

Humboldt-Universität zu Berlin



Automatische Belichtungsregelung in High Dynamic Range-Videos

Studienarbeit im Rahmen des Diplomstudienganges Informatik

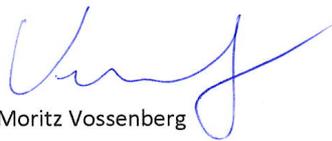
Humboldt-Universität zu Berlin
Mathematisch-Naturwissenschaftliche Fakultät II
Institut für Informatik

Eingereicht von: Moritz Vossenberg
Betreuer: Prof. Dr. Peter Eisert

Berlin, 19.05.2015

Selbstständigkeitserklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbstständig verfasst und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe.



Moritz Vossenberg

Berlin, den 19.05.2015

Inhaltsverzeichnis

1. Einleitung	4
2. Die Kamerafunktion	5
2.1. State of the Art	6
2.1.1. Bestimmung der Kamerafunktion nach Debevec und Malik . .	6
2.1.2. Bestimmung der Kamerafunktion nach Mitsunaga und Nayar .	9
2.2. Modellierung der Kamerafunktion als Bézierkurve	11
2.3. Ergebnisse	12
3. Schätzung des HDR-Histogramms	14
4. Auswahl der Belichtungszeiten	15
4.1. Berechnung nach Gallo et al.	16
5. HDR-Video	20
5.1. Belichtungssteuerung	20
6. Ergebnisse	22
7. Fazit	23
A. Literatur	26

1. Einleitung

Diese Studienarbeit befasst sich mit dem Problem der automatischen und kontinuierlichen Auswahl von geeigneten Belichtungszeiten in HDR-Video-Systemen.

In der HDR-Fotografie lässt sich dieses Problem noch vergleichsweise einfach lösen. Die Auswahl geschieht meistens manuell, basierend auf Erfahrung sowie dem Durchprobieren verschiedener Belichtungen bei statischen Szenen. Bei Videoaufnahmen hat man es jedoch mit komplexeren Problemen zu tun. Die Kamera muss dazu in der Lage sein, sich auf wechselnde Szenen und Belichtungsverhältnisse adaptieren zu können. Dabei muss darauf geachtet werden, dass die Auswahl so rasch geschieht, dass keine Verzögerung im Video auftritt. Andererseits muss die Auswahl aber auch so geregelt werden, dass es nicht zu plötzlichen Helligkeitsschwankungen kommt.

Zudem muss die Auswahl der Belichtungen vollautomatisch ohne Nutzerinteraktion geschehen, damit das System echtzeitfähig sein kann. Unter dieser Voraussetzung ist auch die Anzahl der Belichtungen stark beschränkt.

Der Beitrag dieser Arbeit liegt in der Auswahl von Techniken zu Belichtungssteuerung und -messung in der Fotografie, sowie der Entwicklung eines Modells, um eben diese in geeigneter Weise zu kombinieren und zu einem effizienten HDR-Video-System zu entwickeln.

2. Die Kamerafunktion

Die Kamerafunktion¹ stellt die Beziehung zwischen dem in Richtung Kamerasensor emittierten Licht eines Punktes in der aufgenommenen Szene (engl.: scene radiance) und dem entsprechenden digitalisierten Pixelwert dar.

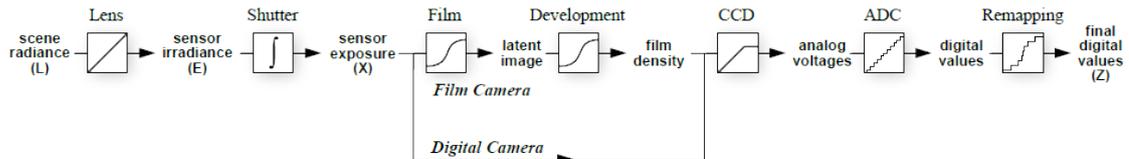


Abbildung 1: Kameramodell [Debevec and Malik, 2008]

Ein typisches Kamerasystem lässt sich wie in Debevec and Malik [2008] als das in Abb. 1 dargestellte Modell beschreiben:

Die scene radiance L passiert das Objektiv der Kamera und trifft auf den Sensor der Kamera als sensor irradiance E . Durch die Belichtungszeit wird nun die sensor exposure X bestimmt. Dies ist die Menge des Lichts, die innerhalb der Belichtungszeit Δt auf den Sensor trifft, d.h. $X = E\Delta t$. Die analogen Messwerte des Sensors werden nun digitalisiert und quantisiert um den endgültigen Pixelwert Z zu erhalten.

Innerhalb dieser Schritte oder nach der Digitalisierung werden je nach Kamera noch unterschiedlichste Maßnahmen zur Optimierung des Bildes, wie z.B. Rauschunterdrückung, Farbgebung, Korrektur von Abbildungsfehlern wie Verzeichnung oder chromatischer Aberration und Komprimierung durchgeführt.

Alle diese Schritte (siehe Abb. 1) lassen sich durch eine nichtlineare Funktion $Z = f(X)$, der Kamerafunktion (Abb. 2), beschreiben.

Ohne Kenntnis dieser nichtlinearen Funktion wäre man ohne Weiteres nicht in der Lage, von einem Pixelwert auf den realen Helligkeitswert der aufgenommenen Szene zu schließen. Dies ist im Bereich der HDR-Fotografie jedoch bedeutsam, um eine Beziehung zwischen verschiedenen Belichtungen oder Belichtungszeiten herzustellen.

¹engl.: Camera Response Function.

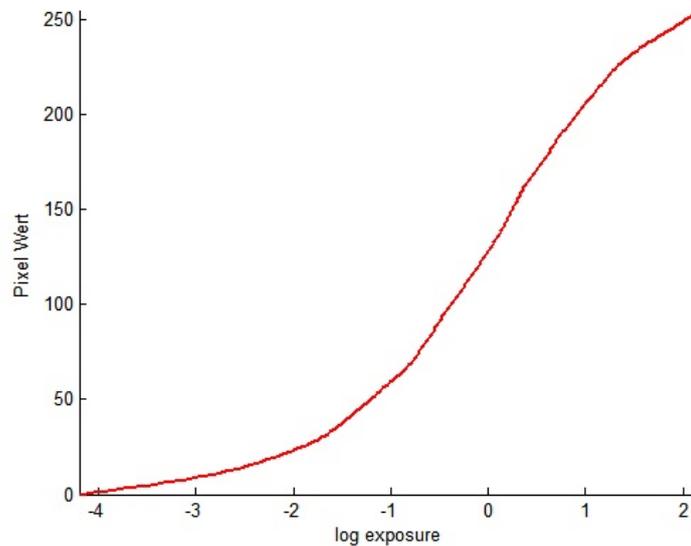


Abbildung 2: Eine Kamerafunktion

Insbesondere wenn die Kamerafunktion nicht linear ist, ergeben sich damit auch nicht lineare Beziehungen zwischen Z und X , bzw. deren Quantisierungsintervallen. Somit sind auch Quantisierungsfehler nicht linear über das Histogramm der Szene verteilt, was eine wichtige Rolle bei der Auswahl einer Belichtungsreihe spielen kann. Darauf wird näher in Kapitel 3 eingegangen.

2.1. State of the Art

Im Folgenden werden zwei Methoden zur Berechnung der Kamerafunktion vorgestellt, die auch Grundlagen der angepassten Methode zur Kamerafunktionsberechnung und Belichtungssteuerung darstellen.

2.1.1. Bestimmung der Kamerafunktion nach Debevec und Malik

Eine der am häufigsten angewandten Methoden zur Bestimmung der Kamerakurve wurde von Debevec und Malik [Debevec and Malik \[2008\]](#) beschrieben. Das Verfahren basiert auf dem Reziprozitätsgesetz. Dieses besagt, dass die Belichtung eines

Filmes/Sensors bei Verringerung der Belichtungszeit gleich bleibt, wenn die Lichtintensität in gleichem Maße erhöht wird.

Somit wird als Grundlage des Belichtungsmodells einzig das Produkt aus der irradiance² E und der Belichtungszeit Δt betrachtet:

$$X = E\Delta t \tag{1}$$

Wobei die Belichtung (radiant exposure H_e) X die Einheit J/m^2 besitzt.

Die Kamerafunktion muss nun aus den einzig bekannten Werten, der Belichtungszeit und dem Pixelwert Z , gewonnen werden. Da die Kamerafunktion die Beziehung zwischen der radiant exposure und dem Pixelwert Z beschreibt, lässt sich die Gleichung

$$Z = f(X) \tag{2}$$

aufstellen. Unter der Annahme, dass f monoton steigend ist, ist die Inverse f^{-1} definiert. Somit lässt sich, sobald f berechnet ist, die Gleichung (2) invertieren und die radiant exposure aus den Pixelwerten zurückgewinnen.

$$X = f^{-1}(Z) \tag{3}$$

Nun, da X bekannt ist, lässt sich auch die irradiance durch Umstellen von (1) berechnen:

$$E = X/\Delta t \tag{4}$$

Das Problem besteht nun darin, die zwei Unbekannten f^{-1} und E mittels nur zwei Bekannter Z und Δt zu finden. Debevec und Malik lösen dies durch Aufstellung eines linearen Optimierungsproblems. Dabei machen sie sich zu nutze, dass für die Funktion f^{-1} nur eine endliche Anzahl an Elementen (die der Pixelwerte 0..255) benötigt wird.

²engl.: Irradiance. Die gesamte Leistung der elektromagnetischen Energie, die auf eine Fläche trifft (bezogen auf deren Größe). Die Einheit ist W/m^2 .

Nach Gleichung (2) gilt für jeden Pixelwert Z_{ij} :

$$Z_{ij} = f(E_i \Delta t) \quad (5)$$

Wobei j das Bild (bzw. die entsprechende Belichtungszeit) aus dem Belichtungsstack bezeichnet, welches als Eingabe des Algorithmus dient und i der Index des Pixels im Bild ist.

Mittels der inversen Kamerafunktion und durch Logarithmierung erhalten wir:

$$\ln(f^{-1}(Z_{ij})) = \ln(E_i) + \ln(\Delta t) \quad (6)$$

Womit sich

$$g(Z_{ij}) = \ln(f^{-1}(Z_{ij})) \quad (7)$$

definieren lässt.

Die Lösung wird nun mittels linearer Optimierung durch Minimierung des MSE gefunden. Dazu wird die folgende, quadratische Zielfunktion aufgestellt:

$$O = \sum_{i=1}^N \sum_{j=1}^P \{w(Z_{ij})[\ln(f^{-1}(Z_{ij})) - \ln E_i - \ln \Delta t_j]\}^2 + \lambda \sum_{z=Z_{min}+1}^{Z_{max}-1} w(Z_{ij})g''(z)^2 \quad (8)$$

Der zweite Term dient zur Sicherstellung, dass f^{-1} glatt ist. Es ist die Summe der quadrierten Werte der zweiten Ableitung $g''(z) = g(z-1) - 2g(z) + g(z+1)$ mit einem Gewichtungsfaktor λ^3 .

Die typische Form der Kamerafunktion entspricht der einer Gammafunktion. Sie ist also zu den Randbereichen hin (Z_{min} und Z_{max}) sehr steil, was bedeutet, dass in diesen Bereichen die Daten schlechter gefittet werden. Die Sättigung verstärkt zudem noch die Ausblendung der Randbereiche.

Aus diesem Grund wird die Gewichtungsfunktion $w(z)$ in die Zielfunktion inte-

³Der Gewichtungsfaktor wird abhängig von der Stärke des Rauschens in Z gewählt.

griert.

$$w(z) = \begin{cases} z - Z_{min} & \text{für } z \leq \frac{1}{2}(Z_{min} + Z_{max}) \\ Z_{max} - z & \text{für } z > \frac{1}{2}(Z_{min} + Z_{max}) \end{cases} \quad (9)$$

Diese ist eine einfache Hut-Funktion und schwächt somit die Glättungs und Fitting Terme eben dieser schlecht aufgelösten Randbereiche der Kamerafunktion ab.

Ein weiteres Problem besteht noch darin, dass $g(z)$ und E nur bis auf einem bestimmten Skalierungsfaktor bestimmt werden können. Die Zielfunktion würde sich nicht ändern, falls g und E_i durch $g + \alpha$ bzw. $E_i + \alpha$ ersetzt würden. Somit kann die Bedingung $g(Z_{mid}) = 0^4$ definiert werden, welche dafür sorgt, dass dem Pixelwert Z_{mid} Einheitsbelichtung zugewiesen wird und die Lage der Kamerafunktion somit fixiert ist. Diese Bedingung erweitert das Gleichungssystem nur um eine Zeile.

Gelöst wird das Gleichungssystem mittels Singular Value Decomposition (SVD). Als Ergebnis erhält man die logarithmische, inverse Kamerafunktion

$$g(Z_{ij}) = \ln f^{-1}(Z_{ij})$$

2.1.2. Bestimmung der Kamerafunktion nach Mitsunaga und Nayar

Mitsunagas und Nayars [Mitsunaga and Nayar \[1999\]](#) Methode ist der von Debevec und Malik [Debevec and Malik \[2008\]](#) prinzipiell recht ähnlich. Sie modellieren jedoch die Kamerafunktion durch ein Polynom:

$$f(Z) = \sum_{n=0}^N c_n Z^n \quad (10)$$

Sie gehen vom gleichen Kameramodell aus wie Debevec und Malik. Die brightness I der Szene ist das Produkt aus irradiance E und Belichtungszeit t mit $E = L \frac{\pi}{4} \left(\frac{d}{h}\right)^2 \cos^4 \phi$. Dabei ist h die Brennweite, d die Blende und L die scene radiance.

⁴ $Z_{mid} = \frac{1}{2}(Z_{min} + Z_{max})$

Ihr ideales, lineares Kamerasystem wäre also

$$I = L * k * e \quad (11)$$

für $k = \cos^4\phi/h^2$ und $e = (\pi d^2/4)t$. e wird als Belichtung bezeichnet.

Die Kamerafunktion stellt nun wieder die Verbindung zwischen den Pixelwerten Z und der scene radiance I her:

$$Z = g(I) \Rightarrow I = f^{-1}(Z) \quad (12)$$

Debevec und Malik nutzen in ihrem Verfahren die Antizipation der Kamerakurvenform mittels der Gewichtungsfunktion, um das überbestimmte Gleichungssystem aufstellen und lösen zu können. Mitsunaga und Nayar hingegen verwenden das Verhältnis zweier aufeinanderfolgender Belichtungen $R_{q,q+1} = e_q/e_{q+1}$ als Basis zur Berechnung der Kamerafunktion. Dies hat den Vorteil, dass die genauen Belichtungszeiten und Blendenwerte nicht bekannt sein müssen.

Dieses Verhältnis lässt sich in folgende Darstellung überführen:

$$\frac{\sum_{n=0}^N c_n * Z_{p,q}^n}{\sum_{n=0}^N c_n * Z_{p,q+1}^n} = R_{q,q+1} \quad (13)$$

Damit lässt sich die Fehlerfunktion aufstellen:

$$\varepsilon = \sum_{q=1}^{Q-1} \sum_{p=1}^P \left[\sum_{n=0}^N c_n Z_{p,q}^n - R_{q,q+1} \sum_{n=0}^N c_n Z_{p,q+1}^n \right]^2 \quad (14)$$

Um alle Messungen zu normalisieren und für die Skalierung $g^{-1}(1) = I_{max}$ zu setzen, wird der constraint $c_N = I_{max} - \sum_{n=0}^{N-1} c_n$ aufgenommen. Dies bedeutet, dass die Dimension des Gleichungssystems um einen Koeffizienten reduziert wird.

Das Minimum lässt sich nun durch Finden der Nullstellen der partiellen Ableitungen

der Koeffizienten finden. Also durch Lösen des Gleichungssystems für

$$\frac{\delta\varepsilon}{\delta c_n} = 0$$

Das Gleichungssystem ist im Gegensatz zur Methode von [Debevec and Malik \[2008\]](#) also wesentlich kleiner und effizienter lösbar.

Ein Problem besteht in der Auswahl des Polynomgrads. Hier wird in [Mitsunaga and Nayar \[1999\]](#) empfohlen, die Kamerafunktion für alle Polynome bis zu einem definierten maximalen Grad zu bestimmen. Das Polynom für welches der Fehler ϵ der Lösung minimal ist, wird dann ausgewählt.

2.2. Modellierung der Kamerafunktion als Bézierkurve

Die Berechnung der Kamerafunktion als Bézierkurve ähnelt der Methode von [Mitsunaga and Nayar \[1999\]](#). Der wesentliche Unterschied besteht darin, dass die Kamerafunktion als Bézierkurve n -ten Grades

$$g(Z_{i,j}) = \sum_{i=0}^n B_{i,n}(Z_{i,j})P_i \quad (15)$$

mit den Bernsteinpolynomen

$$B_{i,n}(t) = \binom{n}{i} t^i (1-t)^{n-i} \quad (16)$$

für $Z_{i,j} \in [0, 1]$ approximiert wird.

Die Fehlerfunktion wird nach den selben Prinzipien aufgestellt und hat somit folgende Form:

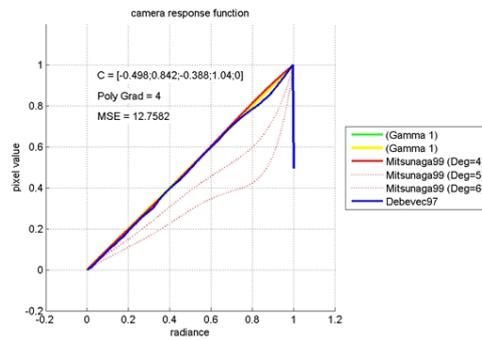
$$\varepsilon = \sum_{q=1}^{Q-1} \sum_{p=1}^P \left[\sum_{i=0}^N B_{i,n}(Z_{p,q})P_p - R_{q,q+1} \sum_{i=0}^N B_{i,n}(Z_{p,q+1})P_p \right]^2 \quad (17)$$

Das Lösen des Gleichungssystems erfolgt nun ebenfalls wie im vorhergehenden Abschnitt beschrieben.

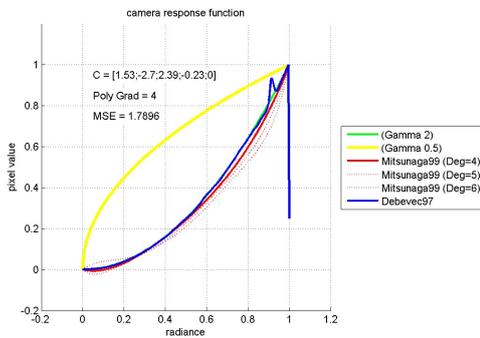
2.3. Ergebnisse

Mit allen drei vorgestellten Verfahren lässt sich für den Anwendungsfall eine ausreichend genaue Approximation der Kamerafunktion berechnen. In Abbildung 3 wurden anhand von Testdaten vier Kamerakurven berechnet. Erkennbar ist, wie die Polynomial berechneten Kurven ab Freiheitsgraden von über 5 anfangen, dem Funktionsverlauf nicht mehr exakt zu folgen. Es zeigt sich, dass das Verfahren von [Debevec and Malik \[2008\]](#) sehr stabil ist, bzw. alle Funktionen gut approximieren kann. Der Nachteil bei diesem Verfahren ist jedoch die vergleichsweise hohe Komplexität des Gleichungssystems. In dieser Hinsicht bieten die Polynom-basierten Verfahren Vorteile, da hier nur die Kontrollpunkte der Polynome berechnet werden müssen. Ein Problem der Polynom-basierten Verfahren ist es jedoch, den richtigen Freiheitsgrad auszuwählen. In [Mitsunaga and Nayar \[1999\]](#) wird dazu vorgeschlagen, die Berechnung der Kamerafunktion mit jedem Freiheitsgrad beginnend bei 1 bis zu einem vorgegebenen, maximalen Freiheitsgrad durchzuführen. Der Freiheitsgrad mit dem kleinsten Fehler (bezogen auf die Lösung des Gleichungssystems) wird dann als beste Lösung gewählt. Ein Vorteil der Bézier basierten Methode gegenüber [Mitsunaga and Nayar \[1999\]](#) liegt in einem geringeren Oszillationsverhalten, welches bei Polynomen höheren Grades auftreten kann.

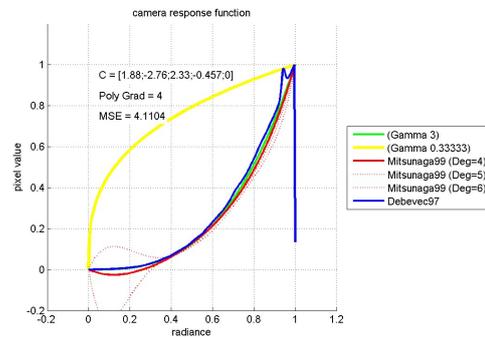
Als Optimierung des Verfahrens ist auch eine Stückweise Approximation der Kamerafunktion denkbar. Dies würde sich positiv auf das Oszillationsverhalten auswirken. Bei dieser stückweisen Approximation mittels Bézierkurven können zudem auf einfache Weise stetige Übergänge zwischen den Kurvensegmenten erreicht werden. Die stückweise Approximation kann auch auf die Form der Kamerafunktion eingehen. Beispielsweise könnte für die Mitte der Funktion ein großes Segment gewählt werden, da die Funktion dort typischerweise relativ gerade ist. An den Rändern der Kamerafunktion können dann kleinere Segmente gewählt werden, um den Funktionsverlauf dort genauer zu approximieren.



(a) Gamma 1



(b) Gamma 2



(c) Gamma 3

Abbildung 3: Approximation von Gammakurven als Beispiel für eine Kamerafunktion mittels der vorgestellten Verfahren

3. Schätzung des HDR-Histogramms

Ein Problem bei der Auswahl der Belichtungszeit eines Fotos, bzw. der Belichtungszeiten für eine Belichtungsreihe, die für eine HDR-Aufnahme benutzt werden kann, ist die Unkenntnis über die tatsächlichen Lichtverhältnisse der aufgenommenen Szene. Ein Großteil des Dynamikumfangs der natürlichen Szene geht durch die Digitalisierung und Quantisierung verloren. Die Histogramme der aufgenommenen (LDR) Bilder lassen deshalb keine Aussage über die tatsächlichen Lichtverhältnisse zu. Vor allem aber lassen sich zwei LDR-Histogramme der gleichen (statischen) Szene mit unterschiedlichen Belichtungszeiten auf Grund der Nichtlinearität der Kamerafunktion nicht vergleichen.

Ideal wäre es natürlich das Histogramm der realen Szene (bzw. dem HDR-Bild) zu kennen. Mit diesem wäre es möglich, geeignete Bereiche und Belichtungszeiten zu wählen, um möglichst viele Details der Szene einzufangen oder geeignete Bilder zur Erstellung einer HDR-Aufnahme zu bestimmen.

Verfahren nach Gallo et al. [2012]

In Gallo et al. [2012] wird ein Verfahren vorgestellt, welches auf der Linearisierung der aufgenommenen Bilder basiert. Mit ihm ist es möglich, aus einer LDR-Belichtungsreihe das originale Histogramm der Szene zu schätzen.

Wieder ausgehend von dem Kameramodell $Z = f(E\Delta t)$, bedeutet eine Änderung der Belichtungszeit Δt eine Verschiebung der Quantisierungsgrenzen im Logarithmischen: $\ln E = \ln f^{-1}(Z) - \ln \Delta t$.

Für eine Belichtungszeit Δt erhält man beispielsweise für das Zentrum des Quantisierungsintervalls für $Z = 0$ das Zentrum $\ln E_0 = \ln f^{-1}(0) - \ln(\Delta t)$ im irradiance Bereich.

Allgemein besteht somit folgender Zusammenhang zwischen den Quantisierungsgrenzen b^Z im digitalen Bereich und den entsprechenden Grenzen $b^{\ln E}$ im irradiance

Bereich:

$$b_{i,j}^{\ln E} = \ln f^{-1}(b_i^Z) - \ln \Delta t_j \quad (18)$$

für i Grenzen und j Belichtungszeiten. Die Quantisierung erfolgt im Regelfall mit 8 Bit, somit existieren 2^8 Intervalle, d.h. $i = 0..255$.

Der erste Schritt um das HDR-Histogramm zu schätzen besteht nun darin, die kumulativen LDR-Histogramme zu berechnen und sie mittels der Kamerafunktion in irradiance Histogramme F_j^L zu überführen. Ihre baselines lassen sich durch Gleichung (18) berechnen.

Die Abszisse des HDR-Histogramms $\{B_k\}$ ist die Vereinigung der Abszissen der LDR-Histogramme:

$$\{B_k\} = \bigcup_j \{b_{i,j}^{\ln E}\} \quad (19)$$

Das HDR-Histogramm setzt sich aus dem jeweiligen Maximum an allen Stellen $\{B_k\}$ aller LDR-Histogramme zusammen. Also die Hüllkurve aller kumulativen LDR-Histogramme.

Das HDR-Histogramm $P_{\ln E}$ erhält man nun durch Ableiten von F^H für $\{B_k\}$ (wegen der nicht homogenen Bin Größe).

4. Auswahl der Belichtungszeiten

Da nun (zumindest theoretisch) das Histogramm der Szene bekannt ist, hat man prinzipiell alle Informationen, um zwei geeignete Belichtungszeiten zu bestimmen.

Das im Folgenden vorgestellte Verfahren geht von einer guten Schätzung des HDR-Histogramms aus. Auf die Problematik, dass die Belichtungen, die zur Berechnung des Histogramms verwendet wurden, unter Umständen nur einen kleinen Teil der Dynamik der Szene abdecken konnten, wird in Kapitel 5 eingegangen.

4.1. Berechnung nach Gallo et al.

Mit Hilfe des HDR-Histogramms lässt sich die Helligkeitsverteilung der Szene nun gut schätzen. Gallo et al. [2012] stellen ein Verfahren vor, welches nur auf dem geschätzten Histogramm und der Kamerafunktion basiert.

Ziel ist es einerseits natürlich ein möglichst gutes Ergebnis zu erhalten. Das bedeutet eine Auswahl an Belichtungszeiten, welche den Helligkeitsumfang der Szene möglichst vollständig abbildet und so zu einem detaillierten HDR-Bild führt. Auf der anderen Seite soll die Anzahl der Belichtungen klein gehalten werden um Speicher und Rechenzeit zu sparen.

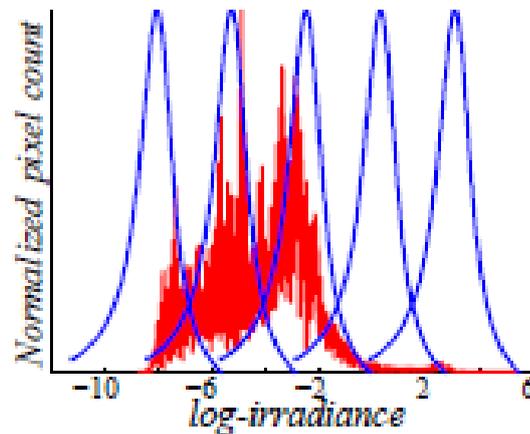


Abbildung 4

Eine triviale Möglichkeit die Belichtungszeiten zu wählen wäre es, wie in Abbildung 4 zu sehen, das Histogramm möglichst vollständig zu überdecken. Dies führt jedoch nicht zu einem optimalen Ergebnis Gallo et al. [2012]. Das Problem liegt wieder in der unterschiedlichen Quantisierung der Histogramm-Bins. Aufgrund der Form der Kamerakurve (siehe Abb. 2) sind die Bins Richtung Anfang und Ende kleiner als in der Mitte, wo die Kamerafunktion annähernd linear ist. Dies bedeutet, dass besonders Richtung Anfang und Ende eine Kompression stattfindet und damit auch die Quantisierungsfehler zunehmen. Eine optimale Belichtungswahl muss also auch diese Quantisierungsfehler in Betracht ziehen. In Abb. 5 ist eine solche (gegenüber Abb. 4 verbesserte) Auswahl zu sehen. In dem Diagramm ist jeweils die Kamerafunktion

⁴(Bei einer nicht linearen Kamerafunktion.)

überlagert dargestellt und die daraus resultierende, geschätzte Quantisierungsqualität.

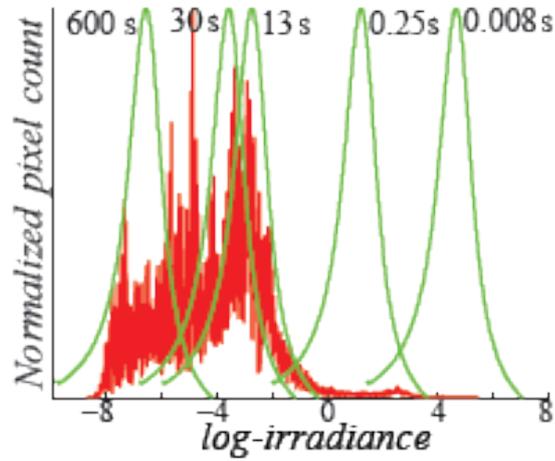


Abbildung 5

Diese Minimierung der Quantisierungsfehler stellt das Prinzip des Verfahrens in Gallo et al. [2012] dar. Die Güte der Belichtungsreihe wird mittels des Spitzen-Signal-Rausch-Verhältnis (PSNR) gemessen und ergibt das zu minimierende Fehlermaß.

Die Qualität einer Belichtungsreihe wird damit durch die Summe der PSNR-Werte der Einzelbelichtungen

$$\sum_{i=1}^N PSNR(\tilde{t}_i) \quad (20)$$

beschrieben. Falls diese minimal ist, ist die beste Belichtungsreihe gefunden.

Berechnung des Spitzen-Signal-Rausch-Verhältnisses (PSNR)

Neben der Quantisierung existieren noch andere Rauschquellen im Prozess der Bildgenerierung. Dies kann Signalabhängiges Rauschen wie Photon Shot Noise (PSN), oder Signalunabhängiges Rauschen wie readout oder thermal noise sein. Diese Fehlerquellen werden mittels eines Modells aus Foi et al. [2008] erfasst:

$$PG(X) = a \frac{X - X_{min}}{X_{max} - X_{min}} + b \quad (21)$$

Dabei bezeichnet X wieder die Belichtung mit $X = Et$, der scene irradiance E und der Belichtungszeit t . Die Parameter a und b sind Kameraabhängig⁵.

Im Logarithmischen:

$$PG(\tilde{X}) = \frac{a}{X_{max} - X_{min}} \exp(\tilde{E}) + b - \frac{a}{X_{max} - X_{min}} X_{min} \quad (22)$$

bzw.

$$PG(\tilde{X}) = a' \exp(\tilde{E}) + b' \quad (23)$$

mit $a' = a/(X_{max} - X_{min})$ und $b' = b - a'X_{min}$.

Die Steigung der Kamerafunktion lässt einen direkten Rückschluss auf die Qualität der Quantisierung zu. Je höher die Steigung ist, desto stärker ist die Komprimierung und damit der Quantisierungsfehler in dieser Region. D.h. die Größe eines Bins steht in keinem linearen Zusammenhang zur Größe eines Bins nach der Digitalisierung.

Damit bietet sich die Ableitung der Kamerafunktion als Basis für ein gutes Maß der Quantisierungsintervallgröße an. Die Größe des Intervalls i , zentriert in $\tilde{X}_i = f^{-1}(Z_i)$ kann folgendermaßen approximiert werden:

$$\Delta_i = \Delta / f'(\tilde{X}_i) \quad (24)$$

Δ bezeichnet dabei die Größe eines Quantisierungsintervalls in Z . Für 8-bit Werte in Z ist $\Delta = 2^{-8}$.

Der Quantisierungsfehler wird als lineare Funktion $e(\tilde{X}) = \tilde{X} - \tilde{X}_i$ mit $\tilde{X} \in [X_i + \Delta i/2, X_i - \Delta i/2]$ innerhalb eines Quantisierungsintervalls i betrachtet.

⁵Beschreibung...

Nach Gallo et al. [2012] kann die Varianz des Quantisierungsrauschens nun unter der Annahme der Gleichverteilung innerhalb des Intervalls durch

$$\sigma_Q^2(\tilde{X}_i) = \frac{1}{\Delta i} \int_{X_i - \Delta i/2}^{X_i + \Delta i/2} e^2(\tilde{X}) de = \frac{\Delta_i^2}{12} = \frac{1}{12} \left[\frac{\Delta}{f'(\tilde{X}_i)} \right]^2 \quad (25)$$

angegeben werden.

Diese Varianz kann nun mit der Varianz des Kamerarauschens zur Gesamtvarianz

$$\sigma_i^2 = PG(\tilde{X}_i) + \sigma_Q^2(\tilde{X}_i) \quad (26)$$

addiert werden⁶.

Das Spitzen-Signal-Rausch-Verhältnis für einen Quantisierungsintervall i ist dann

$$PSNR(\tilde{X}_i) = \frac{1}{PG(\tilde{X}_i) + \sigma_Q^2(\tilde{X}_i)}. \quad (27)$$

Das Rauschverhalten eines einzelnen Bins ist nun bekannt. Jedoch wird ein Fehlermaß gesucht, welches das Rauschverhalten einer Belichtung beschreibt. Durch Formulierung des Problems im Logarithmischen lässt sich dies für eine Belichtungszeit t und die scene irradiance \tilde{E} wegen $\tilde{X} = \tilde{E} + \tilde{t}$ recht einfach ausdrücken:

$$PSNR(\tilde{X}) = PSNR(\tilde{E} + \tilde{t}).$$

Die totale PSNR für eine Belichtungszeit entspricht der Korrelation aus dem Histogramm der logarithmierten scene irradiance \tilde{E} , $P_{\tilde{E}}(\tilde{E}_i)$ und dem PSNR für das entsprechende Intervall:

$$PSNR(\tilde{t}) = \sum_i P_{\tilde{E}} PSNR(\tilde{E}_i + \tilde{t}). \quad (28)$$

⁶Varianzen unabhängiger Variablen können addiert werden, auch wenn sie aus unterschiedlichen Verteilungen stammen Gallo et al. [2012].

Die Güte einer Belichtungsreihe kann damit durch

$$O(\{\tilde{t}_i\}_{i=1}^N) = \sum_{i=1}^N PSNR(\tilde{t}_i) \quad (29)$$

bewertet werden. Ein optimales Ergebnis liefert diejenige Menge an Belichtungen, die diese Gleichung maximiert.

5. HDR-Video

Im Gegensatz zu vergleichsweise statischen Szenen in der HDR-Fotografie, muss die Belichtungssteuerung eines HDR-Videosystems dazu in der Lage sein, sich an dynamische Szenen bzw. wechselnde Lichtverhältnisse, anzupassen. Dazu müssen die Szenenverhältnisse laufend analysiert werden. Dies kostet Zeit und kann zu einer Reduzierung der Framerate führen. Zudem werden mehrere Belichtungen benötigt, wozu für jeden Frame die Belichtungszeit angepasst werden muss. Hier wird mit zwei Belichtungen gearbeitet. Das verwendete Kamerasystem gibt zwei Belichtungszeiten vor, für welche die Kamera abwechselnd die Frames liefert. Ist die Szene sehr dynamisch oder ändern sich die Lichtverhältnisse sprunghaft, dürfen diese Änderungen nicht zu Sprüngen in den Belichtungszeiten führen, welche zu Flimmern oder anderen Artefakten im generierten HDR-Video führen könnten. Natürlich soll die Verzögerung im System so gering wie möglich gehalten werden, weshalb die Belichtungssteuerung dazu in der Lage sein soll, sich an die Dynamik der Szene zu adaptieren.

Dynamische Szenen bringen noch andere Probleme mit sich, wie z.B. Ghosting. Hier wird jedoch nur auf die reine Bestimmung der Belichtungszeiten eingegangen.

5.1. Belichtungssteuerung

Aufgrund der Beschränkungen des Kamerasystems und dem Anspruch, echtzeitfähig zu sein, wird hier mit zwei Belichtungszeiten gearbeitet. Grundsätzlich wird das Verfahren von Gallo et al. [2012] angewandt, welches für zwei Belichtungszeiten

angepasst wird. Dadurch kann zum Beispiel das Optimierungsproblem vereinfacht werden. Die Fehlerfunktion wird nur für ein bestimmtes Fenster um die aktuellen Belichtungszeiten ausgewertet, um große Sprünge zu vermeiden.

Daher gilt es jeweils den PSNR-Wert der beiden Belichtungen zu maximieren. Das bedeutet für die Belichtungszeit $k = 1, 2$ wäre die neue optimale Belichtungszeit $t_{0k} = \sum_i P_E(E_i) PSNR(E_i t_k)$.

Die folgende Funktion beschreibt dann eine optimale Belichtungseinteilung basierend auf den PSNR-Werten für zwei Belichtungen und dem Histogramm der Szene:

$$\min_{t_1, t_2} f(t_1, t_2) = \sum_i (\max(PSNR(E_i t_1), PSNR(E_i t_2))) P_E(E_i) \quad (30)$$

Von den beiden PSNR-Werten der Belichtungszeit 1 und 2 wird jeweils das Maximum mit dem Histogramm der Szene gewichtet.

Dabei ist wichtig, dass die Bilder zur Berechnung des Histogramms nicht überbelichtet sind. Durch die Überbelichtung gehen bei der Berechnung viele Informationen verloren, die bei nur zwei Bildern schnell zu einer schlechten Schätzung führen. Als Maßzahl für den Grad der Überbelichtung, wird das Verhältnis des jeweiligen, aufsummierten LDR-Histogramm zum letzten Bin des Histogramms herangezogen: $w = (b_{255}^Z) / \sum_i b_i^Z$. Für w kann ein Schwellenwert (beispielsweise 10 %) festgelegt werden, was dann die Grenze für die neu zu berechnenden Belichtungszeiten darstellt. Überbelichtete Bilder werden somit abgewertet. Auch braucht dann nur noch in eine Richtung nach besseren Belichtungszeiten gesucht werden.

Da der Algorithmus zu Berechnung der Belichtungszeiten vollständig auf dem HDR-Histogramm basiert, ist er robust gegenüber kleineren Veränderungen der Einzelbilder. Um Sprünge in den Belichtungszeiten abzufedern, werden die aktuellen Belichtungszeiten schrittweise an die neu Berechneten angepasst.

$$t_{neu} = a * t_i + (1 - a) * t_0 \quad (31)$$

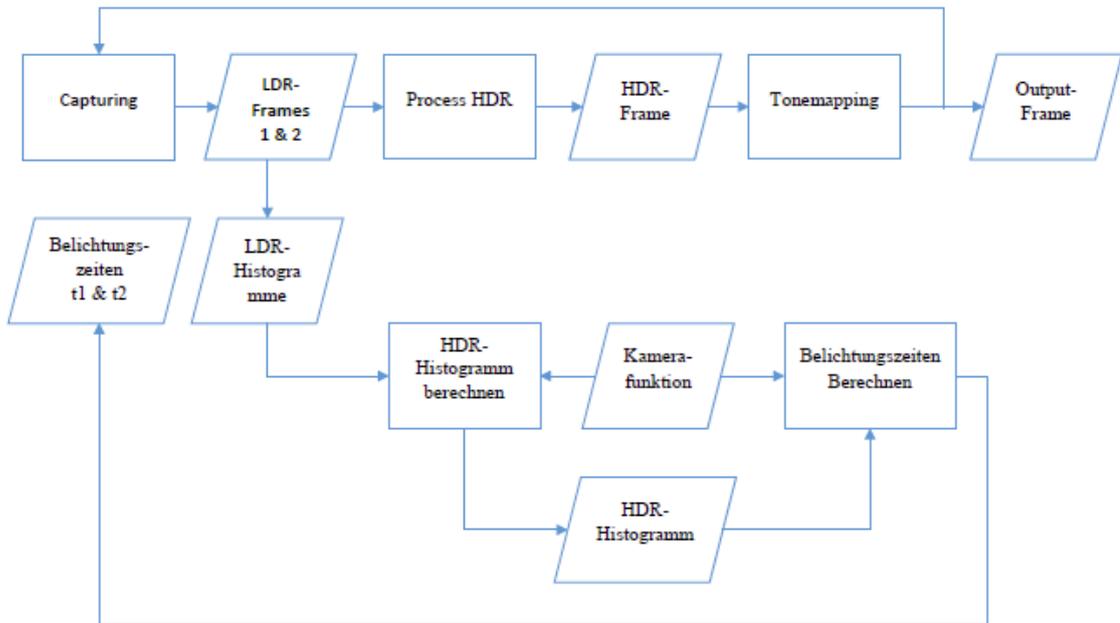


Abbildung 6: Übersicht Belichtungssteuerungssystem

6. Ergebnisse

Nachfolgend werden einige Bildsequenzen dargestellt, in der die Adaption des Belichtungssystem an die jeweiligen Lichtverhältnisse erkennbar ist. Die Bildsequenzen sind jeweils mit dem im Rahmen dieser Arbeit entwickelten Programm aufgenommen wurden. Es ist in C++ implementiert und arbeitet in dieser Konfiguration mit einer Basler Kamera. Die Belichtungszeiten der Kamera werden durch einen Mikrocontroller getriggert, welcher von dem Programm über eine USB-Schnittstelle gesteuert wird. Durch das Triggersignal, welches der Mikrocontroller erzeugt, nimmt die Kamera abwechselnd ein Bild mit der ersten und zweiten Belichtungszeit auf. Der Videostream der Kamera wird nun von der Software empfangen und in einen Stream mit der ersten und einen mit der zweiten Belichtungszeit aufgetrennt.

In Abbildung 7 ist der Beginn einer Belichtungssequenz mit zufälligen Belichtungszeiten⁷ abgebildet. Das Histogramm in der unteren Bildhälfte stellt die Schätzung des HDR-Histogramms aus den beiden oberen Einzelbildern dar. Die LDR-Histogramme

⁷Beispielsweise stammen die Belichtungszeiten aus einer vorhergehenden Aufnahme, passen also nicht zu der gegenwärtigen Szene.

der beiden Einzelbilder sind auf der rechten Seite abgebildet. Man erkennt eine starke Sättigung beider Bilder und sie unterscheiden sich hinsichtlich ihrer Belichtungszeiten nur gering, weshalb auch die Schätzung des HDR-Histogramms zunächst noch relativ ungenau ist. Der Himmel ist z.B. nur eine weiße Fläche und die Fassade des Hochhauses ist nicht gut zu erkennen. Abbildung 8 zeigt nun das Ende dieser Belichtungssequenz. Das Aufnahmesystem hat sich nach ca. 100 Einzelbildern an die Lichtverhältnisse der Szene adaptiert und 0,068 s bzw. 0,524 s für Bild 1 bzw. Bild 2 ausgewählt. Diese Einstellungen bleiben konstant, solange sich die Lichtverhältnisse der Szene nicht ändern. In den Bildern sind nun sehr gut Details zu erkennen. In Bild 1 sind Details in den zu Beginn überbelichteten Bereichen (wie die Wolke am Himmel und die Fassade des Hochhauses) und in Bild zwei Details in den dunklen Bereichen (wie die Fassade des kleineren Gebäudes und die Bäume).

Im zweiten Beispiel lässt sich ein ähnlich gutes Ergebnis beobachten. Abbildung 9 ist der Beginn einer Belichtungsreihe einer Innenaufnahme zu sehen. Die Belichtungszeiten sind zu Beginn bei 0,02 s und 0,05 s. Die beiden Bilder unterscheiden sich hinsichtlich ihrer Belichtung nicht sehr stark und es sind in den schattigen Bereichen kaum Details zu erkennen. Nach kurzer Zeit hat sich das Belichtungssystem dann an die Belichtungszeiten 0,05 s und 0,5 s adaptiert. In Bild 1 sind Details in den hellen Bereichen der Szene (wie in den Reflexionen auf dem Tisch) klar zu erkennen. In Bild 2 hingegen sind nun die Details in den Dunklen Bereichen unter dem Schreibtisch gut zu sehen. Auch das HDR-Histogramm stellt nun eine wesentlich genauere Schätzung der Szenenverhältnisse dar.

7. Fazit

Insgesamt zeigt sich, dass die Kombination und der Ausbau vorhandener Methoden zu einem funktionierenden Belichtungssystem geführt hat. Da das Verfahren nur auf den Histogrammen basiert, ist es relativ unempfindlich gegenüber kleineren Veränderungen der Szene. Das System kann sich auf variable Lichtverhältnisse adaptieren und ist echtzeitfähig. Die Berechnung eines HDR-Bildes bzw. Videostreams ist nicht Teil dieser Arbeit. Jedoch kann nun ein Standardverfahren zur HDR-Berechnung auf die Bilder der zwei Belichtungszeiten angewendet werden. Vorteilhaft dafür ist,

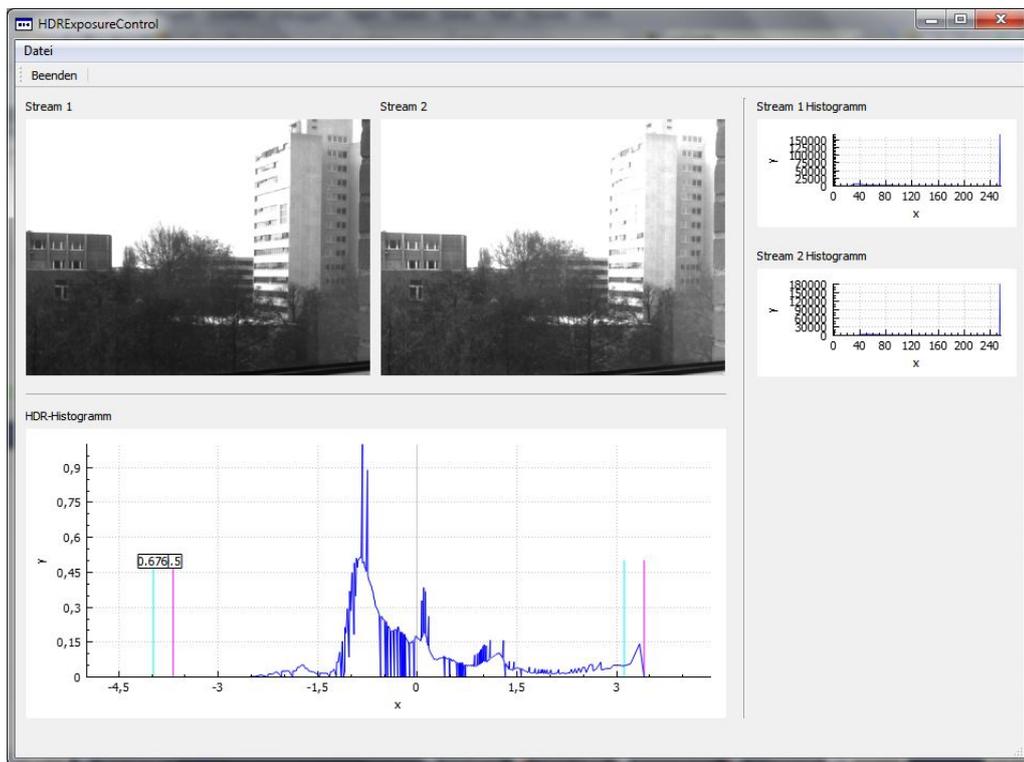


Abbildung 7: Beispiel 1, Belichtungen zu Beginn: 0,5 s und 0.676 s

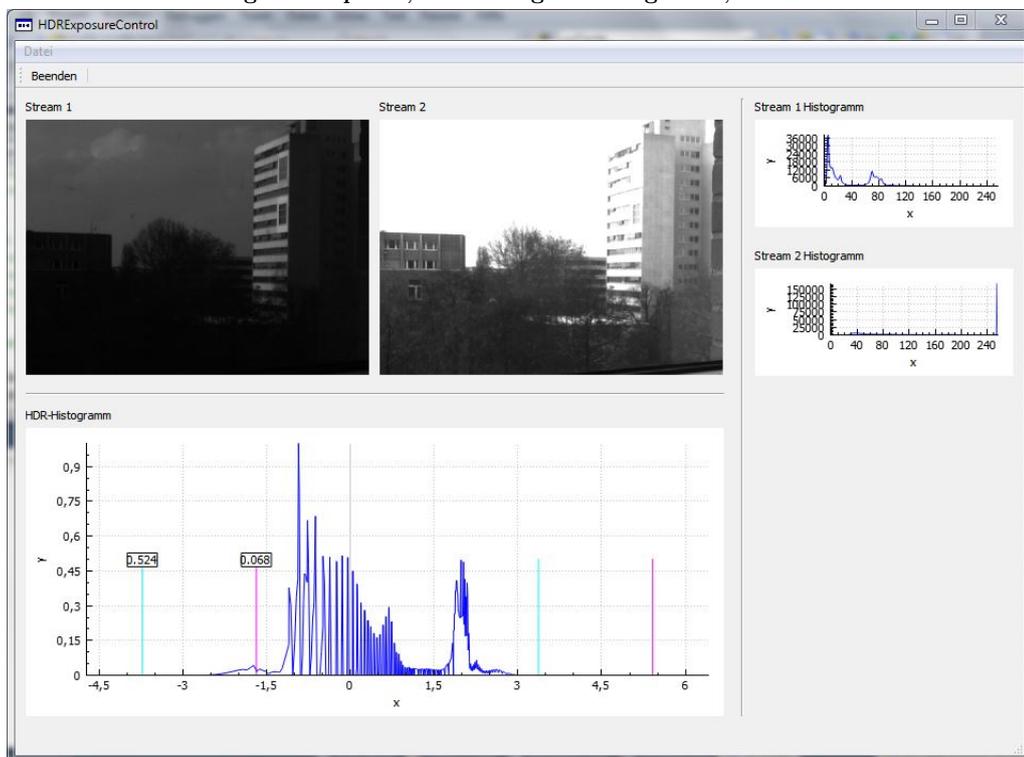


Abbildung 8: Beispiel 1, Belichtungsauswahl: 0,068 s und 0.524 s

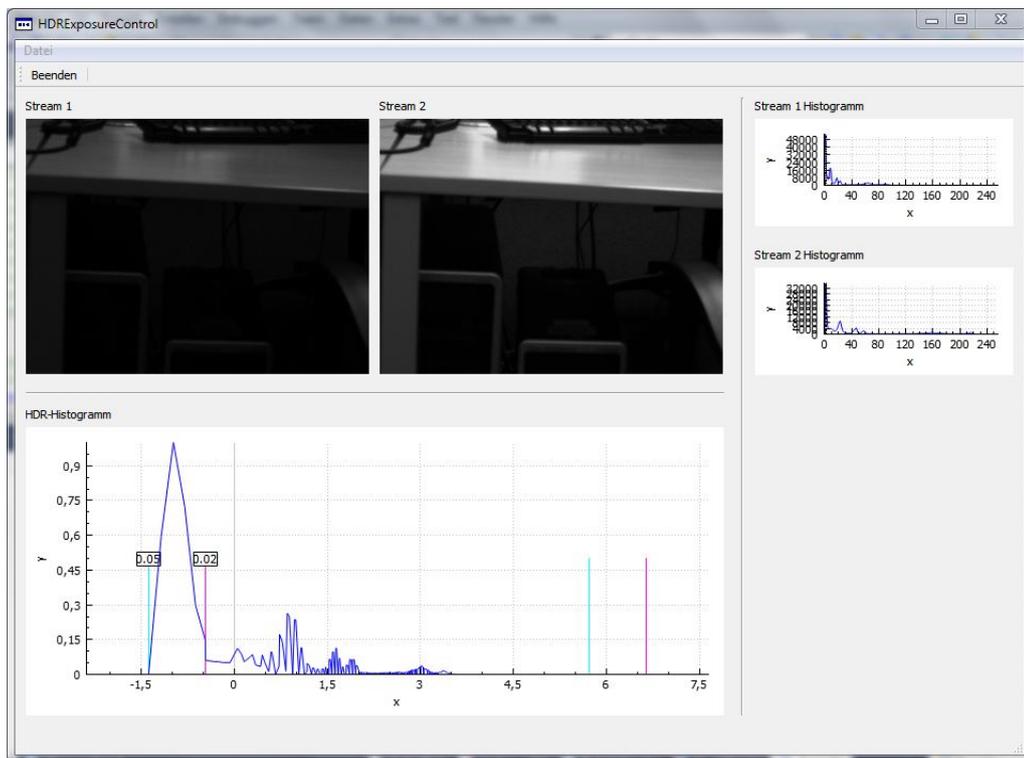


Abbildung 9: Beispiel 2, Belichtungen zu Beginn: 0,02 s und 0,05 s

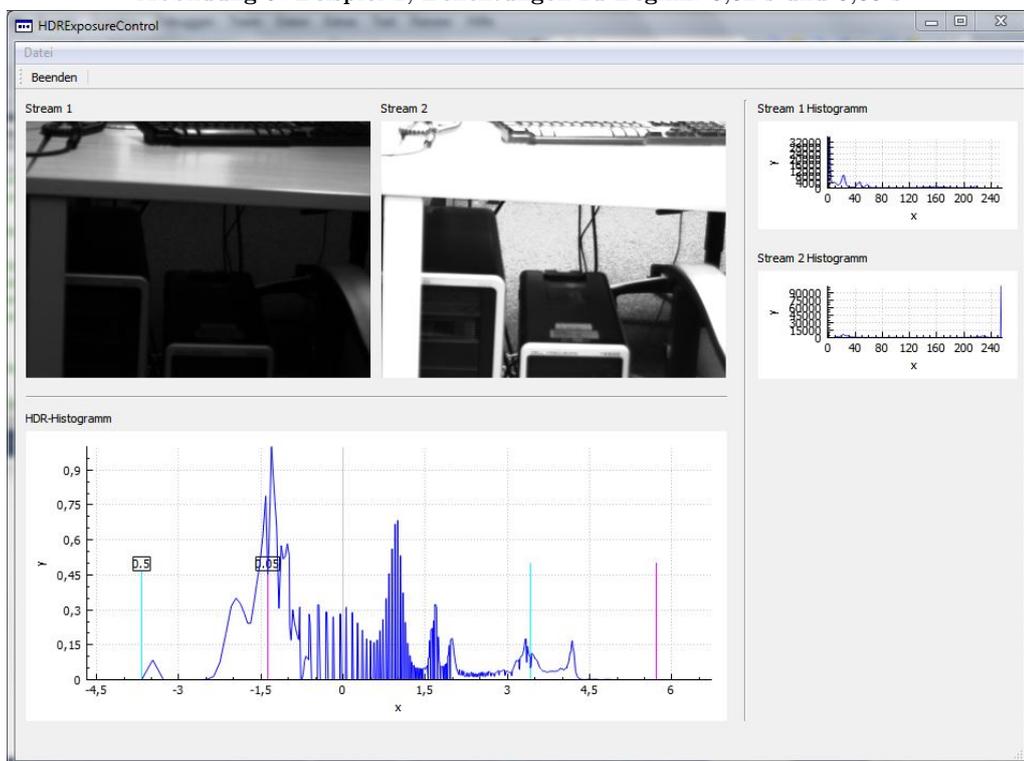


Abbildung 10: Beispiel 2, Belichtungsauswahl: 0,05 s und 0,5 s

dass bereits eine Schätzung über das HDR-Histogramm der Szene vorliegt. Optimierungspotential gibt es noch bei extremen Belichtungssituationen. Des weiteren können noch Techniken wie Bildregistrierung der beiden Belichtungen integriert werden, um Artefakte durch Kamerabewegungen und schnell veränderliche Szenen zu minimieren.

A. Literatur

- F. Benterle, A. Artusi, K. Debattista, and A. Chalmers. *Advanced High Dynamic Range Imaging*. A K Peters. Ltd., 2011.
- Paul E. Debevec and Jitendra Malik. Recovering high dynamic range radiance maps from photographs. In *ACM SIGGRAPH 2008 classes*, SIGGRAPH '08, pages 31:1–31:10, New York, NY, USA, 2008. ACM. doi: 10.1145/1401132.1401174. URL <http://doi.acm.org/10.1145/1401132.1401174>.
- A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazar-Ian. Practical poissonian-gaussian noise modeling and fitting for single-image raw-data. *IEEE Trans. Image Processing* 17, pages 1737–1754, 2008.
- O. Gallo, M. Tico, R. Manduchi, N. Gelfand, and K. Pulli. Metering for exposure stacks. *Computer Graphics Forum (Proceedings of Eurographics)*, 31:479–488, 2012.
- T. Mitsunaga and S.K. Nayar. Radiometric self calibration. In *Computer Vision and Pattern Recognition, 1999. IEEE Computer Society Conference on.*, volume 1, pages –380 Vol. 1, 1999. doi: 10.1109/CVPR.1999.786966.
- B. Quehl. Konzeption und implementierung einer mikrocontrollerbasierten belichtungssteuerung zur generierung von hdr-video. Diplomarbeit, Fachhochschule Brandenburg, 2008.
- E. Reinhard, G. Eard, S. Pattanaik, P. Debevec, W. Heidrich, and K. Myszkowski. *High Dynamic Range Imaging*. Elsevier Inc., 2010.