

HDR FÜR ALLE DISPLAYS

LUCIEN LENZEN, MASTER OF ENGINEERING (FKTG), PROF. DIPL.-ING. MIKE CHRISTMANN (FKTG),
 PROF. DR.-ING. ROLF HEDTKE (FKTG)

Mit dem Begriff High Dynamic Range (HDR) wird ein System beschrieben, das einen weit größeren Umfang an sichtbaren Details, sowohl in den dunklen als auch in den hellen Bereichen eines Bildes, bei der Aufnahme als auch bei der Wiedergabe erlaubt. Damit wird eine deutlich naturgetreuer Bildwiedergabe ermöglicht. In naher Zukunft werden allerdings die wenigsten Konsumenten ein entsprechendes Display besitzen. Um diese trotzdem von einer HDR-Aufzeichnung profitieren zu lassen, wäre eine Bildbearbeitung mit Hilfe sogenannter Tonemapping-Operatoren (TMOs) denkbar. Diese werden heute überwiegend für Standbilder und nicht für Videosequenzen eingesetzt. Außerdem wirken diese Bilder oftmals unnatürlich und es können Artefakte entstehen. Zur Lösung dieser Probleme werden im Folgenden drei Schritte zur Verbesserung der Bildqualität beim Einsatz eines TMOs bei Videosequenzen vorgeschlagen. Das vorgeschlagene Konzept soll den Anforderungen eines Live-Broadcast-Betriebs gerecht werden. Die grundlegende Idee besteht darin, jeden Leuchtdichtewert auf seine Relevanz für das Bild hin zu untersuchen und die unwichtigen Werte zu verwerfen. Dazu wird eine sogenannte „virtuelle Blende“ sowohl auf das gesamte Bild, als auch auf Teile des Bildes angewendet. Schließlich wird noch eine Kompensation der Chrominanz vorgenommen.

► High dynamic range (HDR) allows to capture an enormous range of luminance values. But in the near future the consumers will not have appropriate displays to see the quality difference. To improve the picture quality for SDR as well as HDR displays a new method is described in this paper. The basic idea is to analyse each luminance value for its relevance in the image and discard unimportant ones. This "virtual aperture" will be processed on the whole image and on image sections. Finally the tone mapping result will be composed with the corrected chrominance values.

Einleitung

Der vorliegende Beitrag ist der erste in einer Reihe über das Thema „High Dynamic Range“. In den nächsten Ausgaben der FKT werden weitere Aspekte von HDR genauer beleuchtet, während die Grundlagen von den Autoren bereits in der FKT 5/2016 ausführlich behandelt wurden. Als Start wird in

diesem Artikel noch einmal das klassische Problem einer Down-Konvertierung von HDR nach SDR adressiert. Der Artikel ist an die Veröffentlichung „HDR for legacy displays using sectional tone mapping“ [1] angelehnt.

Durch das gleichzeitige Erfassen von hellen und dunklen Bilddetails kann eine reale Szene sehr natürlich wiedergegeben werden. Die enorme Leistungssteigerung bei Bildaufnahmesensoren in den vergangenen Jahren hat dazu geführt, dass viele Videokameras heute in der Lage sind, einen Szenenkontrastumfang aufzuzeichnen, der jenseits der darstellbaren Möglichkeiten auf einem SDR-Display liegt. Mit Hilfe von neuen Transferfunktionen, wie Hybrid-Log-Gamma (HLG) oder dem Perceptual Quantizer (PQ), kann dieser vergrößerte Kontrastumfang anschließend effizient codiert werden. Doch wie kann es gelingen, dass auch Zuschauer mit einem herkömmlichen Fernsehgerät den deutlichen Mehrwert einer HDR-Produktion genießen können? Der HLG-Standard erhebt zwar den Anspruch, ein abwärtskompatibles Bild auf einem konventionellen Fernsehgerät darzustellen, jedoch wird hierbei die Qualität des SDR-Bildes nicht verbessert.

Anforderungen und Ziel

Wenn HDR-Material auf einem SDR-Display dargestellt werden soll, stellt sich die Frage, wie der enorme Kontrastunterschied behandelt werden kann. Der einfachste Weg wäre, die Werte außerhalb des SDR-Bereichs abzuschneiden. Dies führt jedoch zum Verlust von Zeichnung in hellen und dunklen Bildregionen. Der vielversprechendere Weg ist es, einen Tonemapping-Operator (TMO) zu verwenden. Das Ziel soll daher sein, ein Tonemapping-basierendes System zu entwickeln, das die Anforderungen von Live-Broadcast erfüllt und dabei die typischen Tonemapping-Probleme, die später noch erläutert werden, eliminiert.

Zunächst sollen die Anforderungen für ein solches System definiert werden:

- Das System muss den Zuschauerpräferenzen entsprechen und darf gleichzeitig nicht zu sehr vom bekannten Seherlebnis abweichen.
- Das System muss robust gegen alle Arten von TMO-Artefakten sein (eine nachträgliche Verbesserung ist im Live-Betrieb nicht möglich).
- Das System muss eine automatische, kontinuierliche Anpassung an sich ändernde Lichtverhältnisse durchführen (im Gegensatz zu den heutigen diskreten Blendenstufen bei einer mechanischen Regelung).
- Das System muss in der Lage sein, auf Szenenwechsel adäquat zu reagieren.
- Das System muss die maximal erzielbaren Leuchtdichten eines SDR-Displays optimal ausnutzen.
- Das System muss eine Echtzeitverarbeitung garantieren.

Stand der Technik und Evaluation

Zunächst soll die Zuschauerpräferenz bei einer Kontrastkompression genauer untersucht werden. Dazu wurden fünf professionelle Coloristen gebeten, sieben HDR-Sequenzen

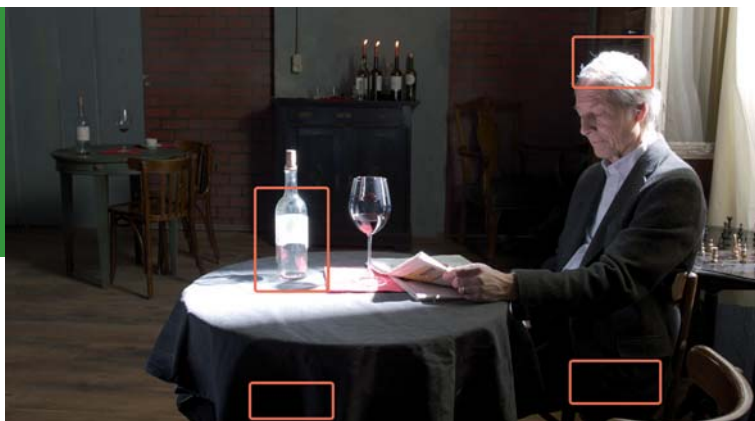


Bild 1. „bistro_01“ aus dem HDR-Testmaterial von Fröhlich et al. [8] – die sehr hellen bzw. dunklen Bereiche sind orange markiert. Bilder falls nicht anders angegeben, die Autoren

für ein SDR-Display zu graden [2]. Im anschließenden Versuch wurden 20 Probanden, mehrheitlich Laien, befragt, welche Version ihnen persönlich am besten gefällt. Dazu wurde jede Version mit jeder Version direkt verglichen. Die Probanden mussten sich für eine der beiden Versionen entscheiden. Es zeigte sich dabei, dass die Versionen mit hohem Kontrast und hoher Sättigung bevorzugt wurden. Dies deckt sich auch mit mehreren anderen Studien [3, 4, 5]. Ein weiteres wichtiges Ergebnis dieser Untersuchungen war die Bevorzugung eines homogenisierten Bildes. Darunter versteht man die Möglichkeit, dass der Zuschauer alle im Bild enthaltenen Details, sowohl in den Lichtern als auch in den Schatten, ohne Anstrengungen noch sehen möchte. Hierfür können eine hohe Sättigung, eine gesteigerte Helligkeit und ein hoher Kontrast in einem gewissen Rahmen nützlich sein. Auch ein Begrenzen der maximalen Helligkeitswerte (Clipping) wird akzeptiert, wenn dadurch keine wichtigen Informationen der Szene verloren gehen. Diese Begrenzung führt zu einer steileren Gradation und somit zu einem erhöhten Kontrast. Es muss also ermittelt werden, welche Leuchtdichtewerte noch relevanten Mehrwert bieten bzw. auf welche Werte zu Gunsten von Helligkeit und Kontrast verzichtet werden kann.

Die genannten Zuschauerpräferenzen mit Hilfe eines TMOs zu erreichen, ist keine einfache Aufgabenstellung. Zunächst sollen die beiden Klassen von TMOs erläutert werden. Es wird grundsätzlich zwischen sogenannten globalen und lokalen Tonemapping-Operatoren unterschieden. Während die globalen Operatoren eine Übertragungskurve erstellen, die meist auf der durchschnittlichen und der maximalen Leuchtdichte des gesamten Bildes beruht, analysieren lokale Operatoren die Nachbarschaft jedes Pixels für eine individuelle Kurve, die auf dem lokalen Adaptionsniveau beruht. Das lokale Verfahren ist rechenintensiver und das Ergebnis birgt auch ein höheres Risiko für Artefakte, wie z. B. Halo-Effekte oder Flackern. Dabei kann das Ergebnis surreal wirken und sich extrem vom heutigen Bildeindruck unterscheiden. Auch bei globalen Operatoren kann es, durch sich zeitlich verändernde Helligkeiten, zu Bildflackern kommen. Dies lässt sich allerdings durch entsprechende Maßnahmen besser verhindert als bei lokalen Operatoren. Das Kernproblem der globalen Operatoren ist jedoch, dass das Bild durch die große Kontrastkompression einen kontrastarmen, flauen Bildeindruck erhält. [3, 5, 6]

Die grundlegende Idee bei dem im Folgenden beschriebenen und an der Hochschule RheinMain entwickelten Verfahren [7], ist deshalb das subjektive Empfinden von Kon-

trast, Sättigung und Schärfe zu verbessern, indem globales Tonemapping mit drei Verbesserungsmaßnahmen kombiniert wird. Das Verfahren trägt die Bezeichnung „EVI“ (Enhanced Video Imaging).

Enhancement-Schritte

1: Virtuelle Blende

Eine logarithmische Übertragungskennlinie besitzt eine flache Gradation. Hierdurch kann ein hoher Kontrastumfang mit einer geringen Anzahl von Bits gegenüber einer linearen Übertragungskennlinie codiert werden. Stellt man eine solche Kurve direkt auf einem standardmäßigen SDR-Display dar, fehlt es dem Bild an Kontrast; es wirkt flach und ohne Tiefe. Je größer der abzubildende Kontrastumfang dabei ist (Abstand von hellstem zu dunkelstem Wert), desto flacher wird auch die Kurve eines globalen TMOs. Die Frage, die sich im Zusammenhang mit der Darstellung von Hochkontrastbildern grundsätzlich stellt, ist, ob es notwendig ist, bei jedem Bild den gesamten Kontrastumfang wiederzugeben und welche Informationen speziell auf einem SDR-Display noch sinnvoll dargestellt werden können. Im Grunde genommen sind dies ähnliche Fragestellungen, wie sie bei der Datenkompression von Video- und Audiosignalen, hinsichtlich der Irrelevanz-Reduktion, diskutiert werden.

Betrachtet man z. B. die Szene „bistro_01“ (Bild 1) von Fröhlich et al. [8], liegt der höchste Leuchtdichtewert bei $58,253 \text{ cd/m}^2$ und der niedrigste bei $0,16 \text{ cd/m}^2$ (inklusive Rauschen). Dies entspricht einem Kontrastumfang von 18,5 Blendenstufen. Reduziert man bei den Maximal- bzw. Minimalwerten lediglich um jeweils ein Promille, so liegen die Werte zwischen $6,653 \text{ cd/m}^2$ und $0,24 \text{ cd/m}^2$. Dies entspricht einer Reduktion von nahezu vier Blendenstufen. Wie unsere Untersuchungen ergeben haben, führt dies beim Zuschauer zu keinem erkennbaren Verlust von Bildinformationen. Das Beispiel zeigt, wie groß der Einfluss der Extremwerte ist. Vergleicht man die beiden Übertragungskennlinien vor und nach der Reduktion der Helligkeitswerte, so erkennt man, dass nach der Reduktion die Übertragungskennlinie eine steilere Gradation aufweist.

Aus diesem Grunde verzichten einige TMOs auf das hellste und dunkelste Prozent der Leuchtdichten, was einer Form von Vorfilterung entspricht. Der Nachteil einer grundsätzlich festen Begrenzung ist allerdings, dass dies bei manchen Bildern subjektiv viel, bei anderen subjektiv wenig Einfluss hat.

Nicht jede Szene braucht eine gleich steile Gradation. Deshalb bewertet das hier beschriebene Verfahren adaptiv und automatisch die Relevanz jedes Leuchtdichtewertes anhand eines Histogramms.

Dazu wird die Leuchtdichte (Y) jedes Pixels logarithmisch in einem Histogramm aufgetragen. Der Kontrastumfang des Histogramms ist dabei so ausgelegt, dass jeder realistisch denkbare Wert abgebildet werden kann. Mit Hilfe eines von der Display-Leuchtdichte abhängigen Kontrastkastens wird eine Position auf dem Histogramm ermittelt, um welche sich die meisten Leuchtdichtewerte befinden. Dies kann als „center of interest“ interpretiert werden. Je größer die Display-Leuchtdichte, desto größer ist auch der Kontrastkasten und der damit verbundene Mindestkontrastumfang. Aus der Lage und der Größe des Kastens lässt sich eine angenäherte Gaußfunktion zeichnen, die zur Gewichtung des Histogramms dient. Die Gewichtung entspricht dabei einer Multiplikation der Werte mit der Gaußfunktion. Das so entstandene Histogramm wird mit zwei, auf der Umgebungsluchtdichte basierenden Grenzwerten, s_1 und s_2 , verglichen. Die kleinste Leuchtdichte, die zu einer Überschreitung von s_1

führt, ist die kleinste relevante Leuchtdichte. Die größte Leuchtdichte, die zu einer Überschreitung von s_2 führt, ist die größte relevante Leuchtdichte. Bild 2 zeigt die Funktionsweise anhand der Histogramme. Wir bezeichnen diesen Vorgang als „virtuelle Blende“, weil es mit dem konventionellen Bestimmen des Arbeitspunktes einer Kamera durch das Nachziehen der mechanischen Blende vergleichbar ist.

In Bild 3 ist an einer realen Szene der Vergleich eines Bildes mit und ohne virtueller Blende gezeigt.

2: Sektionales Tonemapping

Es zeigt sich allerdings, dass eine globale Behandlung alleine nicht ausreichend ist, um eine optimale Kontrastverbesserung zu erhalten. Damit man einen lokalen Einfluss berücksichtigen kann, wird das Bild in Blöcke zerlegt. Eine Größe von 16×9 Blöcken hat sich bei unseren Untersuchungen als geeigneter Wert erwiesen. Bei einer größeren Anzahl der Blöcke sind nicht genügend Werte für eine gute Histogrammauswertung innerhalb eines Blockes vorhanden. Bei einer kleineren Anzahl ist der lokale Einfluss zu gering. Die virtuelle Blende wird nun auf jedem Block einzeln angewendet. Dieses Verfahren bezeichnen wir als „sektionales Tonemapping“. Verglichen zu klassischen lokalen Tonemapping-Verfahren können dabei Probleme, wie Halo-Effekte und ein unnatürliches Erscheinungsbild vermieden werden.

Bei diesem sektionalen Tonemapping-Verfahren gilt es allerdings die sichtbaren Stufen an den Blockgrenzen zu vermeiden (s. Bild 4.2). Dies erreicht man, indem man die Werte für Y_{max} und Y_{min} mit den Werten der Nachbarblöcke filtert. Dafür wird eine nichtlineare Funktion genutzt. Im Vergleich zur globalen Variante sind Y_{max} und Y_{min} jetzt Matrizen der Größe 16×9 und keine Skalare mehr. Anschließend werden die Matrizen an die Größe des Gesamtbildes angepasst. Auch der Mittelwert Y_{mean} kann zusätzlich auf Blockebene ermittelt werden. Bild 4 verdeutlicht den gesamten Bearbeitungsschritt.

3: Modifizierte Farbbehandlung

Eine Kontrastkompression wird lediglich auf der Helligkeitskomponente eines Signals durchgeführt. In der Folge muss dafür gesorgt werden, dass die Farbkomponenten, also der Farbton und die Farbsättigung, bei dieser Bearbeitung unbeeinflusst bleiben. Die meisten TMOs arbeiten dazu im XYZ-Farbraum und übertragen die Änderung von Y (Yin zu Yout) auch auf X und Z. Allerdings ist dieser Farbraum nicht wahrnehmungsgerecht und beeinflusst so den Farbeindruck. In der Regel wird im Broadcast-Umfeld jedoch mit Farbdifferenzsignalen in der Form von Luminanz und Chrominanz ($YCbCr$) gearbeitet. Da dieses Farbdifferenzsignal bei Helligkeitsänderungen auch Farbänderungen hervorruft (schlechte Luminanz/Chrominanz Dekorrelation), muss auch hierbei eine entsprechende Anpassung der Farbwerte, je nach Änderung in der Leuchtdichte, vorgenommen werden. Das

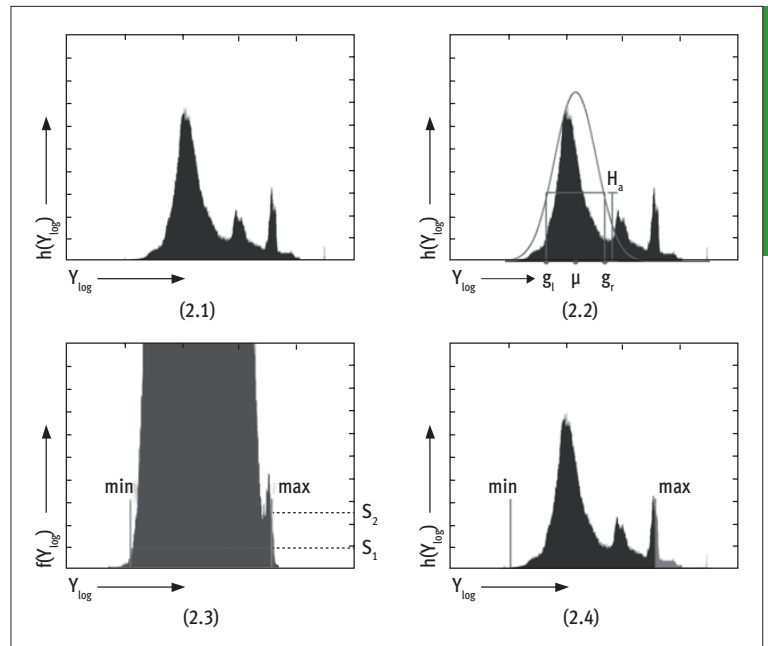


Bild 2: (2.1) logarithmisches Histogramm der Eingangswerte, (2.2) Histogramm mit Kontrastkasten und angenäherter Gaußfunktion, (2.3) gewichtetes Histogramm durch Multiplikation mit dem Histogramm und Schnittpunkte mit den Grenzwerten, (2.4) Histogramm mit den veränderten Maximal- und Minimalwerten

von der Firma Dolby Labs. vorgestellte IC_Cp -System [9] verspricht hier ein besseres Ergebnis. Da es sich hierbei um sehr komplexe Zusammenhänge handelt, wird dieses Thema, in einem späteren Artikel der Serie, noch einmal ausführlich behandelt.

Fazit und Ausblick

Das vorgestellte Verfahren zur Optimierung von TMOs für den Einsatz im Broadcast-Umfeld ermöglicht es erstmals, die Fortschritte bei der Kamertechnik auch bei der Wiedergabe auf einem SDR-Display zu nutzen und gleichzeitig einen deutlich höheren Szenenkontrast darzustellen. Ein weiterer Vorteil des vorgeschlagenen Verfahrens besteht darin, dass die Blende, bei moderaten Helligkeitsänderungen, aufgrund der automatischen Nachregelung unverändert bleiben kann. Dabei werden alle oben angeführten Broadcast-Anforderungen erfüllt.

Das vorgestellte Verfahren ist aber nicht nur in der Lage aus einem HDR-Signal ein verbessertes SDR-Signal zu erzeugen, sondern auch ein HDR-Kamerasignal für die Anzeige auf einem HDR-Display zu optimieren. Durch das sektionale Vorgehen kann die uneinheitliche Ausleuchtung einer Szene kompensiert werden. Schattenpartien können aufgehellt werden, während die Lichter gleichzeitig abgedunkelt wer-



Bild 3: (3.1) Aufnahme mit der Kamera-Transferkurve, (3.2) gleiche Aufnahme mit globalem Tonemapping-Operator bearbeitet, (3.3) gleiche Aufnahme mit globalem Tonemapping-Operator und virtueller Blende



(4.1)

(4.2)

(4.3)

Bild 4: (4.1) virtuelle Blende über das gesamte Bild, (4.2) virtuelle Blende auf 16 x 9 Blöcke angewendet, (4.3) sektionale virtuelle Blende mit Filterung

den. Ein klassisches Beispiel hierfür wäre eine überdachte Moderatorenposition vor einem hellen Hintergrund. In der Folge wären die Details, sowohl der Moderator sowie beispielsweise die des Spielfelds und des Himmels, besser sichtbar.

Derzeit wird das System mit einer echtzeitfähigen GPU-Implementierung in Feldversuchen erprobt. Des Weiteren wird die Entwicklung in Form eines LOEWE-Forschungsprojektes, gefördert durch das Land Hessen, vorangetrieben. Ziel ist es, das Verfahren in eine praxisgerechte Hardware zu überführen.

- Die weiteren Beiträge dieser Serie beschäftigen sich mit der genaueren Analyse der Zuschauerpräferenz und -wahrnehmung,
- dem Einfluss des Tonemappings auf die Farbkomponente sowie
- dem Einfluss des Tonemappings auf die Datenraten bei der Distribution. ◀

Literatur

- [1] Lenzen, L., 2016. HDR for legacy displays using Sectional Tone Mapping. Proceedings of the IBC conference 2016, September 2016
- [2] Schmitt, D., 2015. High Dynamic Range in Videobildern - Akzeptanzermittlung von Coloristen und Konsumenten sowie die Untersuchung der HDR-Fähigkeit bestehender Kameraaufnahmen und Filmmaster. Bachelor Thesis an der Hochschule RheinMain, Mai 2015
- [3] Petit, J. and Mantiuk, R., 2013. Assessment of video tone-mapping: Are cameras' S-shaped tone-curves good enough? J. Visual Communication and Image Representation. Vol. 23 pp. 1020-1030.
- [4] Akyüz, A. O., Fleming, R., Riecke, B. E., Reinhard, E. and Bülthoff, H. H., 2007. Do HDR Displays Support LDR Content? A Psychophysical Evaluation. ACM Trans. Graph. Juli 2007, Artikelnummer 38.
- [5] Eilertsen, G., Wanat, R., Mantiuk, R. K. and Unger, J., 2013. Evaluation of Tone Mapping Operators for HDR-Video. Comput. Graph. Forum. Vol. 32 pp. 275-284.
- [6] Aydin, T. O., Stefanoski, N., Croci, S., Gross, M. and Smolic, A., 2014. Temporally Coherent Local Tone Mapping of HDR Video. ACM Trans. Graph. vol. 33 pp. 196:1-196:13.
- [7] Patent DE102015119137B3 Erfindung betreffend Verfahren zur Blendeneinstellung
- [8] Froehlich, J., Grandinetti, S., Eberhardt, B., Walter, S., Schilling, A. and Brendel, H., Creating Cinematic Wide Gamut HDR-Video for the Evaluation of Tone Mapping Operators and HDR-Displays. Proc. SPIE. Vol. 9023
- [9] Dolby Labs., 2016. ICtCp White Paper Version 7.2 <https://www.dolby.com/us/en/technologies/dolby-vision/ICtCp-white-paper.pdf>

Lucien Lucien



LUCIEN LENZEN Master of Engineering (FKTG)

ist wissenschaftlicher Mitarbeiter/Doktorand im Forschungsbereich UHD TV an der Hochschule RheinMain in Wiesbaden

Mike Christmann



MIKE CHRISTMANN (FKTG)

Lehrt Film- und Broadcasttechnik an der Hochschule RheinMain, Studiengang Medientechnik. Schwerpunkte seiner wissenschaftlichen Arbeit beinhalten unter anderem die Bereicheameratechnik sowie die Broadcast-IT

© Andreas Schlote



ROLF HEDTKE (FKTG)

war Professor im Studiengang Medientechnik an der Hochschule RheinMain und ist inzwischen emeritiert. Er ist dort noch als Berater bei Forschungsprojekten in den Bereichen Bildbearbeitung und UHD TV tätig