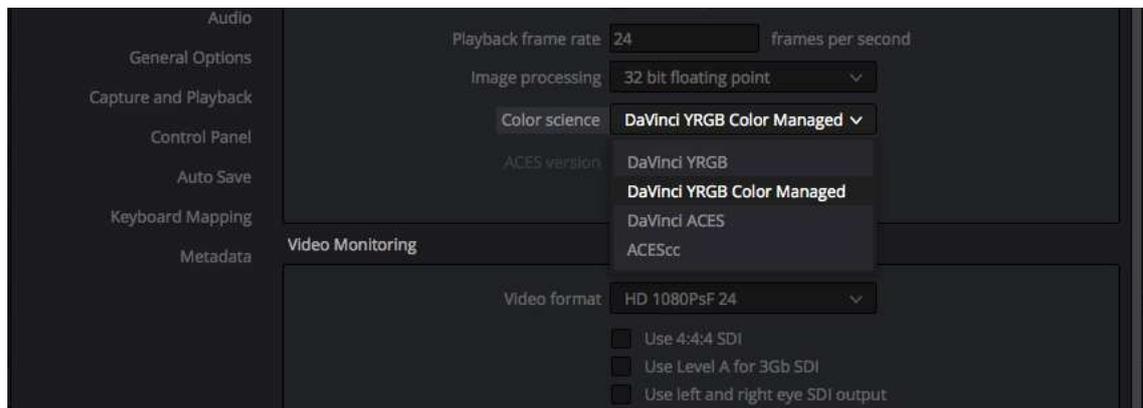


Abbildung 5-2: Einstellung der Color Science in DaVinci Resolve

Project Settings > Color Management > Color Management Settings

Ist Resolves Color Management eingeschaltet, folgt das setzen der Farbräume für Input, Timeline und Output. Zunächst sollte *Use Seperate „Color Space and Gamma“* aktiviert sein, so dass diese getrennt voneinander einstellbar sind. Hierbei empfiehlt es sich, Input Color Space auf Bypass zu setzen, da das RAW-Material durch Resolve debayerd wird

und dadurch das Setzen eines Farbraums unnötig ist. Falls das Material nicht RAW ist kann hier die entsprechende Einstellung gemacht werden. Als Timeline Farbraum entweder Rec.2020 oder P3D65 wählen und als Gamma ST.2084. Für den Ausgabefarbraum die gleichen Einstellungen wählen. Unter „*HDR Mastering is for*“ werden die maximalen Nits auf dem gemastert wird eingetragen. Für HDR10 zum Beispiel 1000 und für Dolby Vision 1000 oder 4000 Nits. Dieser Wert sollte mit dem Wert des tatsächlich benutzten Displays übereinstimmen.

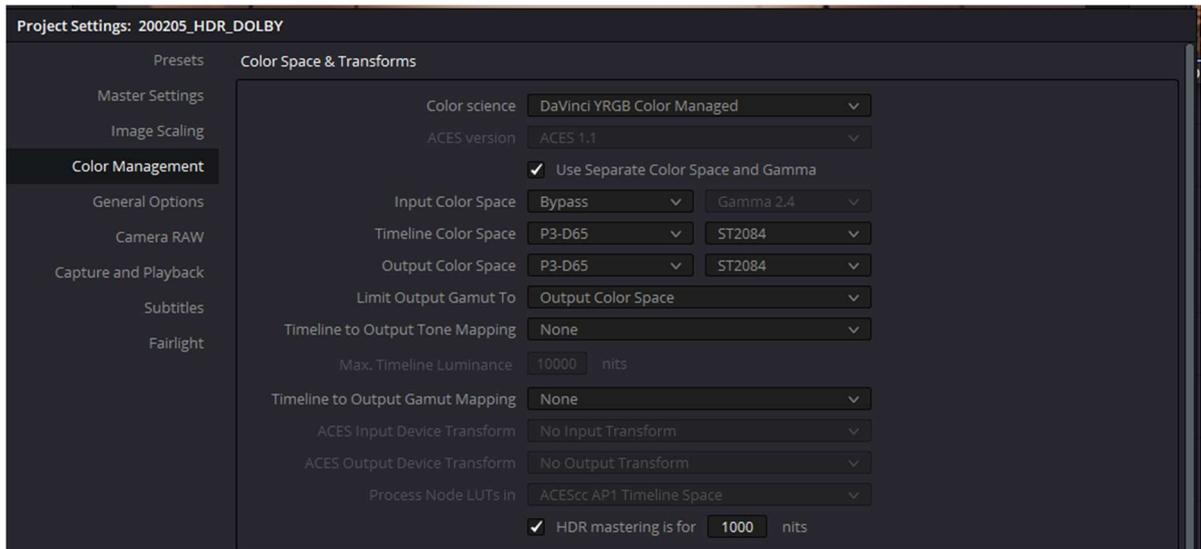


Abbildung 5-3: Einstellung der Farbräume in DaVinci Resolve

Preferences > Color > User > Enable HDR Scopes for ST.2084

Nach dem Festlegen der Farbräume und EOTFs sind die HDR Scopes in Resolve zu aktivieren. Dies ändert die Scope-Anzeige in Resolve von 10-Bit Digitalwerten zu einer logarithmischen Helligkeitsskala, die die Helligkeit in Nits anzeigt.

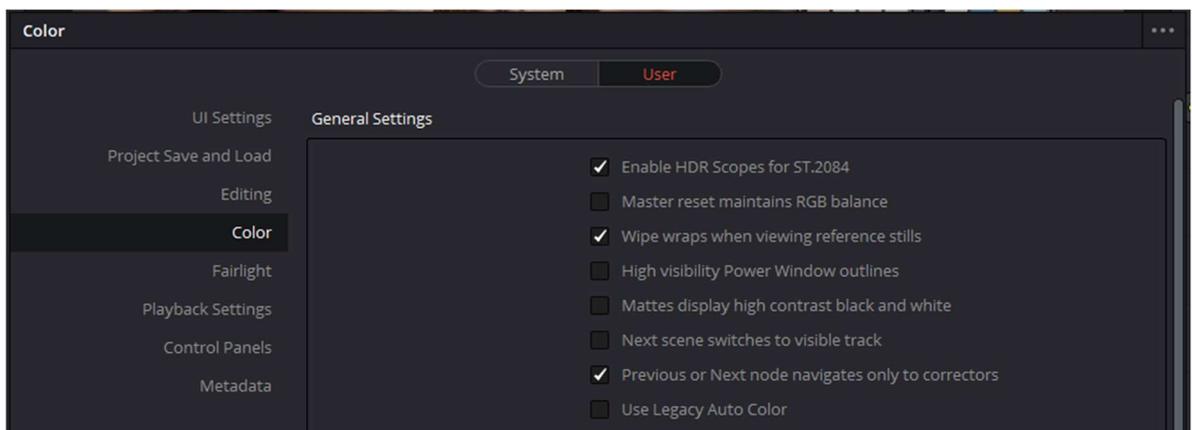


Abbildung 5-4: Einstellung um die Scopes in Resolve in Nits darzustellen

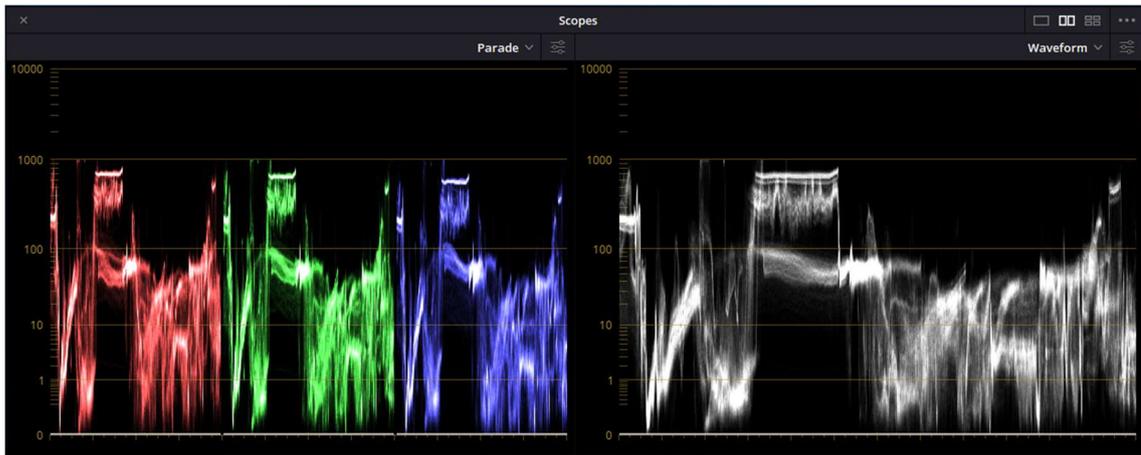


Abbildung 5-5: HDR-Scopes in DaVinci Resolve 16 in Nits

Project Settings > Master Project Settings > Video Monitoring > Enable HDR metadata over HDMI

Ist der Grading Monitor über HDMI mit der Workstation verbunden, müssen die Metadaten über HDMI-Ausgabe aktiviert werden. Andernfalls werden diese nicht mit ausgegeben und der Monitor schaltet nicht in den HDR Modus. Bei den meisten Geräten kann dieser nicht manuell eingeschaltet werden kann. Ist der Monitor über SDI verbunden, werden Farbraum und EOTF entsprechend der Einstellungen in Resolve am Monitor übernommen. Weitere Einstellungen sind für das Grading in HDR vorerst nicht nötig.

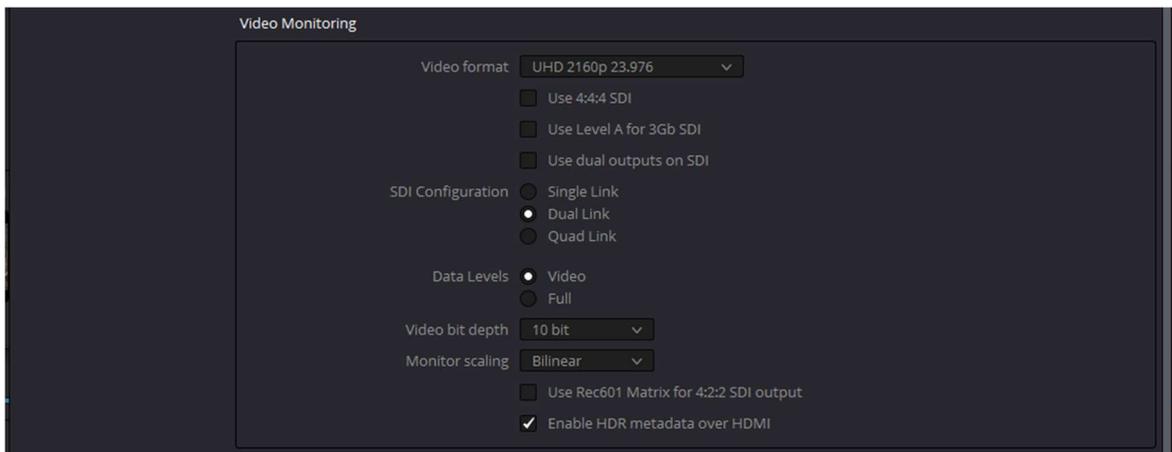


Abbildung 5-6: Aktivieren der Metadaten über HDMI

5.2.2 Anpassung des Helligkeitsbereichs / HDR Mode

Intern arbeitet Resolve mit linearen digitalen Werten, die eine leichte Optimierung für Gamma 2.4 aufweisen. Da die EOTF PQ ST.2084 jedoch stark von einer Linearität abweicht. Wird dem nicht entgegen gewirkt, verschieben sich die Helligkeitswerte schnell in den oberen Bereich der 10.000 Nits und dies wiederum erschwert das Grading. Die eine Möglichkeit diesen Effekt zu vermeiden, ist ein Timeline-Node zu erstellen, die für die gesamte Timeline nach den Korrekturen angewandt wird. In diesem Node werden mithilfe von Curves die Schwarzwerte ausgedehnt und die Lichter komprimiert. Dadurch wirkt das Verhalten von Resolve im Grading der Lichter und Schatten einheitlicher.

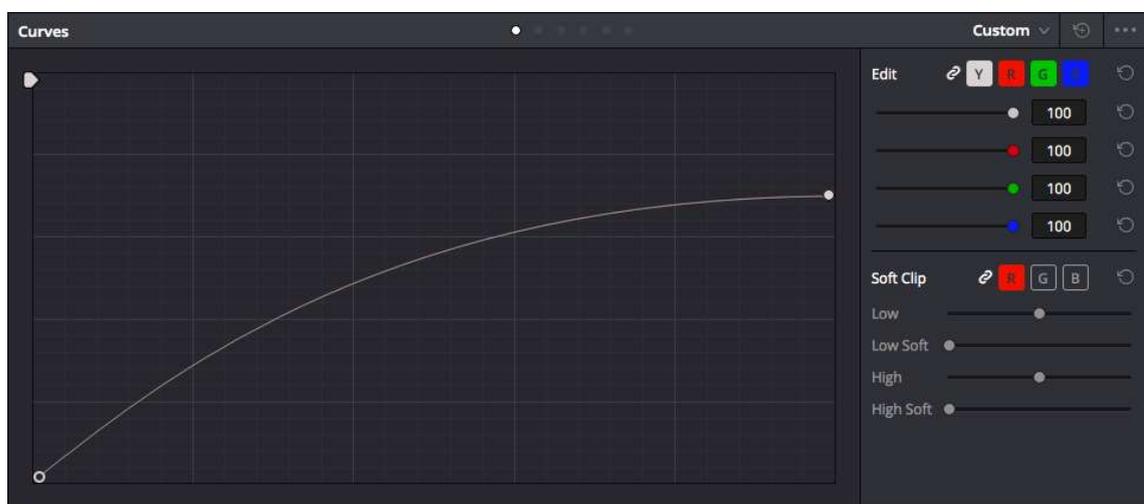


Abbildung 5-7: Anpassung durch eine modifizierte Kurve

Resolve bietet eine ähnliche Anpassung mit dem *HDR Mode* an, der über Rechtsklick auf die Node aktiviert werden kann. Doch dies hat einen seltsamen Effekt auf das Verhalten von Lift, Gamma und Gain. Im HDR Mode ist eine präzise Einstellung der Schatten und Schwarzwerte kaum möglich, da diese schnell nach oben oder unten rasen.

5.2.3 Grading und MaxFALL

MaxFALL bezeichnet den maximalen Helligkeitswert eines Frames, wobei der Helligkeitsdurchschnitt aller Pixel in diesem Frame errechnet wird. Im Grading kann dies zu Problemen führen, denn wenn der für den Monitor spezifische MaxFALL überschritten wird, regelt der Monitor automatisch herunter. Dann kann es passieren, dass das Bild heller gradet werden sollte, es aber nicht heller wird, da der Monitor dagegen steuert. Auch bei dem LG Monitor der für diesen Versuchsaufbau verwendet wurde machte sich dieser Effekt bemerkbar. Professionelle Monitore haben dazu eine extra LED Anzeige

die anzeigt ob der MaxFALL überschritten wurde oder nicht. Der Monitor Regelt ab einer bestimmten Helligkeit herunter. Deshalb muss ein großer Bereich des Frames um die 100 Nits maximal hell sein und nur ein kleiner Teil davon darf diesen Wert übersteigen, so dass der MaxFALL nicht überschritten wird. Gerade bei hellen Shots muss oft das ganze Bild dunkler gradated werden, um den MaxFALL nicht zu überschreiten. Auch bei Shotfolgen, die gleich verortet sind, kann durch MaxFALL eine Uneinheitlichkeit entstehen, wenn z.B. durch Close-Ups von der besonders hellen Stelle der MaxFall überschritten wird und dadurch Helligkeit zurück-genommen werden muss. Um eine Gleichheit in der Helligkeit über alle Shots hinweg zu erhalten, sollten sich die weiteren Shots nach dem Close-Up Shot richten.

5.2.4 HDR Schatten und Lichter

Im Bereich der Schatten und Lichter liegt vielleicht der größte Unterschied zu SDR. Denn in HDR gibt es kein komplettes Schwarz oder Weiß, sondern eine Bandbreite von Schwarz und Weiß. In SDR wurden ohne Probleme die Schatten soweit runter gezogen, dass diese sich unterhalb der Legal Range befanden und dadurch waren diese pures Schwarz, das gleiche mit Weiß. Deshalb sollten die Maximalwerte des Monitors als Orientierung dienen und dort den Weißpunkt bzw. Schwarzpunkt festlegen. Dadurch behält das gesamte Video einen einheitlichen Look. Bei der HDR-zu-SDR Umrechnung sollte beachtet werden, dass die Stellen, die in HDR Weiß aussahen, in SDR nicht mehr Weiß wirken können. Diesem Effekt kann mit den Primary- oder Secondary-Trimtools im Trim-Pass entgegen gewirkt werden. Am unkompliziertesten lassen sich HDR-Weiß und –Schwarz mit *Low Range* und *High Range* in den Log Farbrädern anpassen. Somit kann die Range genau bestimmt werden ab der die Highlight- und Shadow-Farbräder greifen und somit bestimmte Helligkeitsbereiche im Bild exakt angesteuert werden.



Abbildung 5-8: Festlegung der Weiß- und Schwarzwerte durch Low Range und High Range mit den Log Farbrädern

5.2.5 Clipping

Ein Punkt, der beim HDR-Grading noch beachtet werden sollte, betrifft erneut das Weiß bzw. die maximal hellsten Stellen im Video. Wird nach Auge gegraded und nicht nach den Scopes, kann es schnell passieren, dass es Bereiche in Shots gibt, die die Helligkeit des Grading Monitors übersteigen. Dieser Clipped dann das Signal und für den ersten Moment fällt dies auch nicht negativ auf. Doch wird auf einem anderen Display, das eine höhere Helligkeit zulässt, das gegradede Material angeschaut, erscheint das Weiß vielleicht gar nicht mehr Weiß. Das bedeutet, das Master, das wir so gegraded haben, enthält Informationen, die wir nicht auf unserem Monitor sehen konnten. Deshalb ist ein Clipping des Signals auf die maximale Helligkeit des benutzten Grading Monitors durchaus praktikabel. Denn dadurch bleibt der Creative Intent erhalten, denn das Signal ist dann geclipped und enthält keine Informationen mehr. Resolve bietet dazu die Möglichkeit eine Soft Clip LUT (Look Up Table) zu erstellen, die dann beim Export angehängt wird. Diese findet sich in den Einstellungen (*Project Settings > Color Management > Generate Soft Clip LUT*) dort ist das maximale Video Level in Codevalues einzutragen. So entspricht der Wert 767 in einem 10-Bit Video in ST.2084 1000 Nits.

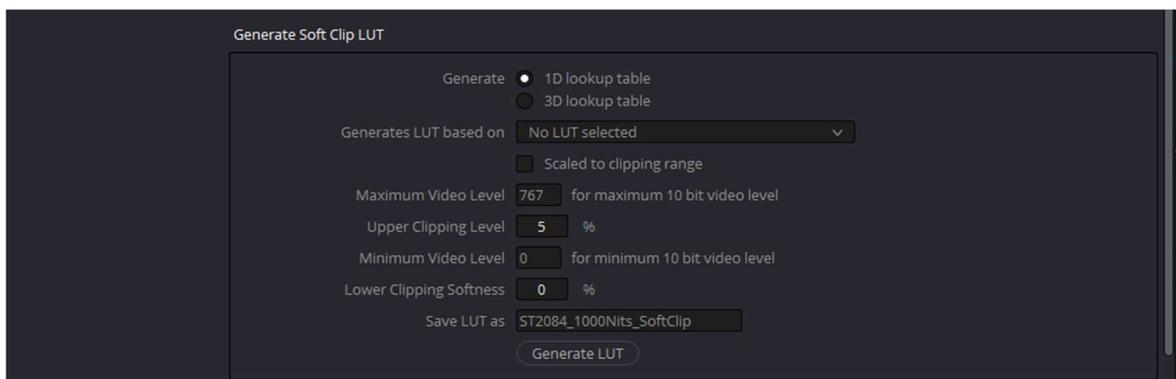


Abbildung 5-9: Soft-Clipping LUT erstellen in Resolve

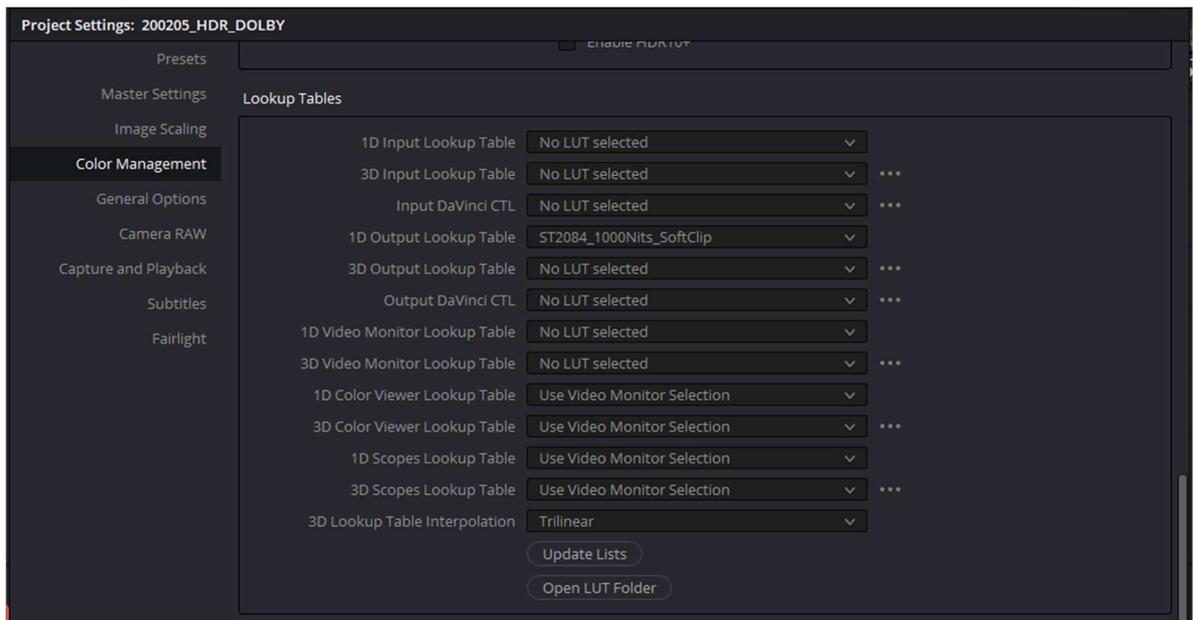


Abbildung 5-10: Aktivieren der Sof-Clipping LUT in Resolve

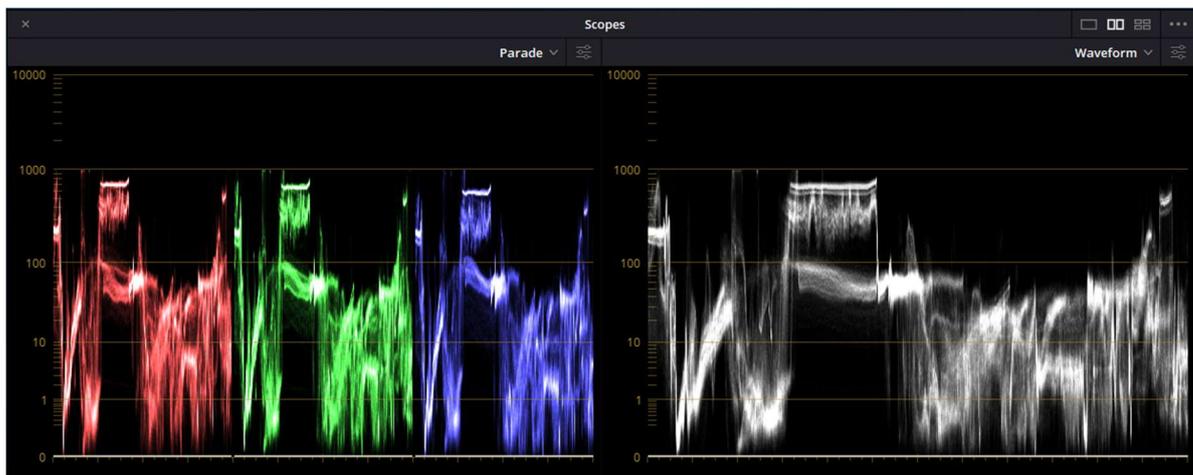


Abbildung 5-11: Durch die Soft Clipping LUT ist ein graden über 1000 Nits nicht möglich. Außerdem besitzt die LUT ein weiches Clipping

5.3 SDR-Trim Pass

Der SDR Trim Pass ist der Prozess, aus einem HDR Grade der für einen bestimmten Helligkeitswert gemastert wurde, durch Algorithmen (L1 Metadaten) und Anpassungen des Coloristen, Versionen für niedrigere Helligkeiten zu erstellen. Zum Beispiel wurde auf 1000 Nit gemastert anschließend wird Shot für Shot Anpassungen für ein 100 Nit Level gemacht. Dadurch kann im Nachhinein eine SDR-Version herausgerechnet werden. Zudem kann der TV des Konsumenten, der nicht die Helligkeitswerte des Grading Monitors erreicht, mithilfe dieser SDR-Trim Metadaten das Bild dementsprechend tonemappen. Ein 100 Nit Trim Pass ist für ein Dolby Vision Mastering pflicht. Es sollten zumindest alle Shots einmal analysiert werden.

Der SDR Trim Pass lässt sich in 3 Schritten gliedern:

- (1) Erstellen des HDR Grades
- (2) Analyse aller Shots durch Dolby Vision Algorithmus (L1 Metadaten)
- (3) Anpassung des Trims durch Primary und Secondary Trim Controls

5.3.1 L1-Metadaten-Analyse

Die L1-Metadaten Analyse liegt einem Algorithmus von Dolby zugrunde. Im Dolby Vision Fenster in Resolve wird hierzu zu aller erst als Target Display Output „100-nit, BT.709, BT.1886, Full“ festgelegt um damit ein Trim für den SDR Standard mit Rec.709 Farbraum und einem Gamma 2.4 zu erstellen. Die Option „Enable Tone Mapping Review“ aktiviert die Ausgabe des berechneten Bildes. Es sollten zuerst alle Shots einmal durch den Algorithmus analysiert werden, da es ansonsten beim Rendering zu Abbrüchen kommen kann. Dieser Schritt des Trim Passes ist vollautomatisch und erzeugt beträchtlich gute Ergebnisse. Bei über der Hälfte der Shots war das Ergebnis zufriedenstellend.

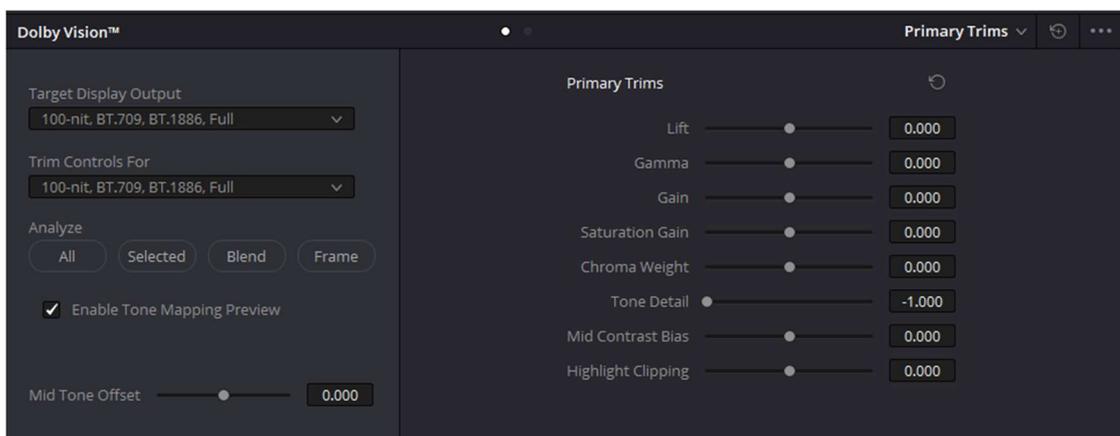


Abbildung 5-12: Globale Trim Controls in Davinci Resolve [34]

Mit der Decklink 4K Mini Monitor Karte ist eine gleichzeitige Ausgabe von SDR und HDR nicht möglich und es ist nötig „*Target Display Output*“ immer hin und her zu schalten zwischen 1000 und 100 Nits um die Versionen miteinander zu vergleichen. Bei Umschaltung auf 100 Nits gibt der HDR Monitor eine verfälschte Ausgabe wieder, wenn die HDR Metadaten noch über HDMI mit ausgegeben werden. Eine Deaktivierung dieser Option schafft Abhilfe. Mit teureren Karten und der eCMU kann ein HDR Bild und gleichzeitig ein SDR Bild ausgegeben werden. Damit ist ein direkter Vergleich beider Versionen möglich und die Anpassung der SDR Version um einiges leichter. Zu beachten ist dass lediglich der erste Video Track analysiert wird. Deswegen sollten, der Einfachheit halber, vor dem Grading die Videospuren auf den ersten Video Track normalisiert und Multi-Track Spuren vermieden werden. Es ist möglich mehrer Video Tracks zu verwenden, dabei sollte jedoch darauf geachtet werden dass diese sich nicht überlagern.

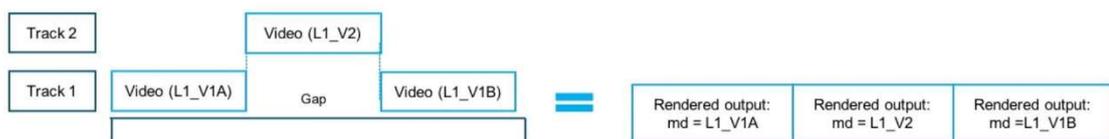


Abbildung 5-13: Multi-Track Workflow [34]

Außerdem muss das Output Blanking for der Analyse gesetzt werden. Dadurch berechnet die Analyse lediglich den aktiven Bildinhalt. Wenn diese Einstellung nicht Korrekt gesetzt sind zieht die Analyse die schwarzen Letterboxen mit in die Berechnung ein.

Set the Global Output Blanking under **TIMELINE** → **OUTPUT BLANKING**

Set per-clip Output Blanking under **COLOR TAB** → **SIZING** → **OUTPUT SIZING** → **BLANKING** → **SHOW BLANKING CLIP OVERRIDE** → **CLIP**

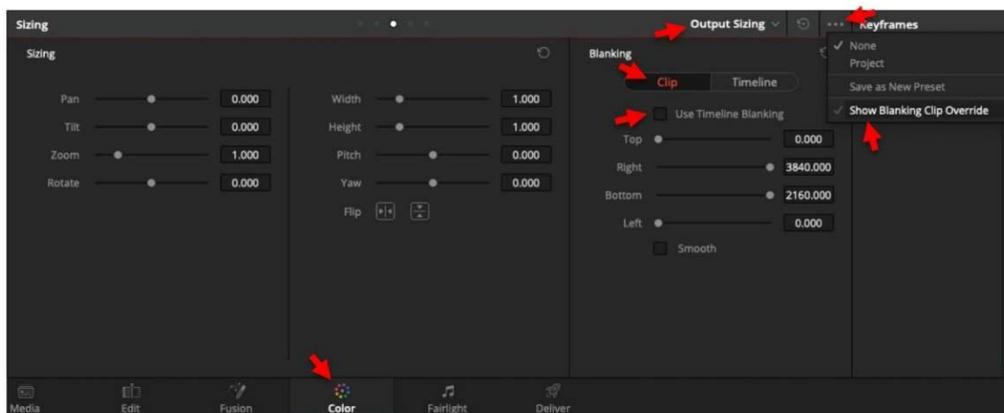


Abbildung 5-14: Resolve Output Blanking Einstellungen [34]

5.3.2 L3 Mid Tone Offset

Der Mid Tone Offset wirkt sich auf Shotbasis aus und dient der Anpassung der Belichtung. Die L1-Metadaten Analyse bietet nicht immer auf Anhieb ein zufriedenstellendes Bild, so ist in Abbildung 5-15 gerade der Protagonist viel zu hell. Ebenfalls der Kontrast zu dunklem Raum und hellem Außenbereich würde verloren gehen.



Abbildung 5-15: Korrektur (rechts) durch Mid Tone Offset

5.3.3 Primary Trim Controls

Die Primary Trim Controls bieten eine bekannte Lift, Gamma und Gain Korrektur sowie verschiedene Einstellungsmöglichkeiten zu Helligkeit und Farbigkeit.

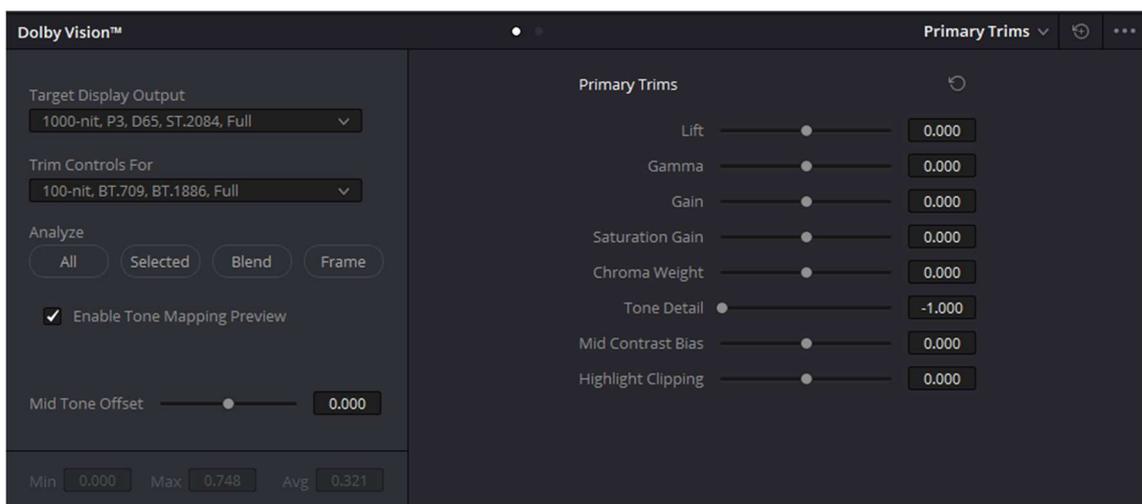


Abbildung 5-16: Dolby Vision Primary Trim Controls in Resolve

Lift/Gamma/Gain

Seit CMv4 verhalten sich diese Regler ähnlich den traditionellem Lift, Gamma und Gain und regeln die Helligkeit von Schatten, Mitteltöne und Lichter. So können nach der Shot Analyse viel zu flache Bilder entstehen die mit Lift, Gamma und Gain wieder angepasst werden können.



Abbildung 5-17: Beispiel für eine Korrektur (rechts) mit Lift / Gamma / Gain eines zu flachen Bildes (links)

Saturation Gain

Saturation Gain ist für die Anpassung der Sättigung des gesamten Bildes verantwortlich und wirkt sich auf alle Farben aus. Oft sind Farben viel zu gesättigt und können mit Saturation Gain korrigiert werden. Um einzelne Farbenkanäle direkt anzusteuern, können die Secondary Trim Controls verwendet werden.



Abbildung 5-18: Beispiel für eine Korrektur (rechts) mit Saturation Gain

Chroma Weight

Besonders beim Mapping von HDR auf SDR können in den Lichtern Details verloren gehen. Mit der Chroma Weight Einstellung lässt sich von maximaler Sättigung und minimaler Helligkeit zu minimaler Sättigung und maximaler Helligkeit regeln. Vor allem bei starken Anpassungen beim 100 Nits Trim Pass ist es empfehlenswert, das Mapping bei anderen Nit-Levels, zum Beispiel 600 Nits, zu kontrollieren.

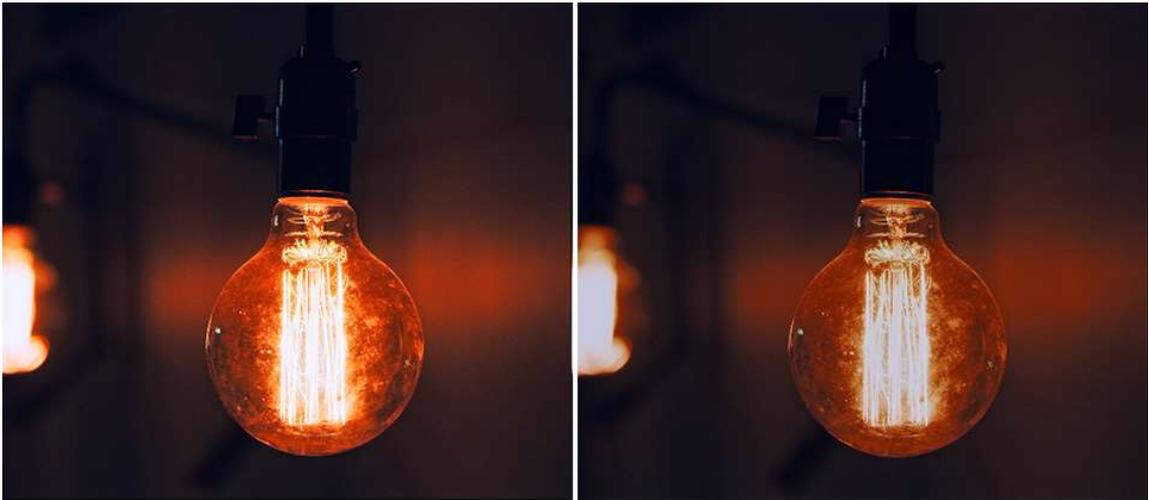


Abbildung 5-19: Beispiel für eine Korrektur (rechts) mit Chroma Weight

Tone Detail

Bei der Konvertierung von HDR zu SDR können Schärfe und Details in den Lichtern verloren gehen. Mit Tone Detail lässt sich diese wiederherstellen. Oft wirkt das Bild um helle Stellen herum unscharf, diesem Effekt lässt sich mit Tone Detail entgegenwirken.

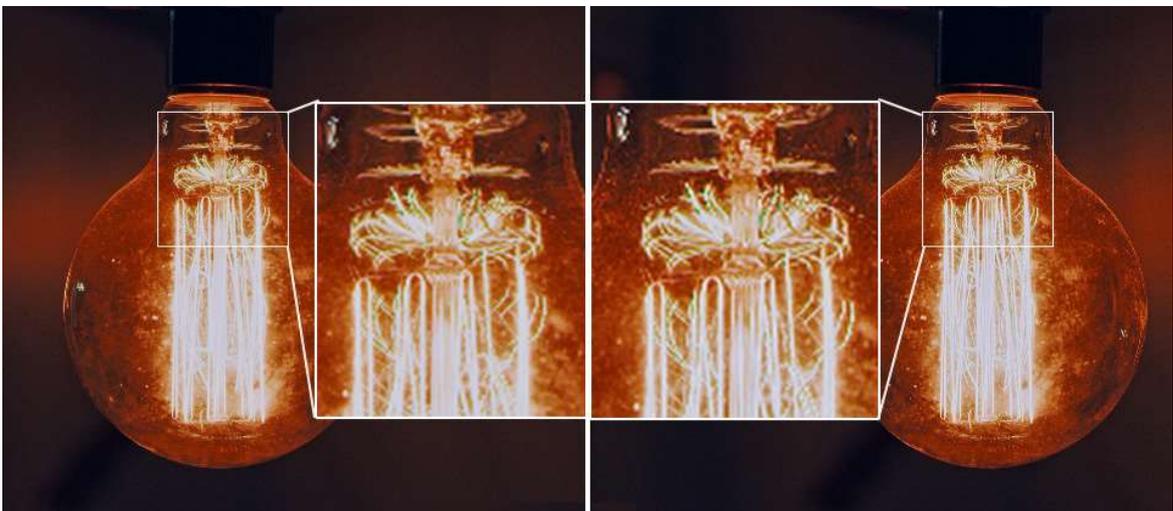


Abbildung 5-20: Beispiel für eine Korrektur (rechts) mit Tone Detail

Mid Contrast Bias

Diese Einstellung regelt den Kontrast in den Mitteltönen. Zusammen mit den Lift und Gain Reglern können präzise Anpassungen gemacht werden. Diese Einstellung ähnelt dem Pivot in Resolve.



Abbildung 5-21: Beispiel für eine Korrektur (rechts) mit Mid Contrast Bias

Highlight Clipping

Gerade in den Highlights können sich Details abbilden, die nicht gewünscht sind. Zudem können die Highlight Levels angepasst werden. Diese Anpassung kann sich jedoch auch auf die oberen Mitteltöne auswirken – dem jedoch mit Gamma und Gain entgegengewirkt werden kann.

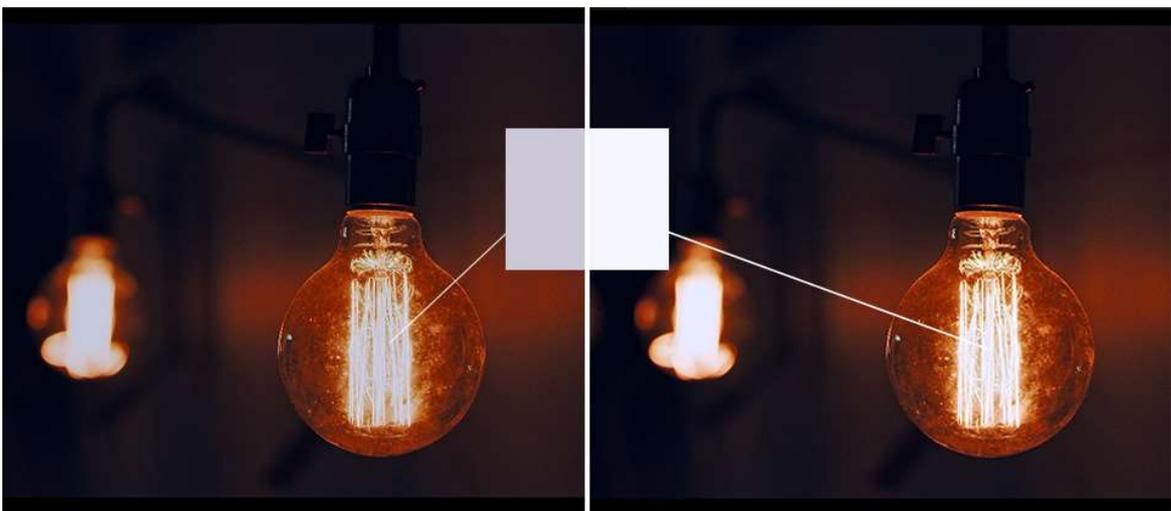


Abbildung 5-22: Beispiel für eine Korrektur mit Highlight Clipping

5.3.4 Secondary Trim Controls

Die Secondary Trim Controls bieten genauere Einstellmöglichkeiten für Sättigung und Farbton. Diese Einstellungen können für die sechs Farbkanäle einzeln vorgenommen werden.

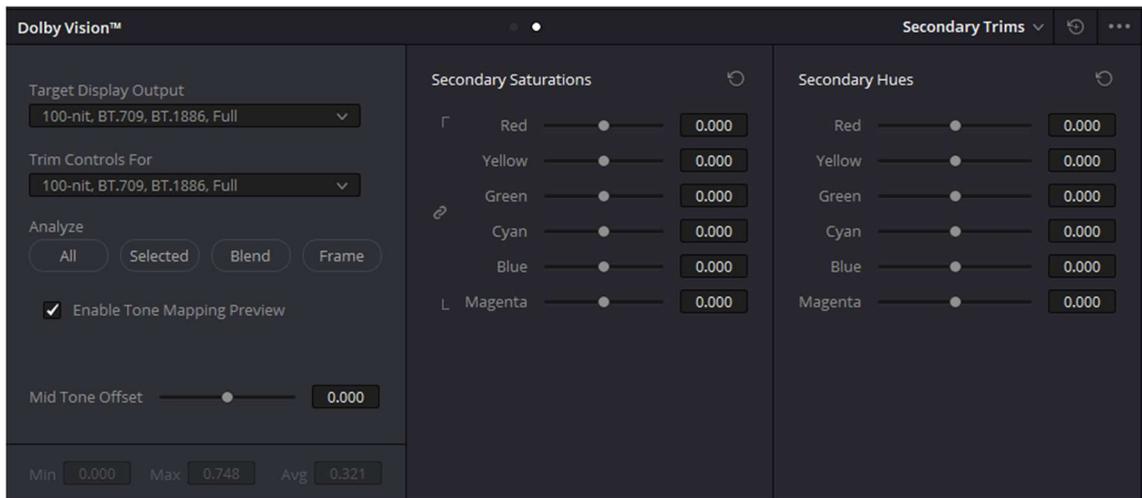


Abbildung 5-23: Dolby Vision Secondary Trim Controls [34]

Saturation Trim Controls

Mit den Saturation Trim Controls hat der Colorist die Möglichkeit, die einzelnen Farbkanäle wie Rot, Gelb, Grün, Cyan, Blau und Magenta anzusteuern und deren Sättigung entweder einzeln oder alle auf einmal anzupassen.

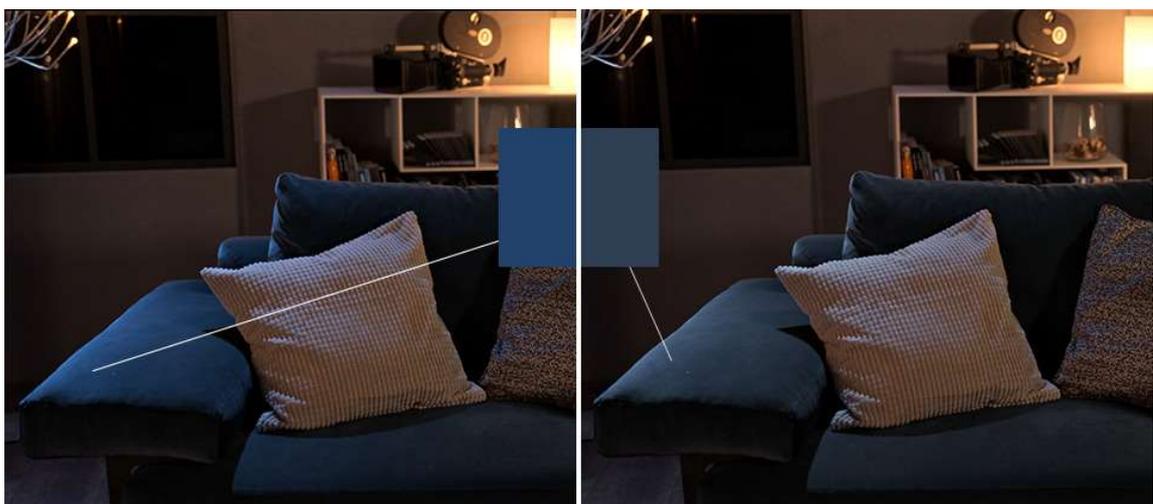


Abbildung 5-24: Beispiel für eine Blau Korrektur (rechts) mit Saturation Trim Controls

Gerade durch das Tone Mapping in einen kleineren Farbraum können viel zu intensive Farben in der SDR Variante entstehen, diese können mit den Saturation Secondary Trim Controls gezielt angesteuert werden.



Abbildung 5-25: Korrektur (rechts) von zu saturierten Rot (links) nach L1 Analyse

Hue Trim Controls

Mit den Hue Trim Controls lässt sich der Farbton jedes einzelnen Farbkanals durch den Coloristen anpassen.

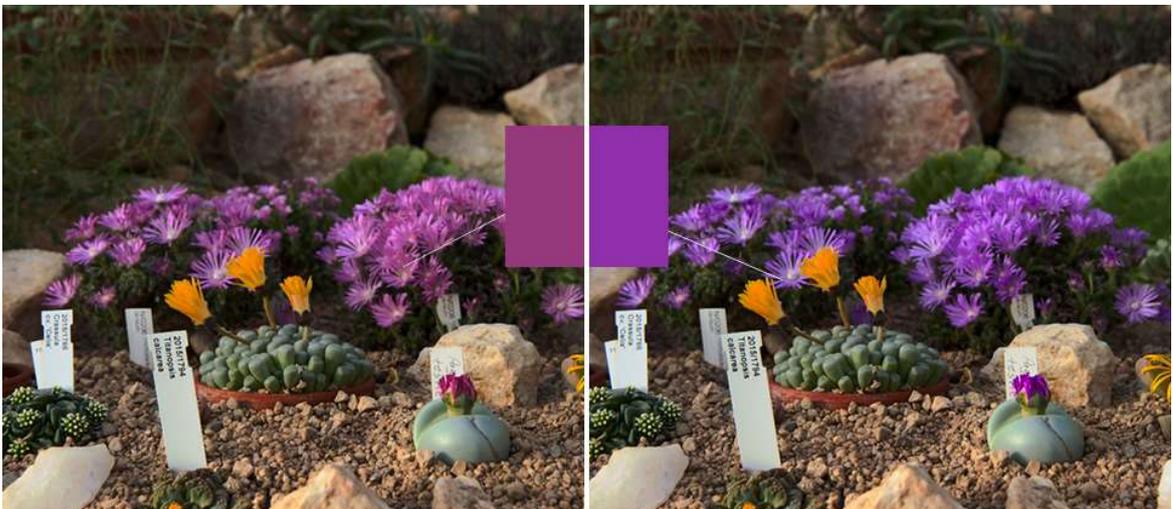


Abbildung 5-26: Beispiel für eine Korrektur (rechts) von falschem Magenta Farbton (links) mit Hue Trim Controls

5.4 Export

Beim Export von Dolby Vision Inhalten gibt es verschiedene Export Varianten:

- (1) Dolby Vision Master: Ein Masterfile der Bildinformationen und des Grades
- (2) Dolby Vision IMF: Deliverable File für Netflix und co.
- (3) Dolby Vision Mezzanine: Eine .mxf Datei die die Metadaten eingebacken hat

5.4.1 Dolby Vision Master

Das Dolby Vision Master, auch Video Display Master (VDM), ist die höchste Bildqualität die für weitere Ausspielungen genutzt werden kann. Dolby empfiehlt eine 16-bit TIFF oder OpenEXR Sequenz mit dazugehöriger Dolby Vision XML Datei [34]:

- Dolby Vision Master:
 - 16-bit TIFF Sequenz
 - ProRes 4444XQ .mov für iTunes
 - ProRes 4444 .mov
- XML Dolby Vision Metadaten

Die XML Metadaten können in Resolve im Editing Panel exportiert werden.

Editing > Export XML, AAF...

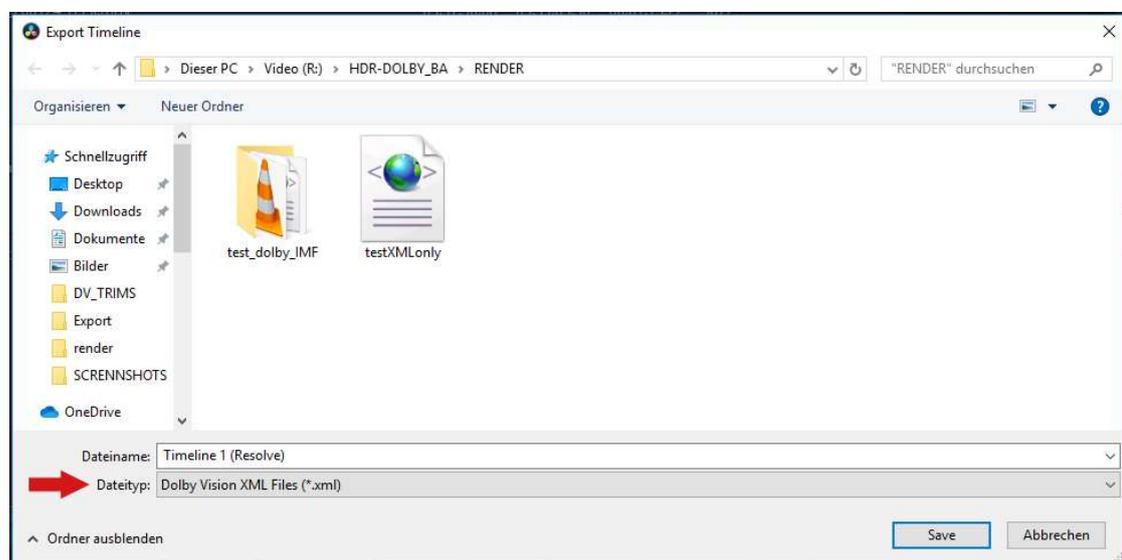


Abbildung 5-27: Dolby Vision XML export findet sich unter Dateityp

5.4.2 Dolby Vision XML

Die Dolby Vision XML enthält alle nötigen Metadaten des Trim Passes. Und ist wie folgt gegliedert:

- (1) Mastering Display: Informationen über das Display auf dem gemastert worden ist
- (2) Target Display: Informationen der Trim Ziele (z.B. 100 Nits, 600 Nits etc.)
- (3) Shot: Enthält die Trimeinstellungen des Shots

Innerhalb des Mastering Display Tags werden die RGB Primaries, Weißpunkt, Spitzenhelligkeit und Minimalhelligkeit des Grading Displays als Fließkommazahlen festgelegt. Anderst als in der CPL des IMFs.

```
- <MasteringDisplay>
  <ID>20</ID>
  <Name>1000-nit, P3, D65, ST.2084, Full</Name>
  - <Primaries>
    <Red>0.68 0.32</Red>
    <Green>0.265 0.69</Green>
    <Blue>0.15 0.06</Blue>
  </Primaries>
  <WhitePoint>0.3127 0.329</WhitePoint>
  <PeakBrightness>1000</PeakBrightness>
  <MinimumBrightness>0.0001</MinimumBrightness>
  <DiagonalSize>42</DiagonalSize>
</MasteringDisplay>
```

Abbildung 5-28: Dolby Vision XML Auszug zu Mastering Display

Innerhalb des Target Display Tags sind die gleichen Displayinformationen gespeichert aber für das Ziel-Display für das Trims angelegt worden sind. Wenn noch andere Trim Pässe gemacht worden sind werden diese unter einer anderen ID auch angegeben.

```
- <TargetDisplay>
  <ID>1</ID>
  <Name>100-nit, BT.709, BT.1886, Full</Name>
  - <Primaries>
    <Red>0.64 0.33</Red>
    <Green>0.3 0.6</Green>
    <Blue>0.15 0.06</Blue>
  </Primaries>
  <WhitePoint>0.3127 0.329</WhitePoint>
  <PeakBrightness>100</PeakBrightness>
  <MinimumBrightness>0.005</MinimumBrightness>
  <EOTF>gamma_bt1886</EOTF>
  <DiagonalSize>42</DiagonalSize>
</TargetDisplay>
```

Abbildung 5-29: Dolby Vision XML Auszug zu Target Display

Ein Shot wird immer mit einer eindeutigen ID (<UniqueID>) versehen und enthält die Nummer des Startframes (<In>) und der die Länge in Frames (<Duration>). Innerhalb des DVDynamicData Tag befinden sich die Dolby Vision Metadaten die in Kapitel 4.2 ausführlich erläutert wurden. Das Tag <TID> bestimmt für welches Trim Ziel die Einstellungen gelten. Die Secondary Trim Controls sind als Vektor abgespeichert (SaturationVectorField, HueVectorField).

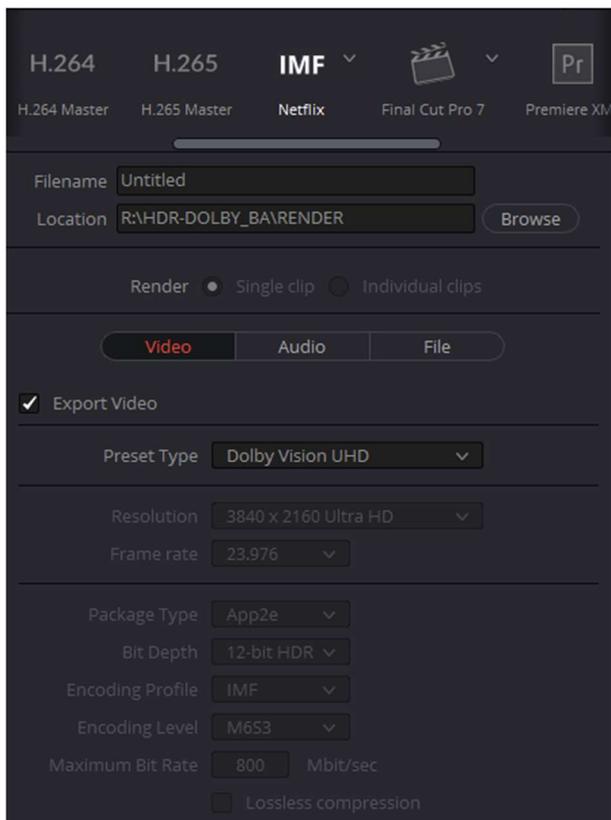
```
- <Shot>
  <UniqueID>c3130655-2d20-424e-91b4-cb58c0f20b4a</UniqueID>
  - <Record>
    <In>86400</In>
    <Duration>165</Duration>
  </Record>
  - <PluginNode>
    - <DVDynamicData>
      - <Level1 level="1">
        <ImageCharacter>0 0.275851 0.74902</ImageCharacter>
      </Level1>
      - <Level2 level="2">
        <TID>1</TID>
        <Trim>0 0 0 0.000483015 -0.035372 0 -0.191 0.177 -0.574</Trim>
      </Level2>
      - <Level3 level="3">
        <L1Offset>0 0 0</L1Offset>
      </Level3>
      - <Level8 level="8">
        <TID>1</TID>
        <L8Trim>-0.007 0 0 -0.191 0.177 -0.574</L8Trim>
        <MidContrastBias>0.106</MidContrastBias>
        <HighlightClipping>0.064</HighlightClipping>
        <SaturationVectorField>-0.02 -0.469 0 0.469 0.408 0</SaturationVectorField>
        <HueVectorField>0.087 0.102 0 -0.008 0.307 0</HueVectorField>
      </Level8>
      - <Level9 level="9">
        <SourceColorModel>255</SourceColorModel>
        <SourceColorPrimary>0.68 0.32 0.265 0.69 0.15 0.06 0.3127 0.329</SourceColorPrimary>
      </Level9>
    </DVDynamicData>
  </PluginNode>
</Shot>
- <Shot>
```

Abbildung 5-30: Dolby Vision XML Auszug zu Shot

```
<?xml version="1.0" encoding="UTF-8"?>
- <DolbyLabsMDF xmlns="http://www.dolby.com/schemas/dvmd/4_0_2">
  <Version>4.0.2</Version>
  - <RevisionHistory>
    - <Revision>
      <DateTime>2020-02-13T14:22:19Z</DateTime>
      <Author>Blackmagic Design</Author>
      <Software>DaVinci Resolve Studio</Software>
      <SoftwareVersion>16.1.2.026</SoftwareVersion>
    </Revision>
  </RevisionHistory>
  - <Outputs>
    - <Output>
      <CompositionName>Timeline 1</CompositionName>
      <UniqueID>cea913fa-f260-46a4-829d-e8e19bf91e66</UniqueID>
      <NumberVideoTracks>1</NumberVideoTracks>
      <CanvasAspectRatio>1.77778</CanvasAspectRatio>
      <ImageAspectRatio>1.85</ImageAspectRatio>
      + <Video>
    </Output>
  </Outputs>
</DolbyLabsMDF>
```

Abbildung 5-31: Dolby Vision XML Wrapper. Details befinden sich innerhalb von Video

5.4.3 Dolby Vision IMF



Das gerade bei Streamingdiensten angeforderte Format für Dolby Vision Inhalte ist ein Dolby Vision IMF, Interoperable Master Format. Der IMF Export ist in Resolve 16 schon als Vorlage implementiert. Es kann zusätzlich zwischen 20th Century Fox und Netflix ausgewählt werden. Die Voreinstellungen des Presets stimmen mit den Vorgaben von Netflix überein.

Abbildung 5-32: Video Export Einstellungen in Resolve

Tabelle 5.4-1: Netflix Spezifikationen für UHD Produktionen [39]

UHD Image Track		
Image Frame Width	3840	
Image Frame Height	2160	
Color Encoding	RGB / 4:4:4 / Full Range	
Frame Rates	23.976 / 24 / 25 / 29.97 / 30 / 50 / 59.94 / 60	
Frame Structure	Progressive	
Stereoscopy	Monoscopic	
JPEG 2000 Profile	Frame rates up to 30 fps use: <ul style="list-style-type: none"> 4k IMF Single Tile Lossy Profile Mainlevel 6 Sublevel 3 (max 800 Mbit/s) Frame rates above 30 fps use: <ul style="list-style-type: none"> 4k IMF Single Tile Lossy Profile Mainlevel 7 Sublevel 4 (max 1600 Mbit/s) 	
Dynamic Range	Dolby Vision™ HDR	SDR
Colorimetry	P3 D65 / SMPTE ST 2084 (PQ)	ITU-R BT.709 / D65 / ITU-R BT.1886 (Gamma 2.4)
Pixel Bit Depth	12-bit	10-bit
Mastering Display Color Volume Metadata	SMPTE ST-2086	n/a
Dynamic Metadata for Color Volume Transformation	Dolby Vision™ created by Dolby Mastering Tools SDK v2.6 or greater	n/a

Der Export liefert einen Ordner mit folgenden Dateien:

- (1) Videodatei als .mxf in JPEG2000 Codec
- (2) Audiodatei als .mxf
- (3) Mehrere XML Dateien, darunter eine CPL_XXX in der die Metadaten enthalten sind (siehe Abb. 3-1)

 ASSETMAP	06.02.2020 16:22	XML-Dokument	3 KB
 AUDIO_0acb9ef4-9580-4b0b-81e5-04c66f78661c	06.02.2020 16:22	MXF Video File (V...	4.496 KB
 CPL_f2898cd7-6fff-4ad5-ba8d-15b455e0a794	06.02.2020 16:22	XML-Dokument	15 KB
 OPL_64beb3a1-08fb-47fa-bd96-ad65099ff04c	06.02.2020 16:22	XML-Dokument	1 KB
 PKL_a431cd69-67ad-46ab-9ca9-2d141a08cfa1	06.02.2020 16:22	XML-Dokument	3 KB
 VIDEO_42f8cec4-f689-4f5b-a93c-0a7000b5e195	06.02.2020 16:22	MXF Video File (V...	1.469.175 KB
 VOLINDEX	06.02.2020 16:22	XML-Dokument	1 KB

Abbildung 5-33: IMF-Ordner Inhalt nach Export aus Resolve

Tabelle 5.4-2: Netflix Spezifikationen für HD Produktionen [39]

HD Image Track		
Image Frame Width	1920	
Image Frame Height	1080	
Color Encoding	RGB / 4:4:4 / Full Range	
Frame Rates	23.976 / 24 / 25 / 29.97 / 30 / 50 / 59.94 / 60	
Frame Structure	Progressived	
Stereoscopy	Monoscopic	
JPEG 2000 Profile	Frame rates up to 30 fps use: <ul style="list-style-type: none"> • 2k IMF Single Tile Lossy Profile Mainlevel 4 Sublevel 2 (max 400 Mbit/s) Frame rates above 30 fps use: <ul style="list-style-type: none"> • 2k IMF Single Tile Lossy Profile Mainlevel 5 Sublevel 3 (max 800 Mbit/s) 	
Dynamic Range	Dolby Vision™ HDR	SDR
Colorimetry	P3 D65 / SMPTE ST 2084 (PQ)	ITU-R BT.709 / D65 / ITU-R BT.1886 (Gamma 2.4)
Pixel Bit Depth	12-bit	10-bit
Mastering Display Color Volume Metadata	SMPTE ST-2086	n/a
Dynamic Metadata for Color Volume Transformation	Dolby Vision™ created with Dolby Mastering Tools SDK v2.6 or greater	n/a

5.4.4 SDR-Trim Version Export

Ab der Resolve Version 15 ist es zudem möglich im Export Fenster von Resolve die getonemappten Versionen einzeln auszuspielen. So kann zum Beispiel eine HDR Version in Dolby Vision gemastert werden und anschließend die SDR Version für den Kunden extra herausgerechnet werden. Dies spart immens Zeit. In den Videoexport Einstellungen unter Advanced Settings findet sich die Option Tone Mapping. Dort den gewünschten Trim Pass, der vorher natürlich erstellt worden ist, auswählen und das Video wie gewohnt exportieren.

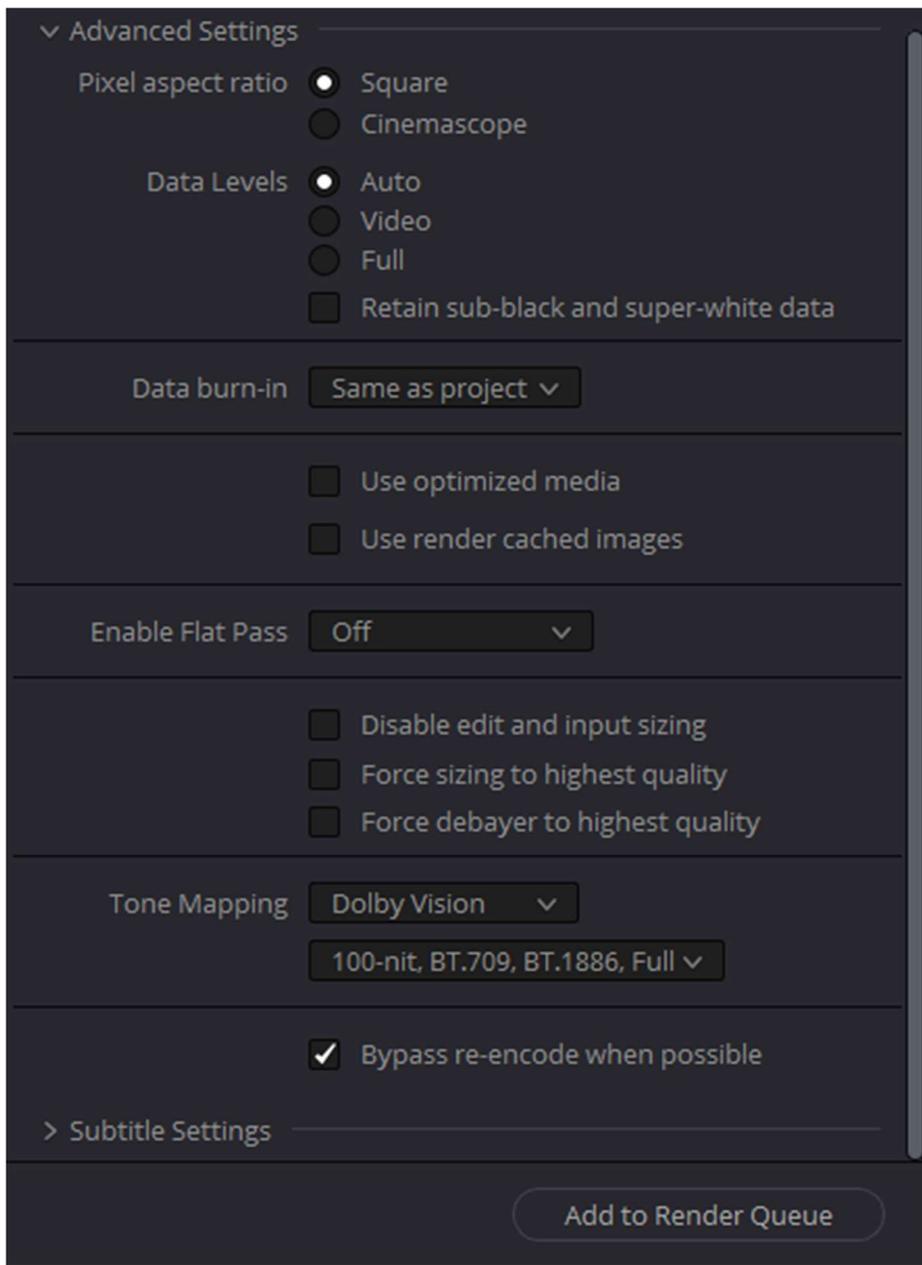


Abbildung 5-34: Tone Mapping Export in Davinci Resolve

6 Schlusswort

Diese Bachelorarbeit hat sich ausführlich mit HDR beschäftigt, um den Zusammenhang zwischen Creative Intent und Dynamischen Metadaten in Dolby Vision sowie deren Bedeutsamkeit zu prüfen.

Nach der intensiven Auseinandersetzung mit HDR und Dolby Vision sowie der praktischen Anwendung von Dolby Vision Mastering in DaVinci Resolve konnte ich mir ein gutes Bild davon machen wie Dolby mit Dolby Vision versucht den Creative Intent beizubehalten. Anhand des praktischen Teils dieser Arbeit kann ich bestätigen dass gerade die Primary und Secondary Trim Controls die Anpassung eines SDR Bildes erheblich vereinfachen. Auch die Zukunftsrelevanz von HDR und die Durchsetzung als neuer Standard wurde mir, beim graden und betrachten der Bilder an einem HDR Monitor, bewusst.

Jedoch ist eine weitere und klarere Adaption von HDR nötig. Gerade im Consumer Bereich ist es für viele Konsumenten noch unklar was HDR ist und auf was bei einem Kauf beachtet werden sollte. Dieser Prozess wird aber vermutlich noch einige Zeit dauern bis HDR in der Mitte der Gesellschaft angekommen ist. Doch die Entwicklung, gerade mit Dolby Vision IQ, geht in die richtige Richtung um dem Konsumenten das best mögliche Ergebnis zu liefern ohne dass dieser sich mit technischen Details herumschlagen muss. Außerdem wird es Interessant wie und ob sich HDR in Kinos etablieren wird und die technischen Errungenschaften sich dort in den nächsten Jahren ergeben werden.

Doch bei all den technischen Neuerungen und den brilliantesten Farben ist es vor allem Bedeutsam wie dieser erweiterte Spielraum im Kreativen Prozess genutzt wird. Denn das beste Bild bringt nichts wenn der Zuschauer nicht emotional vom Gesamtwerk angesprochen wird. Genau auch deshalb bin ich gespannt wie HDR in Zukunft für die narrative Unterstützung eingesetzt wird.

7 Literatur

- [1] askbiologist.asu.edu, *Rods and Cones of the Human Eye*. [Online] Verfügbar unter: <https://askbiologist.asu.edu/rods-and-cones>. Zugriff am: 7. Februar 2020.
- [2] T. Kunkel und E. Reinhard, „A reassessment of the simultaneous dynamic range of the human visual system“ (eng), S. 17, 2010, <http://dl.acm.org/citation.cfm?id=1836248>. Zugriff am: 7. Februar 2020.
- [3] F. Banterle, A. Artusi, K. Debattista und Alan Chalmers, *Advanced High Dynamic Range Imaging: Theory and Practice*. Natick, Massachusetts: A K Peters, Ltd, 2011.
- [4] Steve Holmes, „HDR, Who’s, How“.
- [5] C. Chinnock, „Dolby Vision and HDR10“, 2016.
- [6] T. Schulte und S. Barsotti, „HDR Demystified: EMERGING UHDTV SYSTEMS“, 2016, https://app.spectralcal.com/Documents/White%20Papers/HDR_Demystified.pdf. Zugriff am: 7. Februar 2020.
- [7] T. Fletcher, „2015 Camera Comparison Chart“, <https://drive.google.com/file/d/0B8uBU7GdVMLMM3cyNEhhMDdmVlk/view>. Zugriff am: 7. Februar 2020.
- [8] IHS Markit, „4K TV and UHD: the whole picture“, Jg. 2018.
- [9] ITU-R, „Recommendation ITU-R BT.709-6, Parameter values for the HDTV standards for production and international programme exchange“, 2015.
- [10] C. Chinnock, „Better Pixels in Professional Projectors“, 2016, <http://www.insight-media.info/insight-media-releases-new-white-paper-the-status-of-wide-color-gamut-uhd-tvs/>. Zugriff am: 7. Februar 2020.
- [11] E. Hart, „UHD Color For Games“, 2016, <https://developer.nvidia.com/sites/default/files/akamai/gameworks/hdr/UHDColorForGames.pdf>. Zugriff am: 7. Februar 2020.
- [12] Dolby Laboratories Inc., „Dolby Vision Whitepaper“, 2016.
- [13] K. Goeller, „Building the HDR Economy Nit by Nit“, *SMPTE 2015 Annual Technical Conference & Exhibition*, S. 1–9.
- [14] SMPTE, „SMPTE ST 2086:2018, Mastering Display Color Volume Metadata Supporting High Luminance and Wide Color Gamut Images“, 2018.
- [15] ITU-R, „Recommendation ITU-R BT.601-7, Studio encoding parameters of digital television for standard 4:3 and wide-screen 16:9 aspect ratios“, 2011.

- [16] A.-F. Perrin, M. Rerabek, W. Husak und T. Ebrahimi, „ICtCp Versus Y'CbCr: Evaluation of ICtCp Color Space and an Adaptive Reshaper for HDR and WCG“, *IEEE Consumer Electron. Mag.*, Jg. 7, Nr. 3, S. 38–47, 2018.
- [17] ITU-R, „Report ITU-R BT.2390-3, High dynamic range television for production and international programme exchange“, 2017.
- [18] DGIT, *4k HDR – 8 bit, 10 bit, 12 bit, and all the bits you need to know*. [Online] Verfügbar unter: <https://dgit.com/4k-hdr-guide-51429/>. Zugriff am: 7. Februar 2020.
- [19] S. Bilodeau, *HDR Video Part 1: What is HDR Video?* [Online] Verfügbar unter: <https://www.mysterybox.us/blog/2016/10/18/hdr-video-part-1-what-is-hdr-video>. Zugriff am: 7. Februar 2020.
- [20] ITU-R, „Recommendation ITU-R BT.2100-2, Image parameter values for high dynamic range television for use in production and international programme exchange“, 2018.
- [21] S. Daly, T. Kunkel, X. Sun, S. Farrell und P. Crum, „41.1: Distinguished Paper : Viewer Preferences for Shadow, Diffuse, Specular, and Emissive Luminance Limits of High Dynamic Range Displays“, *SID Symposium Digest of Technical Papers*, Jg. 44, Nr. 1, S. 563–566, 2013.
- [22] S. Miller, M. Nezamabadi und S. Daly, „Perceptual Signal Coding for More Efficient Usage of Bit Codes“, *The 2012 Annual Technical Conference & Exhibition*, S. 1–9.
- [23] ITU-R, „Report ITU-R BT.2246-6, The present state of ultra-high definition television“, 2017.
- [24] S. Miller, „A Perceptual EOTF for Extended Dynamic Range Imagery“, 2014, <https://www.smpte.org/sites/default/files/2014-05-06-EOTF-Miller-1-2-hand-out.pdf>. Zugriff am: 7. Februar 2020.
- [25] SMPTE, „SMPTE ST 2084:2014, High Dynamic Range Electro-Optical Transfer Function of Mastering Reference Displays“, 2014.
- [26] S. Bilodeau, *HDR Video Part 3: HDR Video Term Explained*. [Online] Verfügbar unter: <https://www.mysterybox.us/blog/2016/10/19/hdr-video-part-3-hdr-video-terms-explained>. Zugriff am: 7. Februar 2020.
- [27] Blu-ray Disc Association, „White Paper Blu-ray Disc Read-Only Format, Audio Visual Application Format Specifications for BD-ROM Version 3.1“, 2016.

- [28] L. Borg und R. Yeung, „Content-Dependent Metadata for Color Volume Transformation of High Luminance and Wide Color Gamut Images“, *SMPTE 2015 Annual Technical Conference and Exhibition*, Jg. 2015, S. 1–8.
- [29] L. L.C. HDR10+ Technologies, „HDR10+ System Whitepaper“, 2019, <http://www.hdr10plus.org>. Zugriff am: 7. Februar 2020.
- [30] SMPTE, „SMPTE ST 2094-10:2016, Dynamic Metadata for Color Volume Transform — Application #1“, 2016.
- [31] Dolby Laboratories Inc., „ICtCp Whitepaper“. Zugriff am: 18. Januar 2020.
- [32] R. L. de Valois und K. K. de Valois, „A multi-stage color model“, *Vision Research*, Jg. 33, Nr. 8, S. 1053–1065, 1993.
- [33] P.-C. Hung und R. S. Berns, „Determination of constant Hue Loci for a CRT gamut and their predictions using color appearance spaces“, *Color Res. Appl.*, Jg. 20, Nr. 5, S. 285–295, 1995.
- [34] Dolby Laboratories Inc., „Dolby Vision Color Grading Best Practices Guide“, Version 3.1, 2019.
- [35] C. J. Bartleson und E. J. Breneman, „Brightness perception in complex fields“, *Journal of the Optical Society of America*, Jg. 57, Nr. 7, S. 953–957, 1967.
- [36] A. G. Rempel, W. Heidrich, H. Li und R. Mantiuk, *Video viewing preferences for HDR displays under varying ambient illumination*. New York, NY: ACM, 2009.
- [37] SMPTE, „SMPTE May 2019 Table of Contents“, *SMPTE Mot. Imag. J.*, Jg. 128, Nr. 4, S. 1, 2019.
- [38] Nick Pino, *Dolby Vision IQ is about to make HDR TVs even better to look at*. [Online] Verfügbar unter: <https://www.techradar.com/news/dolby-vision-iq-is-about-to-make-hdr-tvs-even-better-to-look-at>. Zugriff am: 7. Februar 2020.
- [39] Netflix, „Netflix Originals Delivery Specifications OC-3-3“, 2019, <https://drive.google.com/file/d/0B37xotBv1V3Qc1E4VW5sMzNwcmc/view>. Zugriff am: 7. Februar 2020.