

Institut für Parallele und Verteilte Systeme
Universität Stuttgart
Universitätsstraße 38
D-70569 Stuttgart

Diplomarbeit Nr. 3369

Analyse und Optimierung der Farbaufnahme für einen 3D Laser Scanner

Timo Dritschler

Studiengang:	Informatik
Prüfer:	PD. Dr. rer. nat. Michael Schanz
Betreuer:	Dr. rer. nat. Martin Ossig, Dipl.-Inf. Alexander Kramer
begonnen am:	10. Juli 2012
beendet am:	13. Dezember 2012
CR-Klassifikation:	C.3, I.4.1

Zusammenfassung

Das menschliche Verständnis von Raum und Struktur ist stark an die Wahrnehmung von Farbe gebunden. Aus diesem Grund ist die Farbqualität von Fotoaufnahmen heutzutage von enormer Wichtigkeit. In dieser Arbeit soll der Farbaufnahmeprozess des FARO Laser Scanners Focus 3D dahingehend untersucht und optimiert werden, dass möglichst ideale Farbbilder aufgenommen werden können. Dazu müssen verschiedene Schwierigkeiten, die sich aus dem technischen Hintergrund des Farbaufnahmeprozesses ergeben, berücksichtigt werden. Darüber hinaus müssen verschiedenen Rahmenbedingungen eingehalten werden, die sich aus der späteren Verwendung der erhaltenen Farbaufnahmen ergeben. So wird zum Beispiel nicht nur mit einzelnen Aufnahmen, sondern mit einer ganzen Reihe von Aufnahmen gearbeitet, die anschließend zu einem Gesamtbild zusammengesetzt werden müssen. Auch der Farbeindruck, die empfundene Farbhelligkeit und der erkennbare Detailgehalt der Aufnahmen spielen eine entscheidende Rolle.

Zur Optimierung des Farbaufnahmeprozesses wird zuerst der bereits bestehende Prozess auf Mängel untersucht und diese möglichst behoben. Anschließend wird ein Prozess zur weiteren Verbesserung der Bildqualität unter schwierigen Aufnahmebedingungen konzipiert und implementiert.

Inhaltsverzeichnis

1. Einleitung	7
2. Überblick und Aufgabenstellung	9
3. Analyse des bestehenden Farbaufnahmeprozesses	15
3.1. Festgestellte Problematik	17
4. Bestimmung einer optimalen Belichtungszeit	21
4.1. Zusätzlicher Messdurchlauf	23
4.2. Ermittlung der optimalen Einzelbelichtung anhand der Messdaten	25
4.3. Verbleibende Problematik	26
5. High Dynamic Range Imaging	27
5.1. Technische Grundlagen	29
5.1.1. HDR Aufnahme	29
5.1.2. Tone Mapping	33
5.2. Auswahl der HDR Verfahren	38
5.2.1. HDR Synthese Verfahren	38
5.2.2. Auswahl des Tonemapping Verfahrens	40
5.3. Reinhard's Tonemapper	41
6. Integration der ausgewählten Verfahren	47
7. Zusammenfassung und Ausblick	53
A. Anhang	59
A.1. Anhang 1	59
A.1.1. Umrechnung von sRGB nach CIE xyY	59
A.1.2. Rückumrechnung von CIE xyY nach sRGB	59
A.2. Anhang 2	60
A.2.1. Analytische Näherung der Parameter für Reinhard's Tonemapper	60
Literaturverzeichnis	61

Abbildungsverzeichnis

2.1.	FARO Laser Scanner Focus 3D	9
2.2.	Skizze der Laser Projektion beim Scannen	10
2.3.	Scanausschnitt eines Freskos im Ahnenzimmer von Schloss Lichtenstein	10
2.4.	Skizze des Optik Aufbaus des Scanners	11
2.5.	Farbscan mit Farbverlaufs-Problem	11
3.1.	Beispielhaftes Einzelbild	15
3.2.	Benutzbarer Bereich eines Einzelbildes	16
3.3.	Schmiereffekt bei Spiegeldrehung	17
3.4.	Vergleichsscans mit problematischen Lichtsituationen	18
4.1.	Aufbau um die Dauer der Belichtungsmessung zu bestimmen	21
4.2.	Messergebnisse für die Dauer der automatischen Belichtungsmessung	22
4.3.	Messungen für Belichtungszeit mit verschiedenen vielen Positionen	24
4.4.	Bipartite Messreihe mit ungünstigem Median	25
5.1.	Belichtungszeit im Verhältnis zum Dynamikumfang einer Szene	27
5.2.	Beispiel einer ungeglätteten Response Curve	32
5.3.	Beispiel einer geglätteten Response Curve	33
5.4.	Exemplarisches Tonemapping von Grauwerten	35
5.5.	Halo Effekte beim Tonemapping	36
5.6.	Globales Tonemapping	37
5.7.	Unschärfe bei der Erzeugung von Radiancemaps	39
5.8.	Adams Zonensystem	42
5.9.	Einfluss des Key Values auf das Tonemapping	43
5.10.	Grundgestalt von Reinhard's Tonemapping Funktion	44
5.11.	Einfluss des Weißpunkts auf die Gestalt von Reinhard's Tonemapper	45
7.1.	Schlechtes Belichtungseinstellung bei der alten Methode	53
7.2.	Resultat der neuen Methode zur Belichtungsmessung	54
7.3.	Resultat des Tonemappings 1	56
7.4.	Resultat des Tonemappings 2	57

1. Einleitung

Der FARO Laser Scanner Focus 3D ist ein Punktwolken-Messgerät für die dreidimensionale und berührungslose Vermessung von Objekten im Innen- und Außenbereich. Für die berührungslose Vermessung kommt ein phasenmodulierter Laser zum Einsatz, der neben der Distanz zum reflektierenden Objekt auch eine Reflektivitätsstärke ermittelt. Auf Grund der monochromatischen Eigenschaften des Lasers kann durch die gemessenen Reflektivitätswerte lediglich ein Grauwertbild erzeugt werden. Deshalb wurde im FARO Laser Scanner Focus 3D, in der optischen Achse des Laserstrahls, eine Farbkamera integriert, die quasi parallaxenfrei Farbinformationen des zu vermessenden Objekts aufnehmen kann. Um für einen Rundumscan genügend Farbinformationen zwecks Überlagerung mit einem 3D Scan zu erhalten, werden in einem separaten Farbaufnahmeprozess bis zu 85 Farbaufnahmen gemacht, die anschließend zu einem Panoramabild zusammengefasst werden. Die Herausforderung hierbei ist, eine geeignete Einstellung der Farbkamera zu ermitteln, so dass das entstehende Panoramabild keine über-, bzw. unterbelichteten Bereiche zeigt und keine Farbunterschiede gleicher Objekte aus verschiedenen Bildern zu erkennen sind.

Zur Optimierung des Farbaufnahmeprozesses wird zuerst der bereits bestehende Prozess auf Mängel untersucht und diese anschließend behoben. Danach wird ein Prozess zur weiteren Verbesserung der Bildqualität unter schwierigen Aufnahmebedingungen konzipiert und implementiert.

Gliederung

Diese Arbeit ist folgendermaßen gegliedert:

Kapitel 2 – Überblick und Aufgabenstellung: Hier werden auf das Umfeld und die Problemstellung dieser Arbeit eingegangen, sowie einige technische Grundlagen erläutert.

Kapitel 3 – Analyse des bestehenden Farbaufnahmeprozesses: In diesem Kapitel wird der bereits bestehende Farbaufnahmeprozess auf Probleme untersucht.

Kapitel 4 – Bestimmung einer optimalen Belichtungszeit: Untersucht die Lösungsansätze aus Kapitel 3 und geht auf deren Umsetzung und Integration ein.

Kapitel 5 – High Dynamic Range Imaging: In diesem Abschnitt werden die technischen Grundlagen des High Dynamic Range Imaging genauer erläutert, so wie Methoden ausgewählt, um die Bildqualität des Farbaufnahmeprozesses weiter zu verbessern.

Kapitel 6 – Integration der ausgewählten Verfahren: Gibt einen Überblick über die Integration des vorher entwickelten Konzepts zur Verwendung von High Dynamic Range Imaging.

Kapitel 7 – Zusammenfassung und Ausblick: Abschließend werden die Ergebnisse dieser Arbeit zusammengefasst und ein Ausblick über weitere, denkbare Optimierungen gegeben.

2. Überblick und Aufgabenstellung

Die dreidimensionale Dokumentation von Objekten und Strukturen hat durch die sich immer weiter entwickelnden Technologien auf diesem Bereich zunehmend an Bedeutung gewonnen. So werden Bauteile oft nicht mehr nur auf dem Papier spezifiziert, sondern zusätzlich oder teils ausschließlich mit entsprechender Software am Computer modelliert. Diese Digitalisierung von Bauteilen und Objekten ermöglicht es, komplexe Strukturen in räumlichem Kontext am Computer darzustellen und macht es einem Nutzer somit bedeutend einfacher, Zusammenhänge zu begreifen und vereinfacht dadurch den Umgang mit ihnen erheblich.



Abbildung 2.1.: FARO Laser Scanner Focus 3D

Der FARO Laser Scanner Focus 3D ist ein berührungsloses Messgerät für die Erzeugung von dreidimensionalen Modellen in Form von Punktwolken, das zusammen mit der FARO SCENE Software, die zur Darstellung und Verarbeitung der gewonnen 3D Daten entwickelt wurde, ein Gesamtpaket zur dreidimensionalen Dokumentation von Räumen und Objekten bildet.

2. Überblick und Aufgabenstellung

Zur Vermessung und Erzeugung von Scandaten kommt im Focus 3D ein modulierter Laser zum Einsatz, der aus dem Gerät heraus auf einen 45° gekippten Drehspiegel projiziert wird, der sich um die horizontale Längsachse des Scanners dreht. Dadurch wird der Laserstrahl vertikal den Scan-Bereich entlang geführt. Zudem dreht sich das Gerät selbst um die vertikale Achse und führt dadurch den Strahl horizontal entlang des Scan-Bereichs. Dies ermöglicht es dem Focus 3D einen Bereich von 305° vertikal und 360° horizontal mit einer Auflösung von etwa 710 Millionen Scanpunkten abzudecken.

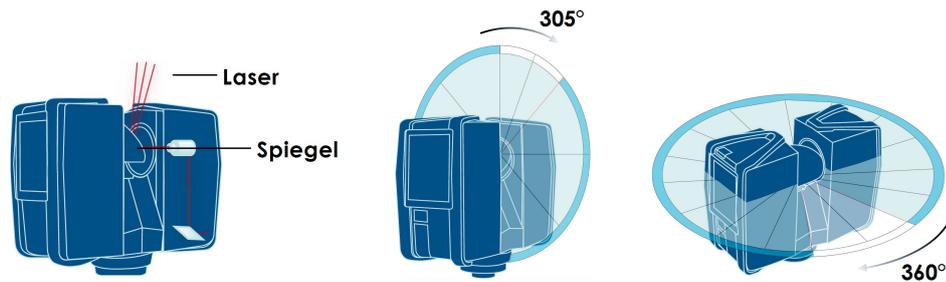


Abbildung 2.2.: Skizze der Laser Projektion beim Scannen

Bei der Vermessung durch den Focus 3D werden neben den gemessenen Distanzinformationen auch Reflektionsstärken des eingesetzten Lasers ermittelt, mit deren Hilfe die erzeugten 3D Daten monochrom eingefärbt werden, um ein besseres räumliches Verständnis der ermittelten Messdaten zu erhalten. Da das menschliche Verständnis von Form und Struktur jedoch zu einem großen Teil auf Farbwahrnehmung basiert, hat die Hinzunahme von Farbinformationen in digitale Modelle einen nicht zu vernachlässigenden Nutzen für deren Verständlichkeit. Aus diesem Grund wurde im FARO Laser Scanner Focus 3D eine Farbkamera in der optischen Achse des Lasers integriert, deren Aufnahmen dazu genutzt werden um sie mit dem aufgenommenen Scan zu überlagern und so eine Einfärbung der sonst nur monochromen Scanpunkte zu erreichen.



Abbildung 2.3.: Scanausschnitt eines Freskos im Ahnenzimmer von Schloss Lichtenstein. Speziell die Zierden um das Fresko und die blau lackierten Kacheln des Kamins (links im Bild) sind im eingefärbten Scan deutlich klarer zu erkennen.

Zur Aufnahme der Farbbilder ist die integrierte Farbkamera auf den Drehspiegel in der Mittelachse des Scanners gerichtet, der nach dem Scannen, in einem separaten Farbaufnahme Prozess, schrittweise so ausgerichtet wird, dass die Kamera den gesamten Scanbereich ausreichend mit Bildern abdecken kann. Dabei werden bis zu 85 Farbaufnahmen gewonnen, die in der FARO SCENE Software zur Einfärbung der 3D Scanpunkte verwendet werden.

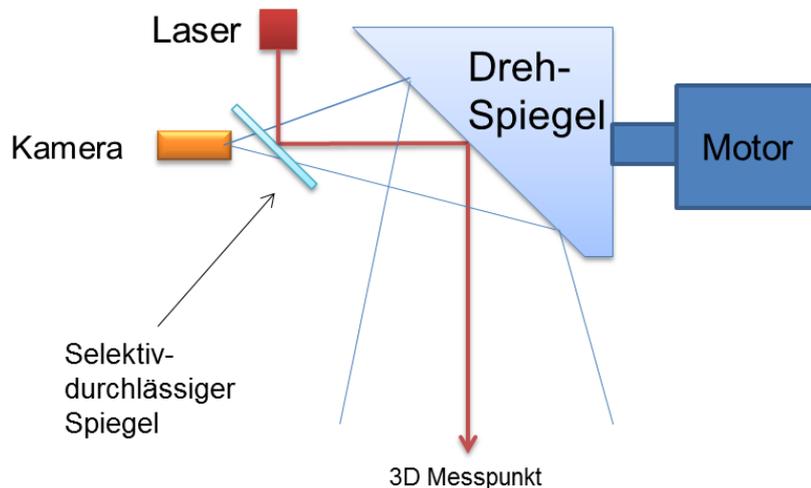


Abbildung 2.4.: Skizze des Optik Aufbaus des Scanners

Da für die Einfärbung von Scandaten also mehrere Bilder zu einem Farbscan zusammengefügt werden müssen, ergeben sich hieraus auch die typischen Probleme, die bei der Zusammensetzung von Einzelbildern zu beobachten sind. Verändern sich die Farbinformationen von Objekten, die auf verschiedenen Bildern gleichermaßen zu erkennen sind, ergeben sich daraus deutliche Farbverläufe in den eingefärbten Scanpunkten. Dadurch entsteht die Herausforderung, geeignete Einstellungen für die Farbkamera zu finden, so dass gleiche Objekte und Oberflächen die auf verschiedenen Bildern zu sehen sind, möglichst einheitlich von der Kamera abgebildet werden.



Abbildung 2.5.: Bei diesem Scan ist während der Aufnahme eine Wolke am Himmel vorbeigezogen und hat kurzzeitig die Sonne abgedunkelt. Die bei der Zusammensetzung der Einzelbilder entstehenden Farbanten sind deutlich zu erkennen.

2. Überblick und Aufgabenstellung

Anmerkung: Wir gehen im kommenden Verlauf dieser Abhandlung ausschließlich von einer Digitalkamera und ihrem elektronischen Farbsensor aus. Die jeweils genannten Zusammenhänge lassen sich jedoch auch analog auf eine klassische fotochemische Kamera mit entsprechendem Farbfilm übertragen.

Um eine sinnvolle Einstellung für die Farbkamera finden zu können, muss zuerst verstanden werden, welche Mechanismen und Prinzipien an einer Abbildung von einer Szene in der Echtwelt auf ein entsprechendes (digitales) Bild beteiligt sind. Dazu betrachten wir die Beschreibung der Abbildung des „Farbreizes“ eines Punktes N im Raum auf einen letztendlichen Farbwert V_N im Pixel-Raum der Farbaufnahme.

$$(2.1) \quad V_N = f(S_N)$$

Hierbei stellt S_N den Farbreiz des Punktes N dar, also die tatsächliche Strahlung, die vom Punkt N auf den Sensor der Kamera fällt. Dieser Farbreiz wird über die Funktion $f()$ von der Kamera auf einen letztendlichen Farbwert V_N im resultierenden Bild abgebildet. Wir bezeichnen $f()$ als die Abbildungsfunktion der Kamera. Die genaue Gestalt dieser Abbildungsfunktion hängt von einer Vielzahl von kamera- und herstellerspezifischen Faktoren ab und ist in der Regel nicht bekannt. Sie wird bevorzugt so eingestellt, dass ein Farbreiz, der von einem menschlichen Betrachter als neutralgrau empfunden wird, möglichst auf einen Farbwert V_N mit gleicher Farbempfindung abgebildet wird. Hierfür besitzen die meisten Kameras interne, automatische Mechanismen, die diese Einstellung vornehmen. Ein typisches Beispiel für solche internen Automatismen ist der automatische Weißabgleich.

Der eingehende Farbreiz S_N kann weiterhin näherungsweise beschrieben werden als das Verhältnis:

$$(2.2) \quad S_N = \frac{E v_N}{k^2} * t$$

$E v_N$ entspricht hierbei der Leuchtdichte des Punktes N und t der Belichtungszeit des Kamerasensors in Sekunden, d.h. der Dauer, mit der der Sensor der Einstrahlung durch den Punkt N ausgesetzt wird. Diese Einstrahlung wird zusätzlich durch die Einstellung der Blendenöffnung k gedämpft, welche die Fläche der Öffnung steuert, durch die Strahlung auf den Sensor trifft. Hierbei beschreibt k die sogenannte „Blendenzahl“¹. Je kleiner die Blendenöffnung eingestellt ist, desto weniger Strahlung trifft auf den Sensor. Halbiert sich die Fläche der Blendenöffnung, verdoppelt sich die Blendenzahl k .

¹<http://de.wikipedia.org/wiki/Blendenzahl>

Aus diesen Formeln erkennt man die vier grundlegenden Faktoren, die den resultierenden Farbwert V_N bei der Bildaufnahme beeinflussen können:

- Leuchtdichte $E v_N$
- Blendenöffnung bzw. Blendenzahl k
- Belichtungszeit t
- Abbildungsfunktion $f()$

Die Blende der im Focus 3D verbauten Kamera ist nicht einstellbar und besitzt eine konstante Blendenzahl von 2.4. Sie ist somit zwar am Abbildungsprozess beteiligt, hat auf die nachfolgenden Betrachtungen in dieser Abhandlung aber keinen weiteren Einfluss. Die verbleibenden Faktoren – Abbildungsfunktion, Leuchtdichte und Belichtungszeit – sind also die maßgeblich ausschlaggebenden Parameter für das Resultat der Bildaufnahme im Focus 3D und liegen somit, unter anderem, im Schwerpunkt der weiter folgenden Betrachtungen.

Um zu gewährleisten, dass gleiche Objekte auf verschiedenen Bildern keine Farbunterschiede aufweisen, muss also gelten, dass der Farbwert $V_{N,p1}$ in Bild $p1$, des Echtweltpunktes N gleich dem Farbwert $V_{N,p2}$ in Bild $p2$, desselben Echtwelt-Punktes ist. Da $E v_N$ somit bei beiden Aufnahmen identisch ist (da es sich um denselben Echtwelt-Punkt handelt), bleiben nur noch $f()$ und t als Parameter von $V_{N,p1}$ und $V_{N,p2}$ übrig. Hieraus folgt direkt, dass, um die Voraussetzung $V_{N,p1} = V_{N,p2}$ zu erfüllen, $t1$ und $t2$ so wie die Abbildungsfunktion $f()$ in beiden Aufnahmen identisch sein müssen. Die Belichtungszeit einer Aufnahme ist in der Regel vom Benutzer einstellbar. Wir können also davon ausgehen, dass wir in der Lage sind die Bedingung $t1 = t2$ sicher zu stellen. Nun muss nur noch dafür gesorgt werden, dass die Abbildungsfunktion $f()$ in beiden Aufnahmen identisch ist.

Die Abbildungsfunktion $f()$ einer Kamera setzt sich zumeist zusammen aus einer Menge an internen Parametern und technisch bedingten Eigenschaften wie beispielsweise Nichtlinearitäten bei der Analog-Digital Wandlung, wie sie in Digitalkameras geschieht, oder des Entwicklungsprozesses des Farbfilms bei fotochemischen Kameras. Oft sind diese Nichtlinearitäten vom Hersteller so angepasst, dass die resultierende Aufnahme ein visuell möglichst ansprechendes Resultat liefert und wir können davon ausgehen, dass viele der Abbildungseigenschaften von $f()$ von uns nicht verändert werden können. Zwei der veränderbaren Eigenschaften von $f()$ sind üblicherweise der Weißabgleich und die Schwarzstufe. Der Weißabgleich gibt dabei an, welcher eingehende Farbeindruck aus der Szene auf reines Weiß bzw. auf ideales Neutralgrau abgebildet werden soll und die Schwarzstufe gibt an, ab welcher eingehenden Signalstärke ein Farbreiz nicht länger auf pures Schwarz abgebildet werden soll. Gehen wir also weiterhin davon aus, dass außer den von uns einstellbaren Parametern, $f()$ nicht weiter veränderlich ist, dann lässt sich unsere Suche nach geeigneten Kameraeinstellung beschränken auf die Suche nach den optimalen Einstellungen für Belichtungszeit, Weißabgleich und Schwarzstufe.

3. Analyse des bestehenden Farbaufnahmeprozesses

Beim bestehenden Aufnahmeprozess werden, im Anschluss an den eigentlichen Scan, eine bestimmte Menge an Bildern aufgenommen, die anschließend auf die Daten des aufgezeichneten Scans angewendet werden sollen. Hierbei soll möglichst der gesamte gescannte Bereich komplett eingefärbt werden. Die Linse der integrierten Kamera besitzt einen Öffnungswinkel von 61.4° , was eine minimale Anzahl von 6 Bildern horizontal an jeweils 6 vertikalen Positionen, insgesamt also 36 Bilder, nötig macht um die gesamte Scankuppel komplett abzudecken. Jedoch sind in der Realität ca. 40% jedes Bildes nicht für die Einfärbung zu gebrauchen. Dies liegt zum Teil daran, dass Teilbereiche der Bilder, die nahe an den Rändern des Drehspiegels liegen, oft prismatisch bedingte Farbverläufe und Spieglungsartefakte, so wie fokale Unschärfe aufzeigen. Der Hauptgrund ist jedoch, dass große Teile des Aufnahmebereichs durch das Gehäuse des Geräts selbst verdeckt werden. Aus diesem Grund muss ein ausreichend große Überschneidung zwischen den einzelnen Aufnahmen gewährleistet werden, um den gesamten Scanbereich zuverlässig mit Farbinformationen abzudecken. In vorangegangenen Untersuchungen wurde deshalb die Farbaufnahme so angepasst, dass insgesamt 85 Einzelbilder für die komplette Scankuppel aufgenommen werden. Die Aufnahmen werden an 14 Positionen entlang des Horizonts mit jeweils 6 vertikalen Positionen vorgenommen, zuzüglich eines einzelnen Bildes senkrecht nach oben. Aus diesen 85 Einzelbildern lassen sich ausreichend viele zuverlässige Farbinformationen gewinnen, um alle gemessenen Scanpunkte damit einzufärben.

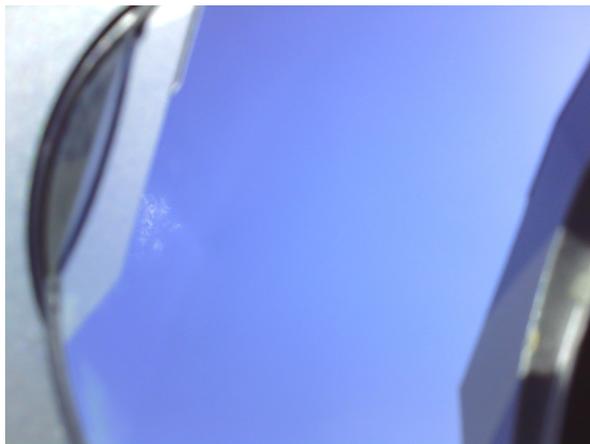


Abbildung 3.1.: Ein exemplarisches Einzelbild der 85 Aufnahmen für die Einfärbung

3. Analyse des bestehenden Farbaufnahmeprozesses

Die Anwendung dieser 85 Einzelbilder auf die Scandaten geschieht innerhalb der FARO SCENE Software bei einem entsprechenden Nachbearbeitungsprozess. Hierbei werden die Bilder jedoch nicht tatsächlich zu einem Panorama zusammengesetzt, sondern für jeden einzelnen Scanpunkt wird ein einzelner oder mehrerer entsprechende Pixel aus verschiedenen Bildern bestimmt, deren Farbwerte dem Scanpunkt direkt zugewiesen werden. Soll ein Scanpunkt Farbinformationen aus mehreren Pixeln erhalten, weil sich mehrere Bilder an der entsprechenden Stelle überschneiden, dann werden diese vorher gewichtet und unter Berücksichtigung dieser Gewichte gemittelt. Ein Punkt im Bild, der nahe am Rand des benutzbaren Bereichs liegt, erhält ein niedriges Gewicht, während ein Punkt, der zentral in der Mitte des benutzbaren Bereiches liegt, ein hohes Gewicht erhält. Unter dem benutzbaren Bereich eines Bildes ist hier der Bereich zu verstehen, der nicht das Gehäuse des Geräts abbildet und der einen bestimmten Mindestabstand zum Rand des Drehspiegels besitzt.

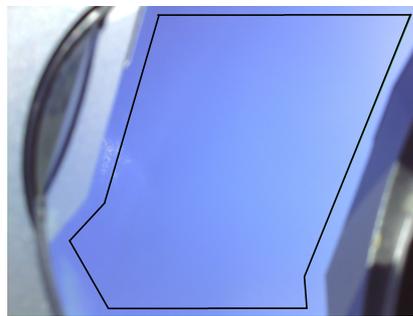


Abbildung 3.2.: Einzelaufnahme mit eingezeichnetem, benutzbarem Bereich

Wie in Kapitel 2 beschrieben werden alle 85 Farbaufnahmen mit derselben Belichtungszeit aufgenommen um Farbverläufe zwischen angrenzenden Bildern zu verhindern. Um dies zu erreichen, muss eine Einstellung für die Belichtungszeit gefunden werden, die in der Gesamtheit der 85 Einzelbilder dafür sorgt, dass möglichst wenige über- oder unterbelichtete Bereiche auf den Bildern vorkommen. Ebenso muss dafür gesorgt werden, dass die Einstellungen für den Weißabgleich und die Schwarzstufe in allen Aufnahmen identisch sind. Zur Ermittlung solcher Einstellungen werden entsprechende Automatismen des Kamerachips selbst verwendet, die in der Lage sind solche Einstellungen für eine einzelne Aufnahme nahezu optimal zu wählen. Da die Einstellung für den Weißabgleich in erster Linie von der „Farbtemperatur“ der momentanen Lichtquellen für die Szene abhängig ist, wird davon ausgegangen, dass es genügt, den Weißabgleich für ein einzelnes Bild im Voraus zu bestimmen und diese Einstellungen dann für die restlichen Aufnahmen gleichermaßen zu verwenden. Die Schwarzstufe unterliegt hierbei ähnlichen Annahmen und wird, zusammen mit dem Weißabgleich, nur einmal zu Beginn der Farbaufnahmen ermittelt und dann für alle nachfolgenden Aufnahmen weiter verwendet. Da die optimale Belichtungszeit für bestimmte Einzelaufnahmen jedoch stark schwanken kann, genügt es nicht, diese für eine einzelne Aufnahme zu Beginn der Farbaufnahmen zu wählen. Sie muss unter Berücksichtigung der gesamten Helligkeitsumgebung gewählt werden, wozu eine Messung an mehreren Stellen im Raum nötig ist.

Um eine möglichst gute Belichtungszeit für den gesamten Scanbereich zu ermitteln, wird die automatische Belichtungsbestimmung der Kamera in regelmäßigen Abständen, während des Scanvorgangs selbst, angestoßen. Der selektiv durchlässige Spiegel vor der Kamera blockt nahezu 100% des reflektierten Lasers, weshalb die Kamera kaum durch den laufenden Scanvorgang beeinflusst wird. Die Belichtungsmessung während des laufenden Scans bietet den Vorteil, dass durch die konstante Drehung des Spiegels ein sehr guter „mittlerer“ Eindruck der momentanen Helligkeitssituation gewonnen werden kann. Bild 3.3 veranschaulicht diesen Effekt sehr gut. Diese Aufnahme ist entstanden, indem während des laufenden Scans, also mit drehendem Spiegel, ein Bild aufgenommen wurde. Man erkennt, dass durch die schnelle Drehbewegung des Spiegels im Mittel ein guter Eindruck der umgebenen Helligkeit gewonnen werden kann, bezogen auf die derzeitige, horizontale Ausrichtung des Drehspiegels. Derartige Messungen werden an mehreren horizontalen Positionen wiederholt und bilden somit eine Reihe an Belichtungsmessungen, die zur späteren Ermittlung einer möglichst optimalen Belichtungszeit dienen. Um aus diesen Werten ein Mittel auszuwählen das möglichst robust gegen Ausreißer in der Messreihe ist, wird im bestehenden Prozess für die optimale Belichtungszeit der Median aus den ermittelten Belichtungswerten gewählt.

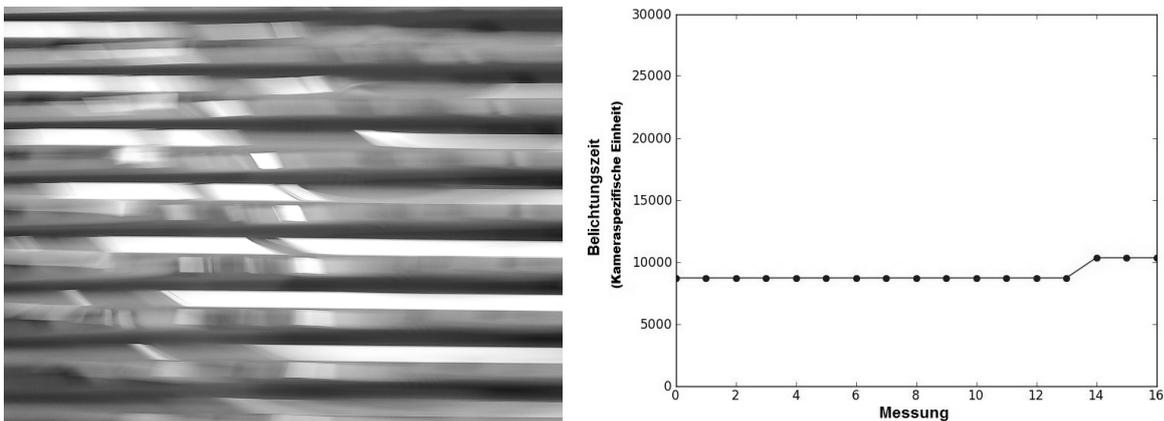


Abbildung 3.3.: Einzelaufnahme die während des Scannens aufgenommen wurde und die Beleuchtungsmessreihe die bei der Szene entstand. Das Bild verdeutlicht den „Schmiereffekt“, der durch die Drehung des Spiegels während des Scannens entsteht.

Der beschriebene Prozess soll nun auf mögliche Mängel untersucht und gegebenenfalls angepasst werden, um entsprechende Mängel zu beheben.

3.1. Festgestellte Problematik

Bei der Untersuchung der Resultate des bestehenden Farbaufnahmeprozesses fiel auf, dass es bestimmte Helligkeitssituationen gibt, bei denen die ermittelte Belichtungszeit zu kurz ist, um ein ausreichend helles Bild zu erhalten. Die resultierende Einfärbung des Scans weist deutlich unterbelichtete Bereiche auf und verfügt über allgemein wenig Kontrast, während

3. Analyse des bestehenden Farbaufnahmeprozesses

die Einfärbung eines anderen Scans, der mit demselben Gerät, nur einige Meter weiter, im selben Raum aufgenommen wurde, ein durchweg gutes Ergebnis liefert.



Abbildung 3.4.: Vergleich zwischen zwei Scans, bei denen Probleme mit der Bestimmung einer idealen Belichtungszeit aufgetreten sind. Beim oberen Scan wurde eine zu kurze Belichtungszeit bestimmt. Beim unteren Scan wurde eine korrekte Belichtungszeit bestimmt.

Bei der Untersuchung der gemessenen Belichtungszeiten von entsprechend ähnlichen Einfärbungen fiel auf, dass die Liste der Messresultate auffällig linear ausfällt. Dies deutet auf eine sehr homogene Helligkeitsumgebung hin, was in den meisten Anwendungsfällen jedoch sehr unwahrscheinlich zu beobachten ist. Die Ursache für dieses Phänomen findet sich, wenn man die Situation der umgebenen Lichtquellen aus Bild 3.4 genauer betrachtet. In Bild 1 befindet sich der Scanner direkt unter einer hellen Deckenlampe. In Bild 2 befindet sich keine aktive Lichtquelle direkt über dem Scanner. Hieraus erklären sich auch sehr einfach die stark linear ausfallenden Messwerte. Durch die Drehung des Spiegels um die Längsachse und die

gleichzeitige Drehung des Scanners selbst um die vertikale Achse, zeigt der Drehspiegel bei jeder Umdrehung einmal, entlang der vertikalen Achse, senkrecht nach oben. Befindet sich in dieser Richtung eine Lichtquelle, so wird diese bei jeder einzelnen Spiegelumdrehung und somit bei jeder einzelnen Belichtungsmessung erneut, mindestens einmal, wahrgenommen.

Dadurch, dass die in Bild 1 zu erkennende Deckenlampe über dem Scanner somit sehr viel öfter als alle andern Lichtquellen den Sensor der Kamera bestrahlt, bestimmt dieser, bei der automatischen Bestimmung der Belichtungszeit, einen entsprechend niedrigeren Wert um auf die anscheinend sehr helle Beleuchtungssituation einzugehen. In der letztendlichen Einfärbung ist dadurch der entsprechende Bereich über dem Scanner gut belichtet, während die Belichtungszeit für den Rest der Abbildungen zu kurz war und diese damit zu dunkel ausfallen.

Die momentane Methode zur Bestimmung der Belichtungszeit weist also sichtbare Mängel auf, speziell dann, wenn sich helle Lichtquellen direkt über dem Scanner befinden. Die Methodik muss also in solcher Form abgeändert werden, dass sie robust gegen derartige Situationen ist und dennoch eine Einzelbelichtung findet, mit der möglichst wenige über- und unterbelichtete Bereiche auf der resultierenden Gesamteinfärbung entstehen.

4. Bestimmung einer optimalen Belichtungszeit

Es wurde gezeigt, dass die aktuell bestehende Bestimmungsmethode für die Belichtungszeit deutliche Mängel aufzeigt, wenn sich der Scanner bei der Bestimmung einer Belichtungszeit direkt unter einer Lichtquelle befindet. Der Prozess muss deshalb derart abgeändert werden, dass solche Probleme vermieden werden. Dazu sollte sichergestellt werden, dass die Belichtungsmessung einzelne Punkte im Raum nicht mehr als einmal betrachtet. Durch die Drehung des Spiegels während der Messung ist dies nur dann möglich, wenn gewährleistet werden kann, dass die einzelnen Messungen der Kamera nur maximal eine volle Umdrehung des Drehspiegels lang andauern. Hierzu soll das Zeitfenster bestimmt werden, innerhalb dessen die Kamera ihre Messungen vornehmen kann. Bei geringster Scanauflösung und damit auch zugleich bei der höchsten Drehfrequenz des Spiegels, erreicht dieser eine Drehfrequenz von 95.4Hz. Hieraus ergibt sich eine Dauer von ca. 0.0104 Sekunden pro voller Umdrehung. Die Linse der Kamera hat einen Öffnungswinkel von 61.4° , also ca. 17% eines Vollkreises. Gehen wir davon aus, dass die Kamera keine Messungen innerhalb des Bereiches senkrecht nach oben durchführen soll, bleiben somit nur noch die verbleibenden 83% des Vollkreises zur Verfügung und somit ein Zeitfenster von 0.0086 Sekunden. Nun soll die Dauer überprüft werden, wie lange die Kamera braucht um eine Belichtungsmessung vorzunehmen. Dazu verwenden wir folgenden Versuchsaufbau:



Abbildung 4.1.: Aufbau um die Dauer der Belichtungsmessung zu bestimmen. Auf den Scanner ist eine Schaumstoff-Haube gesteckt, die verhindert, dass Licht auf die Kamera fällt.

4. Bestimmung einer optimalen Belichtungszeit

Die kontinuierliche Belichtungsmessung der Kamera wird eingeschaltet und der Sensor abgedeckt. Nun wird gewartet, bis die Kamera ihre Messungen abgeschlossen hat und einen stabilen Wert für die Belichtungszeit ermittelt hat. Daraufhin wird die Abdeckung entfernt und der Sensor somit wieder einem Lichtreiz ausgesetzt. Sobald die Kamera erneut einen stabilen Wert für die Belichtungszeit ermittelt hat, messen wir die Zeit, die der interne Mechanismus hierfür benötigt. Diese Messungen werden mehrfach wiederholt, jeweils von Hell nach Dunkel und umgekehrt, und der jeweilige Worst-Case wird notiert.

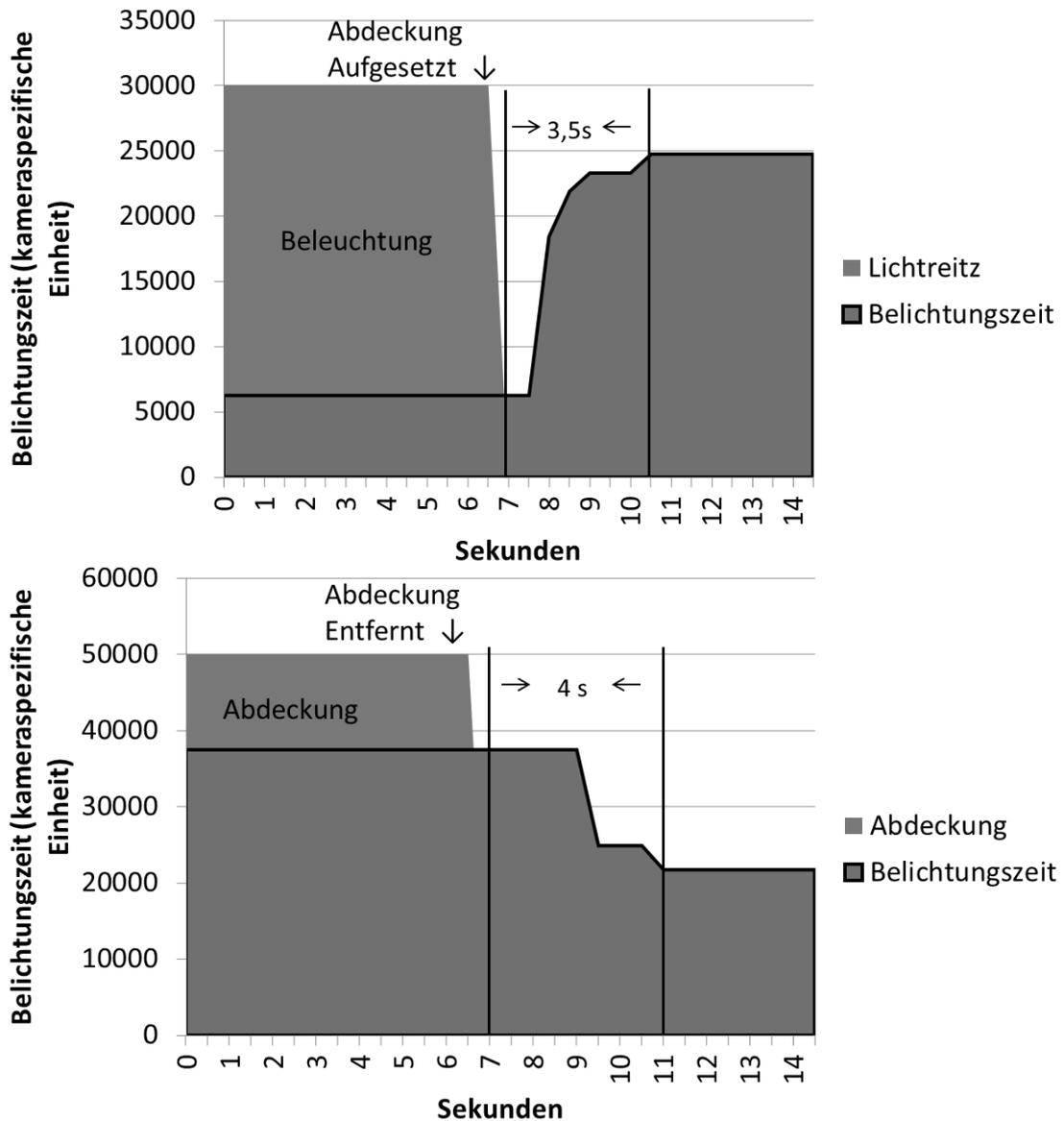


Abbildung 4.2.: Messergebnisse für die Dauer der automatischen Belichtungsmessung. Die Belichtungszeit ist in einer kameranormierten Einheit angegeben.

Aus den Messungen kann man erkennen, dass die Kamera im Worst-Case bis zu 4 Sekunden braucht um eine stabile Messung vorzunehmen.

Die Dauer dieser Messung kann reduziert werden, indem die Frequenz mit der die Kamera Bilder aufnimmt erhöht wird, da die automatische Bestimmung einer Belichtungszeit auf den intern erzeugten Bildern der Kamera basiert. Die Kamera ist in unserem Anwendungsfall auf eine Bildrate von 3.75 Bildern pro Sekunde eingestellt. Dem Datenblatt der Kamera kann entnommen werden, dass die Kamera bei unserer gewünschten Bildauflösung von 2048×1536 Pixeln eine maximale Frequenz von 15 Bildern pro Sekunde erreichen kann. Die Dauer der automatischen Belichtungsmessung könnte demnach um einen Faktor 4 reduziert werden, indem die Kamera auf ihre maximale Aufnahme Frequenz eingestellt werden würde. Demnach würde der Worst-Case für die automatische Belichtungsmessung nur noch 1 Sekunde in Anspruch nehmen.

Dennoch benötigt die automatische Belichtungsmessung der Kamera bedeutend mehr Zeit, also das Zeitfenster von 0.0086 Sekunden, die der Spiegel im schnellsten Fall braucht, bevor die Kamera einen bestimmten Punkt im Raum zum zweiten Mal betrachten würde. Da der Drehspiegel im langsamsten Fall immer noch eine Drehfrequenz von 1.9Hz besitzt und sich somit ein Zeitfenster von 0.436 Sekunden ergeben würde, ist damit gezeigt, dass die Kamera in unserem Anwendungsfall nicht in der Lage ist, schnell genug eine automatische Belichtungsmessung vorzunehmen, die garantiert keinen Punkt im Raum mehr als einmal betrachtet.

Ebenfalls soll betrachtet werden, ob eine manuelle Messung der Belichtung anhand der aufgenommen Kamerabilder möglich wäre. Bei der optimalen Kameraeinstellung von 15 Bildern pro Sekunde würde jedes Bild ca. 0.07 Sekunden für eine Aufnahme in Anspruch nehmen, zuzüglich der Zeit die für die Übertragung aus dem Kamerachip heraus benötigt wird. Für die maximale Drehfrequenz des Spiegels (Zeitfenster: 0.0086 Sekunden) ist dies immer noch bedeutend zu lange.

Es wird also klar, dass eine Messung vorgenommen werden muss, bei der der Drehspiegel entsprechend lange stationär gehalten werden kann. Zu diesem Zweck soll ein zusätzlicher Messdurchlauf nach dem Scannen selbst durchgeführt werden.

4.1. Zusätzlicher Messdurchlauf

Die Auslagerung des separaten Messvorgangs für Beleuchtungssituation ermöglicht es, speziell auf relevante Bereiche innerhalb des Scanbereichs einzugehen. So kann beispielsweise ein bestimmter Ausschnitt des Scanbereichs eigenständig betrachtet werden, um für diesen Bereich eine ideale Belichtungszeit zu bestimmen, während man in Kauf nimmt, dass die verbleibenden Bereiche eventuell ungünstig belichtet sind. Auf dieser Grundlage wurden 2 Basisszenarien gewählt: Betrachtung der Beleuchtungssituation entlang des Horizonts und die Betrachtung der Beleuchtungssituation am Boden, bzw. an der Decke. Da in den meisten Anwendungsfällen die relevanten Strukturen und Informationen entlang des Horizonts zu

4. Bestimmung einer optimalen Belichtungszeit

finden sind, wurde dieser Fall als Standardszenario gewählt. Für die Umsetzung dieses Verfahrens wurden mehrere Versuche angestellt, die die Zuverlässigkeit der Zusatzmessungen in Abhängigkeit zu deren Anzahl untersuchen sollen.

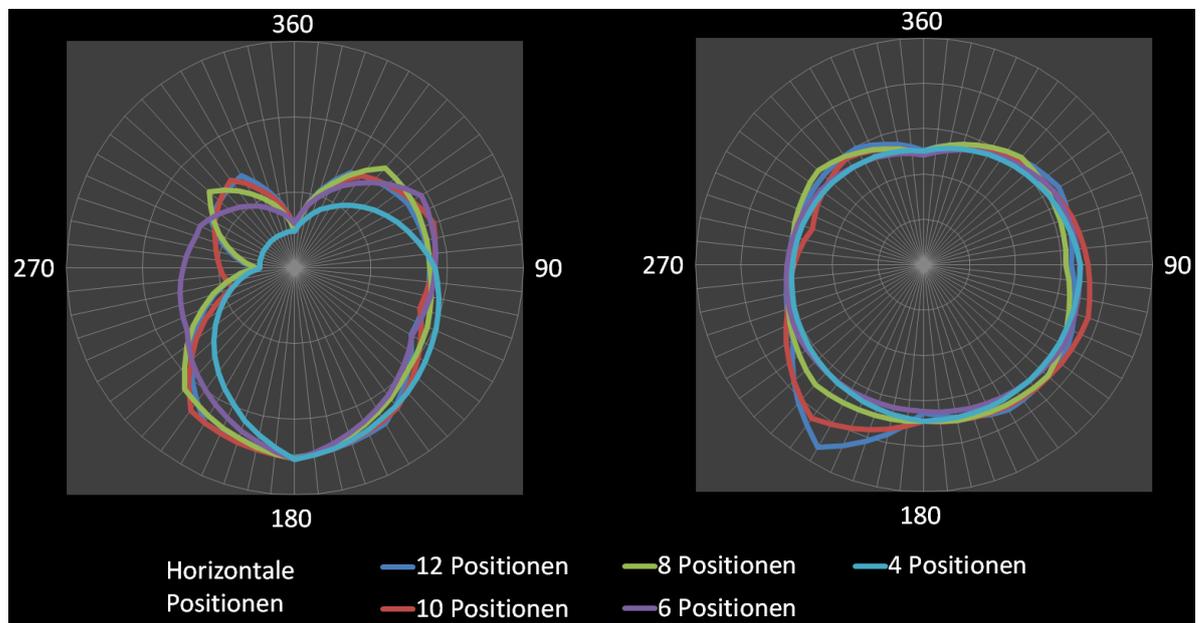


Abbildung 4.3.: Messergebnisse für die verschiedenen Anzahlen an Messpositionen. Man erkennt, dass der Messverlauf für 4 und 6 Positionen sehr ungenau ist. Der Verlauf für 8 und mehr Positionen wird zunehmend genauer. Im linken Bild befand sich der Scanner im Halbschatten, mit einem einzelnen Fenster, durch das direkt die Sonne auf den Scanner fiel. Im rechten Bild befand sich der Scanner in einem gleichmäßig beleuchteten Raum mit einer einzelnen Schreibtischlampe, die auf den Scanner gerichtet war.

Die Verteilung der einzelnen Messpunkte entlang des Horizonts wurde dabei unter Berücksichtigung des Blickwinkels der Kamera vorgenommen. Die Mindestanzahl der Messpunkte wurde deshalb auf 6 äquidistant verteilte Positionen entlang des Horizonts festgelegt, da dies die geringste Anzahl an Positionen ist, die unter Berücksichtigung des Blickwinkels der Kamera (ca. 60 Grad) ausreicht um einmal den 360 Grad Vollkreis mit Messungen abzudecken. Auf dieser Basis werden weitere Positionen hinzugefügt, um mehr Überdeckung zwischen den einzelnen Messungen zu erreichen. Da Messpositionen entlang der Symmetrieachse des Scanners auf die andere Hälfte des Blickfelds übertragen werden, können immer nur gleichzeitig 2 symmetrische Messpositionen hinzugefügt werden. Insgesamt werden fünf Modi untersucht, bei denen jeweils 4, 6, 8, 10 und 12 Messungen entlang des Vollkreises vorgenommen werden, also 2, 3, 4, 5 und 6 Messungen pro Scanhälfte.

Die Resultate der Tests zeigen, dass die Abtastungen an 4 und 6 Positionen im Vollkreis nicht ausreichend genau sind um einen guten Eindruck der umgebenen Belichtungssituation zu gewinnen. Der Grund hierfür liegt an der Abdeckung des Blickfelds der Kamera durch

das Gehäuse des Scanners, was dazu führt, dass bei 4 bzw. 6 Messpunkten bestimmte Teilbereiche des Raums niemals von der Kamera betrachtet werden können. Erst ab 7 Positionen wäre die Überdeckung zwischen den Einzelbildern wieder groß genug, damit Abdeckungen der Aufnahmen durch das Gehäuse nur noch wenig ins Gewicht fallen.

Um die Zuverlässigkeit der Messungen zu erhöhen wurden die Tests noch ein weiteres Mal ausgeführt mit jeweils zwei vertikalen Messpunkten pro horizontaler Position, entgegen den vorherigen Messungen, die nur an einer vertikalen Position durchgeführt wurden. Dadurch wird der vertikale Blickwinkel der Kamera effektiv verdoppelt und somit ein breiterer Bereich entlang des Horizonts betrachtet. Hierbei zeigte sich, dass die Messung mit zwei vertikalen Positionen keine nennenswerte Verbesserung der Messresultate erzielt.

4.2. Ermittlung der optimalen Einzelbelichtung anhand der Messdaten

Im Anschluss an die Belichtungsmessungen der Kamera muss nun noch eine geeignete Einzelbelichtung aus den Messreihen gewonnen werden, die dann für alle weiteren Aufnahmen verwendet werden soll. Bei der Betrachtung des aktuellen Verfahrens (Wahl des Medians) kann man erkennen, dass es bestimmte Situationen gibt, in denen der Median nicht die optimale Wahl für einen Wert aus der Messreihe darstellt. Betrachtet man zum Beispiel die Messreihe 4.4, sieht man einfach, dass der hierbei gewählte Median immer nur für eine Hälfte der Messdaten optimal und für die andere Hälfte völlig unbrauchbar sein kann. Aus diesem Grund soll in der neuen Methode zur Bestimmung der optimalen Belichtungszeit eine möglichst ausreißersichere Mittelung zum Einsatz kommen. Dazu werden jeweils die obersten und untersten 5% der sortierten Messwerte verworfen und über die verbleibenden Messwerte wird ein Mittel gebildet. Auf diese Weise können Worstcase Szenarios, wie beim gezeigten Beispiel mit einer bipartiten Messreihe, effektiv vermieden werden.

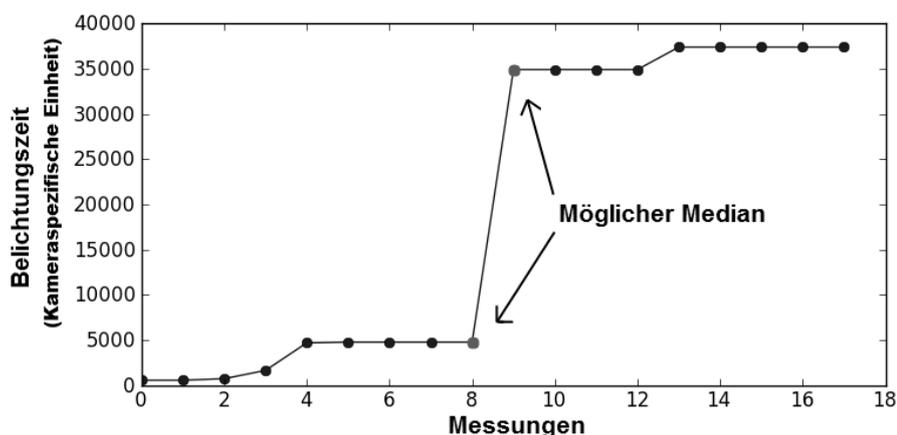


Abbildung 4.4.: Bipartite Messreihe mit ungünstigem Median.

4.3. Verbleibende Problematik

Speziell in sehr stark unterschiedlichen Beleuchtungssituationen versagt das Verfahren weiterhin, da keine Belichtungszeit gefunden werden kann, die dafür sorgt, dass alle Bereiche innerhalb des Scans ausreichend gut belichtet sind. Betrachtet man Bild 4.4, so erkennt man schnell, dass es keine einzelne Belichtungszeit gibt, die für alle Messwerte gleichermaßen, ausreichend nah bei dem gemessenen Optimalwert liegt. Egal welche Belichtungseinstellung letztendlich gewählt wird, werden einige Bildbereiche über- bzw. unterbelichtet sein.

Ein solches Messergebnis deutet darauf hin, dass innerhalb des vermessenen Bereichs sowohl sehr helle, als auch sehr dunkle Bereiche vertreten sind. Diesen Unterschied zwischen Hell und Dunkel bezeichnet man als den „Dynamikumfang“ einer Szene. Auch Kameras besitzen einen Dynamikumfang, der ein Maß für das Verhältnis zwischen dem hellsten und dem dunkelsten Farbreiz ist, den eine Kamera noch korrekt abbilden kann. Korrekt abbilden bedeutet in diesem Fall, dass der eingehende Farbreiz stark genug ist um überhaupt vom Sensor der Kamera wahrgenommen zu werden und gleichzeitig nicht zu stark ist, um den Sensor zu sättigen (auf pures Weiß abgebildet werden).

Dynamikumfang wird typischer Weise in Decibel (dB) angegeben.

$$(4.1) \quad L_{dB} = 10 * \log_{10} \left(\frac{P_1}{P_0} \right)$$

Die Menge an Decibel entspricht demnach dem zehnfachen vom 10er-Logarithmus des Verhältnis der Messgrößen P_1 und P_0 . Im menschlichen Sinne: Für jeden vollen Faktor von 10, der zwischen den beiden Messgrößen liegt, steigt ihr Verhältnis um genau 10 Decibel.

Während ein menschlicher Betrachter einen Dynamikbereich von bis zu 90dB erkennen kann (ca. 1.000.000.000 zu 1), besitzt der eingesetzte Kamerasensor nur einen Dynamikumfang von 60dB, was einen Faktor von 1.000-facher Helligkeit, bzw. Dunkelheit bedeutet, die ein Mensch im Gegensatz zum Kamerasensor immer noch wahrnehmen kann. Aus diesem Grund ist es schwer, mit einer Kamera eine Szene exakt so abzubilden, wie ein menschlicher Betrachter diese wahrnimmt, ohne dass dabei in der Abbildung über- oder unterbelichtete Bereiche auftreten.

Ein Weg um den begrenzten Dynamikumfang unserer Kamera zu erhöhen und damit über- bzw. unterbelichtete Bereiche auf stark unterschiedlich beleuchteten Szenen zu vermeiden, ist der Einsatz von High Dynamic Range Imaging Methoden.

5. High Dynamic Range Imaging

Als *High Dynamic Range Imaging* (**HDRi**) werden eine Menge an Verfahren und Methoden bezeichnet, mit denen sich große Helligkeitsunterschiede auf Bildern genauer beschreiben lassen, als es mit normalen Mitteln möglich wäre. Das menschliche Auge ist noch recht problemlos in der Lage, die Struktur und Einrichtung eines Raumes zu erkennen, der von nichts weiter als einer kleinen Kerze erhellt wird. Der Helligkeitsunterschied zwischen dunklen Bereichen und der Kerzenflamme ist allerdings enorm groß. Für viele Kameras und deren Sensoren ist es daher bereits eine Herausforderung, mehr als den direkt von der Kerze beleuchteten Bereich aufzunehmen.

Treten in Bildern über- bzw. unterbelichtete Bereiche auf, dann ist dies ein Indiz dafür, dass der entsprechende Sensor über nicht genügend Dynamikumfang verfügt um die Szene korrekt aufzuzeichnen. Der *Dynamikumfang* eines Sensors ist das Verhältnis zwischen dem kleinsten Reiz, der über den unteren Schwellwert hinausgeht und dem größten Reiz, der noch unterhalb der Sättigung liegt. Dieser Dynamikumfang ist technisch bedingt und lässt sich –in den allermeisten Fällen– nicht vom Nutzer einstellen. Bei den meisten Kameras ist es jedoch möglich, die Belichtungszeit zu verändern und somit den Bereich des Dynamikumfangs „auszuwählen“, der korrekt abgebildet werden soll.

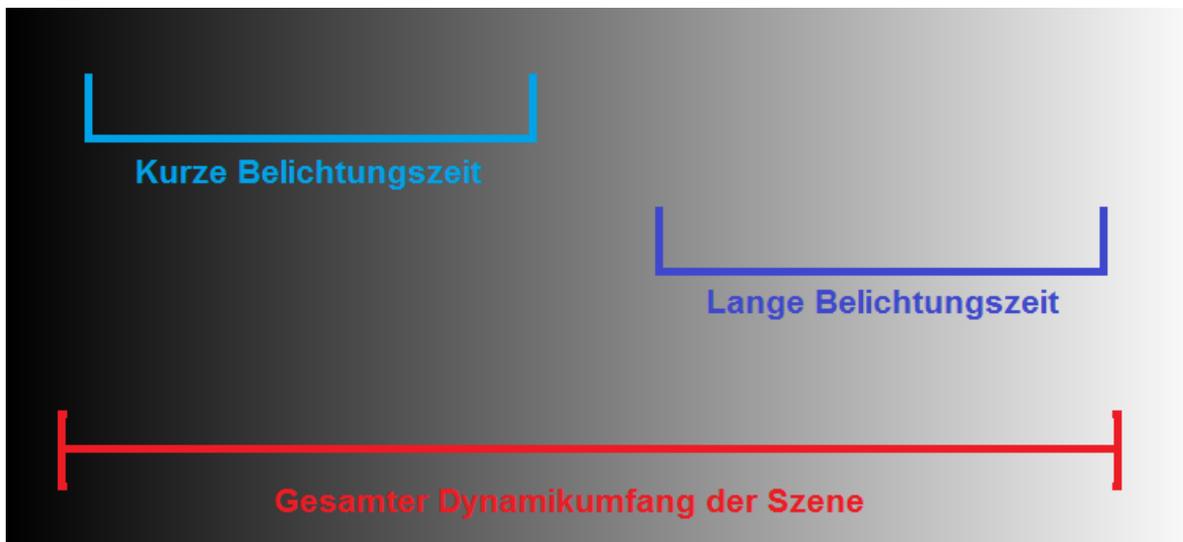


Abbildung 5.1.: Abhängig von der eingestellten Belichtungszeit der Kamera wird ein entsprechender Bereich des gesamten Dynamikumfangs aufgezeichnet.

5. High Dynamic Range Imaging

Es gibt bereits einige digitale Kameras auf dem Markt zu erwerben, die Dynamikumfang von 110 dB und mehr erreichen können¹, dennoch ist derartige Technologie aber immer noch nicht im breiten Einsatz zu finden. Das größte Problem von High Dynamic Range Imaging findet sich nämlich nicht in der Sensorik und der Aufnahmemethodik wieder, sondern in der Darstellung solcher HDR Bilder, die man auch als Radiance Maps bezeichnet, auf konventionellen Displays.

Wie bereits erwähnt, ist das menschliche Auge in der Lage, einen recht großen Dynamikumfang von bis zu 90 dB zu erfassen, was etwa dem Helligkeitsunterschied gleich kommt zwischen einem Objekt, das von schwachem Mondlicht beschienen wird, und einem Objekt das sich unter direkter Sonneneinstrahlung befindet. Allerdings ist das menschliche Auge nicht in der Lage, diesen Kontrast zur selben Zeit zu erfassen, sondern muss sich auf die jeweilige Situation erst einstellen (Adaption).

Dagegen besitzen handelsübliche LCD Displays meist einen dynamischen Kontrast von 2.000.000 zu 1, also lediglich 63 dB. Hierbei handelt es sich allerdings um dynamischen Kontrast, der geringer ausfällt, wenn das volle Spektrum der Helligkeit von 0% bis 100% im Eingangssignal vertreten ist. Um diese dynamischen Kontrastverhältnisse zu erreichen wird meist die Hintergrundbeleuchtung des Displays mit verwendet, was effektiv eine Verschiebung der gesamten Helligkeitsskala des Displays bewirkt, ähnlich wie die Einstellung der Belichtungszeit bei einer Kamera. Der absolute Kontrast, den solche Displays typischerweise erreichen, liegt bei ca. 1.200 zu 1, was gerade einmal 31dB entspricht. Man erkennt also, dass selbst wenn wir in der Lage wären Bilder mit beliebigem Dynamikumfang aufzunehmen, wir keinen direkten Nutzen daraus ziehen könnten, da sich solche HDR Bilder nicht ohne weiteres auf konventionellen Displays und somit auch nicht auf den meisten Computerbildschirmen darstellen ließen.

Displays codieren üblicherweise ihre Farben im 8-Bit RGB System. Dieses System beschreibt die Farbeindrücke Rot, Grün und Blau getrennt voneinander jeweils als 8-Bit Zahl. Dies entspricht einer Farbtiefe von 0 bis 255. Somit kann jeder Farbkanal auch nur eben jene 256 Helligkeitsabstufungen besitzen, was gerade einmal einem Dynamikbereich von 24 dB entspricht. Zusammen mit dem geschickten Einsatz der Hintergrundbeleuchtung des Displays lassen sich diese 256 Helligkeitsstufen zusätzlich noch linear verschieben und somit ein insgesamt größerer Umfang an darstellbaren Gesamthelligkeiten erreichen. Dennoch ist der Unterschied zwischen der dunkelsten, darstellbaren Helligkeitsstufe und der hellsten, darstellbaren Helligkeitsstufe eines Displays und damit auch sein Dynamikumfang, bedeutend geringer als die Helligkeitsunterschiede, die in natürlichen Szenen vorkommen können.

Radiance Maps, also HDR Bilder, beschreiben oft ihre jeweiligen Farbkanäle als 32-Bit Zahl, also einer Farbtiefe von 0 bis 4.294.967.295, was einem theoretischen Dynamikumfang von ca. 96 dB entspricht; fast 17 Millionen mal mehr, als ein konventionelles Display üblicherweise darstellen kann.

Um ein HDR Bild also auf einem normalen Display anzeigen zu können, muss sein Dynami-

¹Zum Beispiel der Omnivision OV10121 Chip

kumfang wieder weit genug reduziert werden, um durch die verwendete 8-Bit Codierung darstellbar zu sein. Diesen Vorgang nennt man *Dynamikkompression* oder auch *Tone Mapping*. Diese Kompression wird meistens so gestaltet, dass der subjektive Eindruck des komprimierten **LDR** (Low Dynamic Range) Bildes (im Rahmen der technischen Machbarkeit) möglichst nahe an dem der abgebildeten Szene liegt.

5.1. Technische Grundlagen

Das Verfahren des High Dynamic Range Imaging für die Darstellung von Bildern auf konventionellen Low Dynamic Range Displays lässt sich grob in zwei Schritte unterteilen:

- Die HDR Aufnahme selbst.
- Anschließendes Tone Mapping.

Die Grundgedanken der beiden Teilschritte sollen hier genauer betrachtet werden.

5.1.1. HDR Aufnahme

High Dynamic Range in der Fotografie bedeutet, dass der Helligkeitsunterschied zwischen dem dunkelsten und dem hellsten Punkt der Szene sehr groß ist. Sollen solche Szenen korrekt abgebildet werden, ergeben sich schnell Schwierigkeiten, wenn man dazu einen Sensor verwendet, der nicht in der Lage ist, einen entsprechend großen Dynamikbereich zu erfassen. Um diese Probleme zu überwinden kann man sich einiger mathematische Methoden bedienen, die es ermöglichen, auch mit einem Sensor der nicht über einen genügend großen Dynamikumfang verfügt, die Szene korrekt aufzuzeichnen. Hierbei wird die Aufnahme einer Szene mehrfach wiederholt mit variierenden Belichtungszeiten. Dadurch werden, abhängig von der eingestellten Belichtungszeit, jeweils andere Teile des Dynamikbereichs der Szene auf dem Bild korrekt abgebildet. Auf diese Weise kann ebenfalls der volle Dynamikbereich der Szene abgebildet werden, jedoch auf mehrere Einzelbilder verteilt, anstelle nur eines einzigen Bildes. Diese Aufnahmen müssen dann letztendlich wieder entsprechend miteinander verrechnet werden. Hierbei werden jeweils die optimal belichteten Bereiche der einzelnen Bilder verwendet und letztendlich zu einem Einzelbild zusammengesetzt.

Ein solches Verfahren wurde 1997 von Paul Debevec und Jitendra Malik entwickelt ([DM97]) und ist eines der gängigsten Verfahren zur Erzeugung von HDR Bildern aus Belichtungsreihen, die mit einer gewöhnlichen LDR Kamera aufgenommen wurden. Alternativ zu dem von Debevec und Malik entwickelten Verfahren wurde 2003 von Mark Robertson et. al. ([RBS99]) ein Verfahren entwickelt, das auf ähnlichen theoretischen Grundansätzen basiert, jedoch einen anderen mathematischen Ansatz zur Lösung der resultierenden Gleichungssysteme verwendet. Um die Grundgedanken der Erzeugung von HDR Bildern aus einer LDR Belichtungsreihe zu erläutern wird hier das Verfahren von Debevec erklärt, da dieses weniger kompliziert ist als das Verfahren von Robertson und somit die Grundideen besser transportieren kann.

5. High Dynamic Range Imaging

Um mehrere Einzelbelichtungen so in einem HDR Bild zusammenzufassen, dass die jeweiligen Farb- und Helligkeitsinformationen der einzelnen Bildpunkte unabhängig von den jeweiligen Belichtungszeiten beschrieben werden können, müssen diese auf eine unabhängige Größe zurückgeführt werden. Dazu sollen die einzelnen Bildpunkte auf ein proportional korrektes Maß ihrer physikalischen Leuchtkraft zurückgeführt werden, um von den Nichtlinearitäten des Abbildungsprozesses unabhängig zu sein. Da der echte (physikalische) Leuchtwert eines Punktes in der Szene normalerweise nicht bekannt ist, muss er entsprechend approximiert werden. Dazu betrachten wir erneut die formale Beschreibung des Abbildungsprozesses eines Punktes i im Raum:

$$(5.1) \quad V_i = f(E_i * \Delta T)$$

- E_i : Echte Leuchtkraft des Punktes in der Szene (unbekannt)
- ΔT : Belichtungszeit des Sensors (in Sekunden)
- V_i : Abgebildeter Wert des Punktes (Pixelwert oder auch einfach Pixel)
- $f(x)$: Abbildungsfunktion des Sensors (i.d.R. unbekannt)

Anmerkung: Wie in Kapitel 2 beschrieben, wurde die Blendenöffnung in dieser Formulierung vernachlässigt, da diese in unserem Modell unveränderlich ist und somit nur in Form einer linearen Skalierung auf die Abbildung Einfluss hat, die die proportionalen Verhältnisse der einzelnen Lichtwerte zueinander nicht verändert.

Diese Repräsentation soll so umformuliert werden, dass sie sich nach den Werten E_i auflösen lässt. Durch mathematische Umformung erhalten wir dann:

$$(5.2) \quad g(V_i) := \ln f^{-1}(V_i) = \ln E_i + \ln \Delta T$$

$g(x)$ ist hierbei die sogenannte „Camera Response Curve“. Da die Belichtungszeit bekannt ist, müssen wir, um die unbekanntenen E_i Werte zu finden, nun zuerst $g(x)$ bestimmen. Da sowohl die Funktion $g(x)$ als auch die E_i Werte unbekannt sind, können wir $g(x)$ mit unseren Ausgangsdaten nicht vollständig analytisch bestimmen. Wir formulieren deshalb folgende Forderung an unser System um $g(x)$ und E_i in gemeinsamem Zusammenhang zu beschreiben:

$$(5.3) \quad 0 = \sum_{i=0}^N [g(V_i) - E_i - \Delta T]$$

Dies soll für N aus den Bildern ausgewählte Bildpunkte geschehen, die die gesamte Dynamik der Szene möglichst vollständig beschreiben (d.h.: Alle vorkommenden Helligkeiten, über die Einzelbelichtungen hinweg, abdecken). Da $g(x)$ nur diskrete Werte annimmt (die Pixelwerte des Bildes: $0 \dots 255$ (8-Bit Farbtiefe)), lässt sich das Problem reduzieren auf eine Parametersuche nach mehreren Unbekannten. Gesucht werden hierbei die 256 diskreten Werte der Funktion $g(x)$ und die E_i Werte der N ausgewählten Bildpunkte.

Dieses Problem kann sich als ein Optimierungsproblem in Least-Squares Gestalt formulieren lassen:

$$(5.4) \quad \sum_{i=0}^N \sum_{j=0}^P [g(V_{ij}) - E_i - \Delta T_j]^2 \stackrel{!}{=} \text{Min.}$$

Hierbei werden über eine Menge an Bildpunkten N und alle Einzelbelichtungen P summiert. Dabei muss nicht notwendigerweise über alle vorkommenden Bildpunkte N_p summiert werden, sondern es genügt eine Menge an Bildpunkten, so dass das formulierte Least Squares Problem ausreichend überbestimmt ist. Da $g(x)$ nur 256 diskrete Werte annimmt, lässt sich das Problem als eine Parametersuche nach $M \times N$ Unbekannten auffassen, mit $M = 256$ und N , der Anzahl an Bildpunkten, die für die Bestimmung der Kamerakurve herangezogen werden. Dieses Problem lässt sich mittels Singulärwertzerlegung effektiv lösen. Hieraus ergeben sich dann die 256 diskreten Werte der Funktion $g(x)$ so wie die E_i Werte der N ausgewählten Bildpunkte.

Um das Resultat weiter zu verbessern kann zusätzlich noch ein Glattheitstherm angehängt werden, der $g(x)$ durch einen erhöhten Fehler bestraft, je weniger glatt die Funktion ist. Dazu wird das Integral der zweiten Ableitung der Funktion $g(x)$ verwendet. Zudem können wir davon ausgehen, dass Pixelwerte, die nahe am Rand des Wertebereichs liegen, eine hohe Wahrscheinlichkeit haben, verrauscht zu sein. So kann ein Bildpunkt, der den Sensor bereits vollständig sättigt, durch elektronisches Rauschen in der Analog-Digital-Wandlung des Sensors auf einen Messwert abgebildet werden, der kurz unterhalb der Sättigung liegt, und somit das Resultat verfälscht. Aus diesem Grund sollen Punkte, die nahe am Rand des Wertebereichs liegen, weniger Einfluss auf den Gesamtfehler haben als Punkte, die nahe der Mitte des Wertebereichs liegen. Hierzu eignet sich eine entsprechende Gewichtung, beispielsweise eine simple Dreiecksfunktion oder eine Gauss-Glocke.

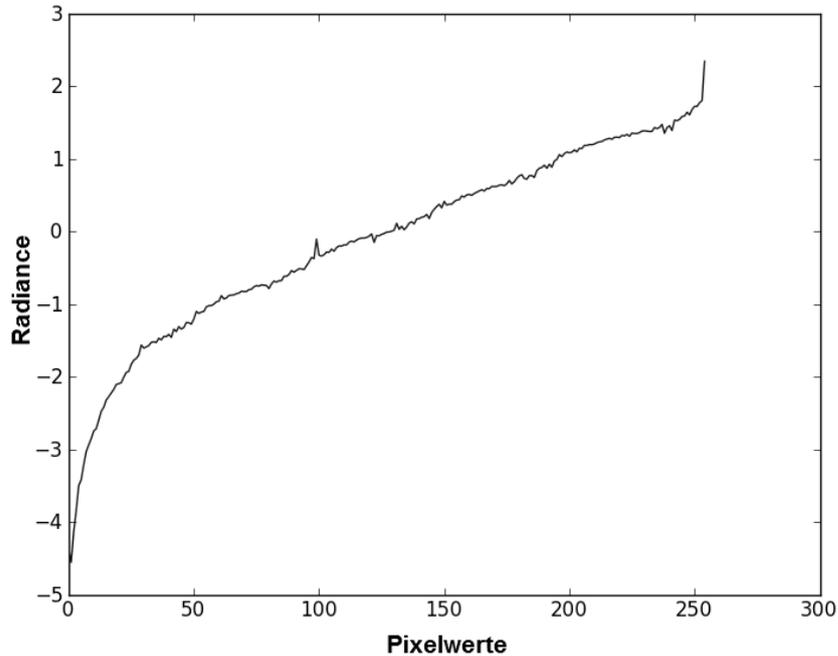


Abbildung 5.2.: Beispiel einer ungeglätteten Response Curve

Wir gewichten unsere Funktion also noch zusätzlich mit der Gewichtungsfunktion $w(x)$ und erhalten somit, zusammen mit dem Glattheitstherm:

$$(5.5) \quad \theta := \sum_{i=0}^N \sum_{j=0}^P [w(V_{ij}) * (g(V_{ij}) - E_i - \Delta T_j)]^2 + \lambda \sum_{v=V_{min}}^{V_{max}} [w(v) * g''(v)]^2$$

$$\theta \stackrel{!}{=} Min.$$

Der Parameter λ steuert hierbei den Einfluss der Glattheitsforderung am Gesamtfehler der Optimierung.

Nachdem wir nun die Response Curve $g(x)$ bestimmt haben, können wir jetzt die verbleibenden E_i aller anderen Bildpunkte bestimmen, die nicht bereits als Teil der Least Squares Optimierung im vorangegangenen Schritt ermittelt wurden.

$$(5.6) \quad \ln E_i = \frac{\sum_{j=0}^P [g(V_{ij}) - \ln \Delta T_j]}{P + 1}$$

Zur Robustheit mitteln wir dazu über alle Einzelbelichtungen hinweg. Hierbei ergibt sich noch eine weitere Optimierungsmöglichkeit: Da Pixel nahe der Sättigung weniger verlässliche

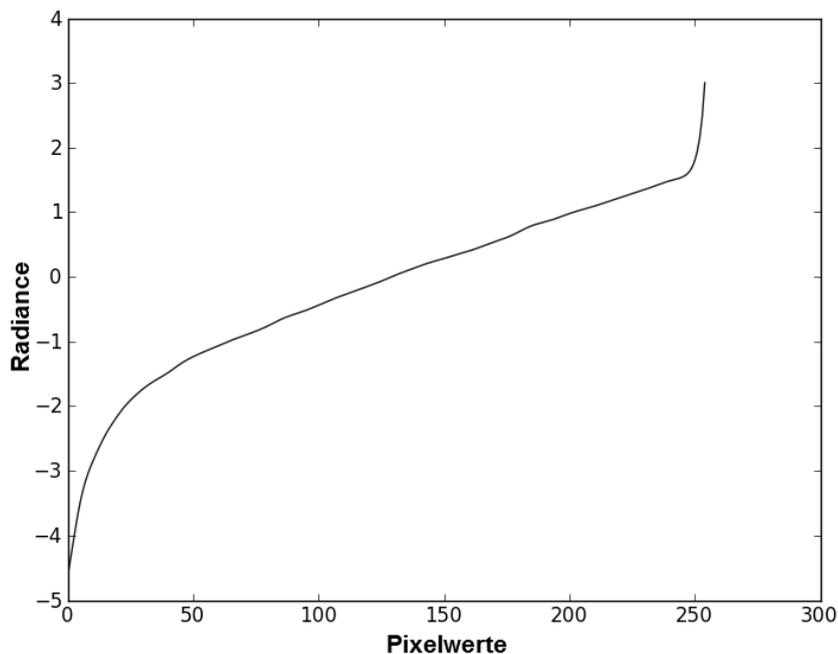


Abbildung 5.3.: Beispiel einer geglätteten Response Curve

Informationen transportieren als Pixel in der Mitte des Wertebereichs, sollten wir solche Pixel entsprechend weniger Einfluss auf das Gesamtergebn haben lassen. Wir verwenden dazu erneut die Gewichtungsfunktion $w(x)$ aus der Berechnung der Response-Curve. Hieraus ergibt sich letztendlich folgende Formel für die Berechnung der (logarithmischen) Leuchtkraft der einzelnen Punkte der Szene:

$$(5.7) \quad \ln E_i = \frac{\sum_{j=0}^P [w(V_{ij}) * (g(V_{ij}) - \ln \Delta T)]}{\sum_{j=0}^P w(V_{ij})}$$

5.1.2. Tone Mapping

Ein HDR Bild umfasst in der Regel einen erheblich größeren Farbumfang als ein konventionelles Display darstellen kann (8-Bit Farbtiefe gegenüber 32-Bit). Deshalb muss ein HDR Bild entsprechend verarbeitet werden, bevor es auf einem solchen Display dargestellt werden kann.

Tone Mapping beschreibt hierbei im Allgemeinen den Vorgang der Abbildung des Farbumfangs eines Bildes auf den Farbumfang eines anderen Bildes. Im Speziellen geht es darum, den großen Farbumfang eines HDR Bildes auf den verhältnismäßig geringen Farbumfang eines LDR Bildes zu reduzieren, so dass lokale Kontrastverhältnisse und globaler Farbeindruck im LDR Bild möglichst weit erhalten bleiben. Das resultierende LDR Bild hat dann einen

5. High Dynamic Range Imaging

ausreichend kleinen Farbumfang, um auf einem normalen (nicht HDR) Display dargestellt zu werden.

Besitzt ein Bild, beispielsweise, einen Farbumfang von 20 diskret darstellbaren Helligkeitsstufen und soll auf einem Display angezeigt werden, das nur 10 Helligkeitsstufen abbilden kann, so muss entschieden werden in welcher Form die 20 Helligkeitsstufen des ersten Bildes „komprimiert“ werden müssen um sie auf 10 diskrete Helligkeitsstufen zu reduzieren.

Hierfür sind eine Reihe von Abbildungen denkbar. Die einfachste Abbildung wäre eine lineare Abbildung in solch einer Gestalt, dass sie jeweils 2 Wertigkeiten des Ausgangsbildes zu einer Wertigkeit zusammenfasst. Kompliziertere, nicht lineare Abbildungen könnten zusätzlich noch Informationen über die Verteilung der Helligkeitsstufen im Ausgangsbild berücksichtigen. Kommen die Helligkeitsstufen 0 bis 4 im Ausgangsbild beispielsweise gar nicht oder nur sehr selten vor, so können diese, ohne nennenswerten Detailverlust, zu einer Helligkeitsstufe zusammengefasst werden und ermöglichen dadurch die verbleibenden Werte weiter aufzufächern.

Die Grafik 5.4 veranschaulicht dieses Problem. Hier wurden die Grauwerte aus dem linken Teil des Bildes (256 diskrete Werte) auf nur noch ca. ein Drittel der Werte im mittleren und rechten Teil des Bildes reduziert. Im mittleren Teil wurde dazu eine lineare Abbildung verwendet. Man erkennt, dass der subjektive Gesamteindruck des Bildes annähernd erhalten bleibt. Einige Details in den Haaren bzw. im Auge gehen allerdings verloren. Im rechten Teil der Abbildung wurden die Werte aus dem linken Teil nicht mehr linear abgebildet, sondern stattdessen folgende Abbildungsfunktion verwendet.

$$(5.8) \quad V_n = \frac{(P_n * 85)}{(P_n + 85)}$$

Es ist zu erkennen, dass viele Flächen des Bildes jetzt sehr homogen wirken (auf den gleichen Wert abgebildet wurden), speziell die Details in den Haaren und im Auge allerdings immer noch erhalten blieben. Welcher Ansatz der geeigneter ist hängt dabei vom speziellen Anwendungsfall ab.

Analog zu diesen Beispielen und Überlegungen, gibt es eine Vielzahl von verschiedenen *Tonemapping Operatoren* (TMO. Hier im Plural: TMOs) in der Literatur zu finden, die sich nach zwei Hauptkriterien unterscheiden lassen:

- Lokale TMOs
- Globale TMOs

Lokale TMOs passen den Wert eines Pixels in starker Abhängigkeit zu dessen Nachbarschaft im Bild an. Dies führt dazu, dass pixelweise Kontrastverhältnisse, bzw. Helligkeitsverhältnisse



Abbildung 5.4.: Der volle Dynamikumfang aus dem linken Teil der Grafik (256 Werte) wurde mit Hilfe eines globalen Tone Mappings auf einen drei mal kleineren Dynamikumfang im mittleren Teil der Grafik reduziert. Der rechte Teil wurde mittels eines nicht linearen Mappings auf den gleichen Umfang reduziert.^a

^aAusgangsgrafik entnommen aus der University of Southern California SIPI Image Database <http://sipi.usc.edu/database/?volume=misc&image=12>

erhalten bleiben und somit feine Details auf den Bildern klar zu erkennen sind. Allerdings führt dies auch oft dazu, dass Pixel die im HDR Bild den gleichen Pixelwert besitzen, im LDR Bild unterschiedlich eingefärbt sein können, abhängig von deren jeweiliger Nachbarschaft. Dieser Effekt zeigt sich oft durch das Auftreten sogenannter „Halos“ (Heiligenscheine) an den Helligkeitskanten im Bild.

Betrachtet man die Abbildung 5.5 so erkennt man speziell an den Rändern der Brücke diese typischen Halo-Artefakte. Hier wurde vom entsprechenden TMO ([FLW02]) erkannt, dass ein sehr dunkler Bildpunkt, direkt an der Kante der Brücke, nahe an einem sehr hellen Bildpunkt im Himmel liegt. Der Kontrast zwischen den beiden Bildpunkten ist im HDR Bild enorm hoch, weswegen die entsprechenden Bildpunkte im resultierenden LDR Bild jeweils aufgehellt bzw. abgedunkelt wurden, um deren subjektiven Kontrast zueinander mittels des geringeren Dynamikumfangs des LDR Bildes anzunähern. Bildpunkte, die nicht mehr in die Nachbarschaft des dunklen Bildpunktes am Brückenrand gefallen sind, haben demnach eine solche Aufhellung nicht erfahren, weswegen sie dunkler gewählt werden. Das Resultat ist ein Helligkeitsgradient den wir als das beschriebene Halo-Artefakt wahrnehmen können.



Abbildung 5.5.: Bei der Anwendung von lokalen Tonemappern (Hier der Tonemapper von Fattal ([FLWo2])) können ungewünschte Artefakte, wie die hier zu sehenden „Heiligenscheine“ auftreten.^a

^aBild Eigentum von Frank Dürr: http://www.frank-durr.de/tonemapping_comparison.html

Ebenso erkennt man aber auch die Vorteile dieses Verfahrens deutlich anhand des gezeigten Beispielbilds. Betrachtet man speziell die Tragestruktur unter der Brücke, so ist diese sehr klar und detailreich zu erkennen, da auch hier die einzelnen Bildpunkte in Abhängigkeit zu deren Nachbarschaft eingefärbt werden. Dies hebt speziell die Kanten zwischen den einzelnen Strukturen deutlich hervor.

Im Gegensatz zu den lokalen TMOs stehen die globalen TMOs, bei denen versucht wird, eine Abbildungsfunktion zu finden, die auf das gesamte Bild gleichermaßen angewendet wird. Durch diese „absolute“ Abbildung von HDR Pixelwert nach LDR Pixelwert ist gewährleistet, dass Bildpunkte mit demselben Pixelwert aus dem HDR Bild auch immer auf denselben Pixelwert im LDR Bild abgebildet werden. Die weiter oben verwendete Beispielfunktion 5.1.2 stellt eine solche Abbildungsfunktion dar. Betrachtet man die dazugehörigen Resultate, wird auch sofort klar, wo die Schwächen dieses Verfahrens liegen. Sie beschreiben Unterschiede in den Helligkeiten von Bildpunkten nur im Zusammenhang zum gesamten Bild und können dadurch auf lokal begrenzte Bereiche mit hohem Kontrast nicht gesondert eingehen.

Dies führt oft dazu, dass solche Bereiche einen geringeren Kontrast aufweisen, als für deren korrekte Darstellung nötig wäre.



Abbildung 5.6.: Bei der Anwendung von globalen Tonemappern (Hier der Tonemapper von Durand ([DDoo])) können Bereiche mit lokal begrenztem, starkem Kontrast eventuell nicht optimal aufgearbeitet werden.^a

^aBild Eigentum von Frank Dürr: http://www.frank-durr.de/tonemapping_comparison.html

Man erkennt hier speziell an der Tragestruktur unter der Brücke diese Problematik. Die Abbildung der Pixelwerte unter der Brücke basiert rein auf deren Helligkeit im Verhältnis zum Rest des Bildes. Da sie in diesem Bezug sehr dunkel sind, werden sie auch auf einen entsprechend niedrigen Helligkeitswert im LDR Bild abgebildet. Ein weiterer wichtiger Vorteil von globalen TMOs im Bezug zu lokalen TMOs ist jedoch, dass globale TMOs meist deutlich weniger komplex sind und dadurch oft sehr viel schneller zu berechnen sind. Dies ist speziell im Bereich von Echtzeitanwendungen entscheidend, wo die Anzahl an berechenbaren Bildern pro Zeiteinheit eine wichtige Rolle spielt. Ebenfalls erzeugen globale TMOs im resultierenden LDR Bild meist einen „realistischeren“ Eindruck als die lokalen Varianten. Mit entsprechender Parameterwahl sind allerdings auch diese lokalen Varianten durchaus in der Lage, ein fotorealistisches Resultat zu erzeugen, das nur sehr wenige bis hin zu gar keinen Artefakten aufweist. Ein weiterer wichtiger Vorteil von globalen TMOs im Bezug zu lokalen TMOs ist jedoch, dass globale TMOs meist deutlich weniger komplex sind und dadurch oft sehr viel schneller zu berechnen sind. Dies ist speziell im Bereich von

Echtzeitanwendungen entscheidend, wo die Anzahl an berechenbaren Bildern pro Zeiteinheit eine wichtige Rolle spielt. Ebenfalls erzeugen globale TMOs im resultierenden LDR Bild meist einen „realistischeren“ Eindruck als die lokalen Varianten. Mit entsprechender Parameterwahl sind allerdings auch diese lokalen Varianten durchaus in der Lage, ein fotorealistisches Resultat zu erzeugen, das nur sehr wenige bis teils gar keine Artefakte aufweist.

5.2. Auswahl der HDR Verfahren

Um ein geeignetes Verfahren für HDR Imaging im Kontext des Focus 3D Scanners auszuwählen, müssen eine Menge von Faktoren berücksichtigt werden, auf die ein entsprechend gestaltetes Verfahren passen sollte. Dazu betrachten wir die Ziele und Belange des Scanvorgangs selbst. Das Resultat eines Scanvorgangs ist eine möglichst realitätsgetreue, 3-dimensionale Abbildung der Wirklichkeit als Punktwolke und sollte möglichst keine künstlichen Effekte und Artefakte enthalten. Dieser Anspruch erhält sich auch bei der Einfärbung der Scandaten. Die realitätsgetreue Abbildung des Farbeindrucks der Scanumgebung soll erhalten bleiben. Jedoch muss immer noch im Blick behalten werden, dass der Focus 3D in erster Linie ein Messgerät ist und als solches die Gewinnung bestimmter Fakten und Daten aus der Scanumgebung eine der höchsten Prioritäten des Gesamtsystems darstellt. Die Hervorhebung von solchen Informationen über das Mittel der Visualisierung mit Farbinformationen ist dabei von enormem Nutzen. Dies wird klar, wenn wir noch einmal das Bild 5.5 betrachten. Durch die deutlichen Artefakte die durch das Tone Mapping in die Szene eingebracht wurden hat das Bild stark an Realismus verloren. Jedoch ist, im Sinne der Dokumentation von tatsächlich sichtbaren Objekten, die deutliche Hervorhebung der Tragestruktur unter der Brücke absolut wünschenswert.

Ein optimales Einfärbungsverfahren muss also in der Lage sein, auf die Ansprüche der realitätsgetreuen Abbildung und des einfach zugänglichen Informationsgehalts gleichermaßen eingehen zu können, ohne dabei ungewollt Informationen zu verfälschen. Ebenfalls muss es sich möglichst nahtlos in den bestehenden Workflow eingliedern, um dem Nutzer den Umgang mit dem Gerät nicht zu erschweren.

5.2.1. HDR Synthese Verfahren

Für das Verfahren der HDR Synthese, also zur Gewinnung des HDR Bildes selbst, stehen zwei grundlegende Verfahren zur Verfügung:

- Direkte Erzeugung mittels einer entsprechenden HDR Kamera
- Berechnung aus mehreren Einzelbildern mit einer LDR Kamera

Der Einsatz einer HDR Kamera hätte einige Vorteile für das Verfahren, da die Erzeugung von HDR Bildern mit solch einer Kamera bedeutend schneller und weniger störanfällig ist als die Berechnung aus Einzelbildern. Einzelbelichtungen müssen der Reihe nach aufgenommen

werden, was je nach gewünschter Menge an unterschiedlichen Belichtungszeiten einige Zeit in Anspruch nehmen kann. Ebenfalls muss bei der Aufnahme von Belichtungsreihen für die HDR Berechnung stark darauf geachtet werden, dass alle aufgenommenen Einzelbilder an exakt derselben Position aufgenommen sind, da ansonsten bei der Berechnung des HDR Bildes Unschärfe entstehen kann.



Abbildung 5.7.: Unschärfe, die entstehen kann, wenn die verwendeten LDR Bilder nicht sauber zueinander ausgerichtet sind oder wenn sich Objekte in der Szene während den Aufnahmen bewegen.^a

^aBelichtungsreihe für die Erzeugung der Radiance Map, Eigentum von Erik Reinhard: http://www.cs.bris.ac.uk/~reinhard/tm_comp/flicker_hdr/Flicker%20HDR.html

Diese Unschärfe entsteht dadurch, dass bei der Berechnung von Radiance Maps, unter anderem, über alle Eingangsbilder gemittelt wird. Sind die Eingangsbilder nicht korrekt zueinander ausgerichtet, werden Bildpunkte miteinander verrechnet, die nicht zueinander gehören, was zu deutlichen Verfälschungen im Ergebnis führt, das sich als Unschärfe an den falsch ausgerichteten Bereichen äußert. Eine vorherige Ausrichtung der Bilder ist zwar möglich und man kann entsprechende Verfahren, die diesen Prozess automatisieren, in der Literatur finden, jedoch sind solche Verfahren enorm aufwändig und verbrauchen sehr viel Rechenzeit. Der Vorteil bei der Aufnahme von Einzelbildern ist jedoch, dass nach der Aufnahme bereits mehrere LDR Bilder vorhanden sind, die als Preview oder als Backup verwendet werden können. Fallen bei der Aufnahme ausschließlich HDR Bilder an, so müssen diese immer zuerst Tonemapped oder anders aufbereitet werden, bevor die Ergebnisse auf einem konventionellen Display dargestellt werden können. Bei der Aufnahme von LDR Bildern könnte der Nutzer nachträglich noch entscheiden, eine der Einzelbelichtungen für die Einfärbung zu verwenden. Diese Möglichkeit entfällt bei der ausschließlichen Aufnahme von HDR Bildern.

Da im Focus 3D nur eine gewöhnliche LDR Kamera verbaut ist, würde eine Umrüstung auf eine HDR Kamera einen erheblichen Integrationsaufwand bedeuten. Aus diesem Grund beschränken wir uns bei der Wahl des Verfahrens für die HDR Erzeugung auf die Berechnung aus Einzelbelichtungen. Dazu wird das Verfahren aus Kapitel 5.2.1. (Debevec: [DM97]) angewendet. Das Verfahren von Robertson ([RBS99]) zeigte sich als weniger günstig für unseren Anwendungsfall, da es bedeutend rechenaufwändiger ist als das Verfahren von Debevec und keine bis kaum Kontrolle über den Verlauf der eingesetzten Camera-Response-Kurve (5.1.1) zulässt.

5.2.2. Auswahl des Tonemapping Verfahrens

Für die Auswahl des Tonemapping Verfahrens wurde eine Vielzahl von TMOs untersucht und verglichen. Hierzu kam hauptsächlich die Freie Software „Luminance HDR“² von Davide Anastasia zum Einsatz. Da leider kein ausschließlich objektives Verfahren zur automatisierten und objektiven Bestimmung der Qualität des Resultats eines Tonemappings existiert, müssen die Resultate hauptsächlich nach ihrem subjektiven Eindruck bewertet werden. Dazu wurde eine Mitarbeiter-Befragung durchgeführt, bei der die Resultate mehrerer TMOs mit jeweils demselben HDR Bild als Grundlage zum Vergleich präsentiert wurden. Die TMOs, die für diese Befragung ausgewählt wurden, waren Reinhard02, Mantiuko8 und zusätzlich das optimal belichtete Einzelbild der Belichtungsreihe, aus der das zugrunde liegende HDR Bild berechnet wurde. Die oben genannten TMOs wurden zuvor bereits grob gegeneinander verglichen und sind aus verschiedenen Gründen in die engere Auswahl für die Umfrage gekommen, die hier kurz erläutert werden sollen:

Reinhard02:

- Simple mathematische Grundlagen
- Sehr schnell berechenbar
- Globales Tonemapping sorgt für ansprechendes Resultat
- Geringe Menge an Parametern die für die Berechnung notwendig sind
- Notwendige Parameter können sich anhand der Bilddaten im Voraus sehr gut annähern lassen
- Globales Tonemapping führt unter Umständen zu Bildbereichen mit ungünstiger Abbildung. Besitzt jedoch eine lokal arbeitende Abwandlung die sehr wenig Artefakte erzeugt

Mantiuko8:

- Mit moderatem Zeitaufwand berechenbar
- Globales Tonemapping sorgt für ansprechendes Resultat

²<http://qtpfsgui.sourceforge.net/>

- Sehr komplexe mathematische Grundlagen
- Geringe Menge an Parametern die für die Berechnung notwendig sind
- Durch die geringe Zahl an Parametern auch geringe Anpassbarkeit
- Standardparameter liefern akzeptables Ergebnis

Andere Tonemapper, wie Fattal ([FLW02]), Durand ([DD00]), Pattanaik ([PFFG98]) und Mantiuk06 ([MMS06]) wurden bei der Befragung nicht mit angeboten, da sie im Voraus für weniger geeignet erachtet wurden. Diese Tonemapper produzierten entweder zu unrealistische Ergebnisse, wären enorm komplex zu integrieren, benötigen eine große Menge an Parametern oder sind nur schwer so zu parametrisieren, dass sie ein visuell ansprechendes Resultat erzeugen. Ebenfalls war die teils sehr hohe Laufzeit der Tonemapper ein wichtiges Ausschlusskriterium.

Den Mitarbeitern wurde bei der Umfrage nicht bekannt gegeben welches Resultat durch welches Verfahren entstanden ist und jeder Befragte sollte sich für eines der Tonemapping Resultate entscheiden, das seiner Meinung nach das optimalste Ergebnis liefert. Hierbei waren alle Beteiligten in der Wahl ihrer Bewertungskriterien völlig frei.

Von den etwa 40 Befragten entschied sich eine knappe Mehrheit für das Verfahren von Mantiuk. Nur einige Stimmen weniger erhielt das Verfahren von Reinhard. Die optimale Einzelbelichtung erhielt nur eine Stimme.

Da sich die beiden Tonemapping Verfahren von ihrer Stimmanzahl nur unerheblich unterschieden, entscheiden wir uns für das Verfahren von Reinhard, da dieses deutlich weniger komplex ist als das Verfahren von Mantiuk und zusätzlich eine Möglichkeit bietet, die benötigten Parameter für das Tonemapping, analytisch anhand der HDR Daten zu bestimmen. Dies vereinfacht dem Nutzer den Umgang mit dem System erheblich, da auf diese Weise nur noch dann eine Interaktion erforderlich wird, falls die automatisch bestimmten Parameter kein ausreichend gutes Resultat produzieren sollten.

Weiterführende Untersuchungen zur Güte von verschiedenen TMOs wurden ebenfalls von P. Ledda. et. all. ([LCTSo5]) und F. Drago et. all. ([DMMS03]) durchgeführt. Diese Untersuchungen bestärken unsere Entscheidung für Reinhard's TMO zusätzlich, denn in beiden Untersuchungen schnitt Reinhard's TMO ebenfalls gut ab.

5.3. Reinhard's Tonemapper

Bei diesem Tonemapping Verfahren handelt es sich um einen globalen Tonemapper der 2002 von E. Reinhard et. all. entwickelt wurde. Das Verfahren arbeitet dabei ausschließlich auf den Helligkeitswerten der Farben (Luminance). Um die Helligkeit einer Farbe zu verändern, ohne ihren chromatischen Eindruck zu verfälschen, muss diese zuerst in einen entsprechenden Farbraum konvertiert werden, in dem die Helligkeit einer Farbe unabhängig von ihren

5. High Dynamic Range Imaging

Chromatizitäten gegeben ist. Ein solcher Farbraum ist der CIE xyY^3 Farbraum, in dem der Y Kanal für die Helligkeit der Farbe steht und x und y jeweils die entsprechenden Chromatizitäten in Komplementärfarben codieren. Die Abbildungsfunktion von normalem sRGB (der von unserer Kamera verwendete Farbraum) nach CIE xyY ist im Anhang A.1.1 zu finden.

Für die Dynamikkompression bedient sich Reinhard's Tonemapper bei einigen Methoden aus der klassischen Photographie und des Schwarz-Weiß-Drucks. Genauer greift Reinhard dabei auf das von Ansel Adams und Fred Archer begründeten „Zonensystem“⁴ zurück. Dieses System wurde entwickelt, um es einem Fotografen zu erleichtern, eine Abschätzung darüber zu machen wie eine aufgenommene Szene nach dem Druck ausfallen wird. Dazu teilt Adams Zonensystem alle Grauwerte in 11 Zonen ein, von 0 bis 10 (in der Literatur als römische Ziffern dargestellt), bei denen die Zone 0 ein Grau mit 0% Reflektivität (pures Schwarz) und die Zone X ein Grau mit 100% Reflektivität (pures Weiß) beschreibt. Dabei sind die Reflektivitäten so verteilt, dass sich jeweils alle 2 Zonen die Reflektivität verdoppelt. Die Annahme der Verdopplung der Reflektivität pro Zone beruht dabei auf der Erkenntnis, dass sich die menschliche Wahrnehmung von Helligkeit in etwa logarithmisch verhält (Weber-Fechner-Gesetz).



Abbildung 5.8.: Einteilung der Grauwerte in Adams Zonensystem.

Aus den Erkenntnissen des Zonensystems kann abgeleitet werden, dass ein mittleres Grau (empfundenermaßen genau auf „halber Strecke“ zwischen Schwarz und Weiß) einem „physikalischen“ Grau mit 18% Intensität entspricht. Reinhard's Tone Mapper zielt darauf ab, das empfundene Mittelgrau einer Radiance Map auf ein 18%iges Grau im LDR Bild abzubilden. Dazu wird zuerst das logarithmische Mittel über alle N , in der Radiance Map vorkommenden Werte gebildet.

$$(5.9) \quad \bar{L}_w = \exp \left(\frac{1}{N} \sum_{x,y} \log (\delta + L_w(x, y)) \right)$$

³<http://de.wikipedia.org/wiki/CIE-Normvalenzsystem>

⁴<http://de.wikipedia.org/wiki/Zonensystem>

Der Wert δ steht hierbei für einen sehr kleinen Wert, größer 0, der die Singularität verhindern soll, die bei $\log(0)$ auftritt. Danach werden alle Werte der Radiance Map so skaliert, dass der ermittelte, geometrische⁵ Mittelwert exakt auf 0.18 (18%) abgebildet wird.

$$(5.10) \quad L(x, y) = \frac{a}{\bar{L}_w} L_w(x, y)$$

Der Wert a steht hierbei für den so genannten „Key-Value“ und wird, im Standardfall, auf den erwähnten Wert von 0.18 gesetzt. In manchen Fällen kann es vorkommen, dass der logarithmische Mittelwert einer Radiance Map weit ab von einem empfundenen Mittelgrau ist. In solchen Fällen wäre eine Skalierung auf 0.18 nicht sinnvoll, da die resultierenden Werte sonst entweder zu dunkel oder zu hell werden würden.

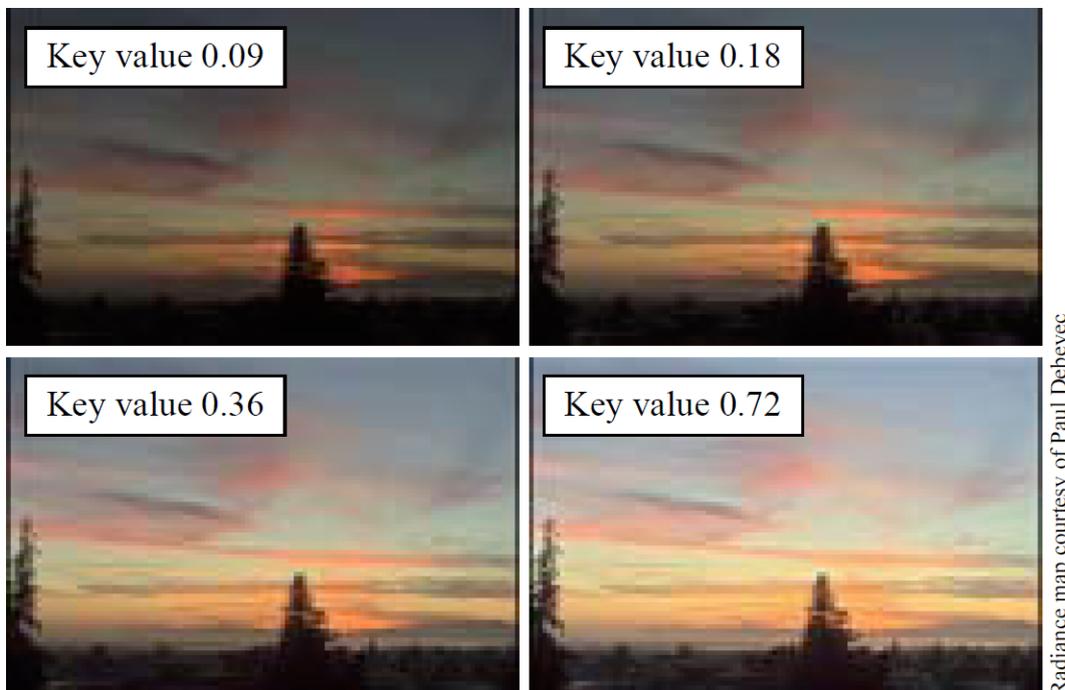


Abbildung 5.9.: Diese Abbildung aus Reinhard's Paper ([RSSFo2]) veranschaulicht den Einfluss des Key Value auf das Tonemapping

Der Wert 0.18 ist daher einer der Nutzerparameter von Reinhard's Tonemapper. Fällt eine Szene enorm hell aus (weißer Hase im Schnee), so sollte der Key-Value entsprechend höher gewählt werden. Analog müsste bei einer sehr dunklen Szene der Key-Value kleiner gewählt werden.

⁵Reinhard spricht in seinem Paper ([RSSFo2]) hier von einem „logarithmischen“ Mittel. Tatsächlich entspricht seine Formulierung dieses Mittelwerts aber genau dem sogenannten „geometrischen“ Mittel.

5. High Dynamic Range Imaging

Letztendlich müssen die skalierten Werte noch gemapt werden. Für das Tonemapping kommt in Reinhard's Tonemapper eine recht einfache Grundfunktion zum Einsatz, die dann entsprechend an das jeweilige Tonemapping Problem angepasst wird.

$$(5.11) \quad L_d(x, y) = \frac{L(x, y)}{1 + L(x, y)}$$

Diese primitive Funktion besitzt bei Eingangswerten nahe 0 eine sehr starke Steigung, die abflacht, je größer die Eingangswerte werden. Auf diese Weise werden sehr kleine Eingangswerte enorm aufgefächert (dunkle Bildpunkte werden zunehmend aufgehellt) und große Eingabewerte werden stark gestaucht (helle Bildpunkte werden zunehmend abgedunkelt). Eine Skizze der Funktion ist in Abbildung 5.10 zu finden.

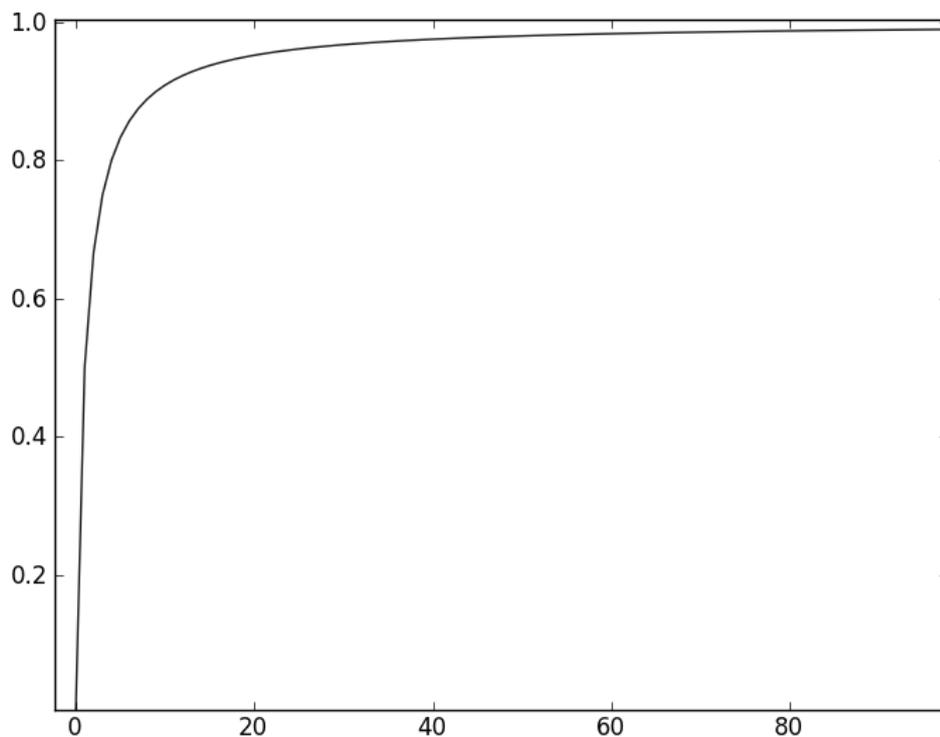


Abbildung 5.10.: Grundgestalt von Reinhard's Tonemapping Funktion

Da der Verlauf dieser Funktion sich dem Maximalwert von 1 nur asymptotisch annähert, wird der volle Wertebereich von $[0, 1]$ niemals voll ausgeschöpft werden. Außerdem möchten wir steuern können, wie stark in den entsprechenden Wertebereichen der Eingangswert gefächert oder gestaucht werden soll. Zu diesem Zweck erweitert Reinhard die Tonemapping Funktion derart, dass gesteuert werden kann bei welchem Eingabewert ein Ausgabewert von 1.0 erreicht werden soll. Hiermit lässt sich effektiv der Wertebereich der abzubildenden Eingangswerte definieren.

$$(5.12) \quad L_d(x, y) = \frac{L(x, y) \left(1 + \frac{L(x, y)}{L_{white}^2}\right)}{1 + L(x, y)}$$

Den Wert L_{white} nennt Reinhard den „Weißpunkt“. Er ist der zweite wichtige Nutzerparameter von Reinhard's Tonemapper. Der Weißpunkt gibt die Schwelle an, ab der alle Werte auf pures Weiß abgebildet werden sollen. Entspricht der Eingabewert der Funktion 5.12 genau dem Weißpunkt, wird das Resultat auf genau 1.0 abgebildet. Durch diese Erweiterung ist es jetzt allerdings nicht mehr gewährleistet, dass alle abgebildeten Werte in das Intervall $[0,1]$ fallen, weswegen nachträgliches Clipping nötig ist (Alle Werte größer als 1.0 werden auf genau 1.0 abgebildet). Durch die beiden Nutzerparameter (Key-Value α und Weißpunkt L_{white}) kann letztendlich die Krümmung von Reinhard's Tonemapping Kurve gesteuert werden.

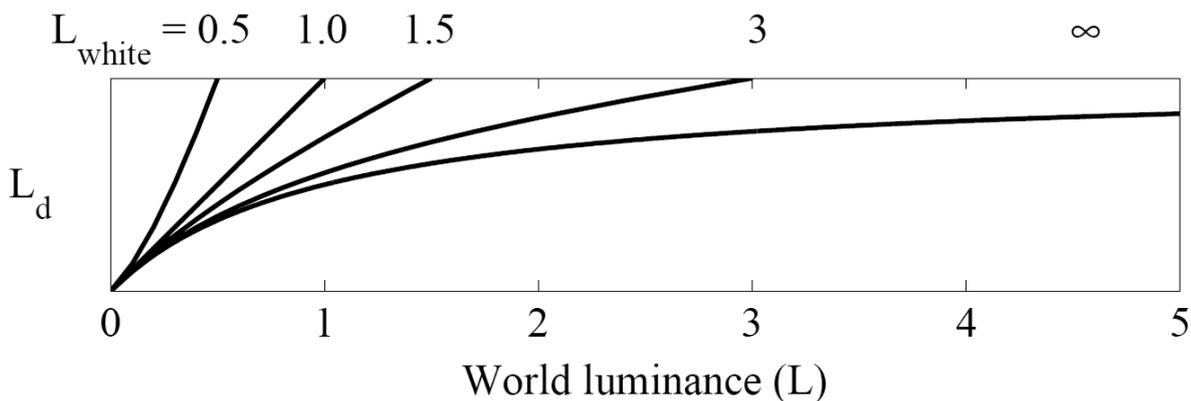


Abbildung 5.11.: Diese Grafik wurde aus Reinhard's Paper ([RSSFo2]) entnommen und veranschaulicht den Einfluss der verschiedenen Weißpunkte auf die Krümmung der resultierenden Tonemapping Kurve haben.

Um diese beiden Parameter möglichst automatisch zu bestimmen und einem unerfahrenen Nutzer somit den Umgang mit dem Tonemapper zu vereinfachen, untersuchte Reinhard in seinem Paper [Reio2] wie diese beiden Werte möglichst analytisch ermittelt werden können. Eigene Tests mit seinen Ergebnissen haben gezeigt, dass diese automatische Parameter Abschätzung in den allermeisten Fällen ein akzeptables Ergebnis liefert. Die jeweiligen Formeln für Key Value und Weißpunkt können im Anhang A.2 gefunden werden.

Reinhard's Tonemapper besitzt außerdem noch eine Abwandlung, die nicht global, sondern lokal arbeitet. Diese definiert einen (einstellbaren) Bereich um den zu mappenden Pixel und versucht einzuordnen, ob sich dieser in einer Umgebung mit starken lokalen Kontraständerungen befindet oder nicht. Abhängig von dieser Einschätzung wird dann die Abbildung des entsprechenden Pixels beeinflusst. Tests mit Reinhard's eigener Implementierung seines Tonemappers zeigten, dass die Verwendung der lokalen Abwandlung speziell den Kontrast

5. High Dynamic Range Imaging

in hochfrequenten Bildbereichen verbessern kann, während der subjektive Gesamteindruck des Bildes dadurch aber unverändert bleibt. Der Einsatz der lokalen Variante kann deshalb vorerst als eine Möglichkeit zur Optimierung des Resultats eingestuft werden. Aus Zeitgründen konnte in dieser Abhandlung allerdings nur auf die globale Variante von Reinhard's Tone Mapper eingegangen werden.

6. Integration der ausgewählten Verfahren

Nachdem die Verfahren jetzt ausgewählt sind, müssen sie entsprechend in den bestehenden Farbaufnahmeprozess integriert werden. Der geplante HDR Imaging Workflow sieht hierbei folgendermaßen aus:

1. Bestimmung der optimalen Einzelbelichtung
2. Bestimmung der Menge an Einzelbelichtungen
3. Aufnahme der Belichtungsreihe
4. Berechnung der Radiance Maps aus den Einzelbildern
5. Tone Mapping der Radiance Maps um das resultierende LDR Bild zu erhalten.
6. Scan einfärben und Resultate speichern

Die jeweiligen Verfahren für die Bestimmung einer optimalen Einzelbelichtung, der Berechnung der Radiance Maps und das anschließende Tone Mapping wurden bereits in den Kapiteln 5 und 6 behandelt. Die dort vorgestellten Lösungen und Verfahren wurden entsprechend in das System integriert. Die verbleibenden Integrationsschritte, die weiterhin nötig sind um ein komplettes Framework zur Verfügung zu stellen, sind somit noch die Aufnahme der Belichtungsreihe und das Speichern der Resultate.

Die Auswahl der Menge an Einzelbelichtungen für eine HDR Aufnahme wird dem Nutzer im Interface des Scanners als Einstellung zur Verfügung gestellt. Die Untergrenze beträgt 3 Einzelbelichtungen in der gesamten Belichtungsreihe. Die Schrittweite der einzelnen Belichtungen ist dabei abhängig von der geplanten Anzahl an Einzelbelichtungen. Man spricht bei der Variation von Belichtungszeiten für HDR Aufnahmen von sogenannten „Stops“. Hierbei entspricht ein Stop einer Verdopplung der Belichtungszeit von einer Aufnahme zur nächsten. Eine Wahl von 2 Stops pro Zwischenschritt ist für HDR Fotografie ein gängiges Maß. Tests mit dieser Einstellung haben zufriedenstellende Ergebnisse geliefert, weswegen wir diese Einstellung auch für unser Verfahren einsetzen. Die maximale Anzahl an Einzelbelichtungen, die in unserem Anwendungsfall möglich sind, beträgt damit 9 Bilder. Dieser Wert ergibt sich aus der maximalen Belichtungsdauer, die abhängig von dem im Scanner eingesetzten Kamerasensor erreicht werden kann.

Das kameraspezifisch einstellbare Verhältnis zwischen der kürzesten und der längsten Belichtung beträgt ca. Faktor 100.000, was 16.6 Stops und somit 8.3 Zwischenschritten je 2 Stops entspricht.

6. Integration der ausgewählten Verfahren

Sofern der Nutzer eine ungerade Anzahl an Einzelbelichtungen wählt, wird die optimale Einzelbelichtung als die mittlere Belichtungszeit der Belichtungsreihe eingestellt. Andernfalls wird die optimale Einzelbelichtung als „Entwicklungspunkt“ gewählt um den herum alle anderen Belichtungen berechnet werden. Beispiel:

Ermittelte, optimale Einzelbelichtung: 0.2 Sekunden.

Eingestellte Anzahl Einzelbelichtungen: 3 → 0.05s, 0.2s, 0.8s

Eingestellte Anzahl Einzelbelichtungen: 4 → 0.025s, 0.1s, (0.2s), 0.4s, 1.6s

Nun soll die Aufnahme der Belichtungsreihe betrachtet werden. Es gibt zwei Möglichkeiten, diese umzusetzen. Es kann ein kompletter Durchlauf pro ermittelte Belichtungszeit vorgenommen werden, oder es kann ein gemeinsamer Durchlauf für alle Belichtungszeiten vorgenommen werden. Da die Ergebnisse der Bildaufnahmen von der gewählten Reihenfolge der Bildpositionen und Belichtungen nicht weiter beeinflusst werden, können wir hier ausschließlich auf den notwendigen Zeitaufwand für die beiden Methoden eingehen. Dazu betrachten wir den Workflow der Aufnahme eines einzelnen Bildes:

1. Positionierung des Drehspiegels für die Aufnahme
2. Bestimmung der Aufnahmeparameter
3. Aufnahme des Bildes
4. Nachbearbeitung und Übertragung des Bildes

Die reine Zeit, die für die Positionierung des Drehspiegels in einem kompletten Aufnahmedurchlauf nötig ist, beträgt etwa 20 Sekunden. Die Ermittlung der Aufnahmeparameter des Bildes hängt von der Kamera ab und beträgt in unserem Anwendungsfall ca. 2 Sekunden. Die Aufnahme des Bildes selbst benötigt maximal 2 mal die jeweils eingestellte Belichtungszeit. Der Zeitaufwand für Nachbearbeitung und Übertragung hängt vom Bildinhalt ab (Kompressionsaufwand) und von der Bereitschaft des Gesamtsystems, das Bild zu empfangen. In unserem Anwendungsfall bewegt sich dieser Zeitaufwand in der Größenordnung von einigen Millisekunden (typischerweise 250 ms).

Der Zeitaufwand für die Nachverarbeitung und Übertragung der Bilder sowie der Zeitaufwand für die Positionierung des Drehspiegels mindestens eines vollen Durchlaufs lässt sich nicht verhindern. Um Zeit einzusparen soll möglichst nur ein Durchlauf verwendet werden, um die Spiegel Positionierungen nur einmal vornehmen zu müssen. Die Aufnahmeparameter, mit Ausnahme der Belichtungszeit, lassen wir von der Kamera selbst ermitteln. Da im Allgemeinen davon ausgegangen werden kann, dass sich die Art der Lichtquelle während eines Durchlaufs der Bildaufnahme nicht ändert, genügt es meist den Weißabgleich an einer einzelnen Position vorzunehmen und diese ermittelten Werte für alle nachfolgenden Einzelaufnahmen des Durchlaufs gleichermaßen anzuwenden. Der Schwarzlevel der Kamera hängt in erster Linie von der eingestellten Belichtungszeit ab und ist somit ebenfalls auf die nachfolgenden Einzelaufnahmen gleichermaßen anwendbar. Zu deren Ermittlung stellen wir also der Reihe nach die gewünschten Belichtungszeiten ein und lassen die Kamera Weißabgleich und Schwarzlevel automatisch bestimmen. Die

ermittelten Werte werden dann aus der Kamera ausgelesen und abgespeichert. Jedes Mal, wenn die Belichtungseinstellungen für die Kamera geändert werden (für die Aufnahmen der entsprechenden Einzelbelichtungen), werden diese gespeicherten Werte wieder direkt an die Kamera übertragen. Auf diese Weise kann die Zeit gespart werden, die die Kamera benötigen würde, um die Werte bei jeder Bildaufnahme erneut automatisch zu bestimmen.

Durch dieses Vorgehen benötigen wir nur einen einzelnen Durchlauf für die Aufnahme aller Einzelbilder und Belichtungen und es wird pro Aufnahmeposition und Belichtungseinstellung keine zusätzliche Zeit für die automatische Bestimmung der Aufnahmeparameter benötigt. Der gesamte Aufnahmeprozess verlängert sich damit nur um die Dauer der zusätzlichen Aufnahmen der Einzelbelichtungen und um die Zeit, die benötigt wird, um die einzelnen Aufnahmeparameter im Voraus einmal zu bestimmen.

Letztendlich müssen die Resultate des Gesamtprozesses noch in einer Datei gespeichert werden können. Dazu wurden mehrere mögliche Dateiformate zum Speichern von Radiance Maps untersucht. Zu den untersuchten Kandidaten gehören die Formate:

- OpenEXR¹
- LogLuv Tiff²
- JPEG-HDR ([WS05])
- Bitmap³
- Radiance-RGBE ([Warg1])

Diese einzelnen Codierungen unterscheiden sich zum Teil erheblich in der anfallenden Datenmenge und im Umfang der gespeicherten Informationen. So können zum Beispiel im OpenEXR Format eine große Menge an Zusatzinformationen zu dem HDR Bild gespeichert werden, wie Reflektionswerte, Transparenz oder Winkelnormalen, bei gleichzeitig geringem Speicherbedarf. Jedoch ist dieses Format enorm komplex und erhöht den letztendlichen Speicherbedarf für eine Radiance Map unnötig, wenn man die angegebenen Zusatzinformationen nicht verwenden will. Die einzelnen Vor und Nachteile der Formate werden hier gegenübergestellt:

OpenEXR:

- Offener, industrieller Standard zum Speichern von 3D HDR Szenen.
- Implementierungen in freien Bibliotheken sind vorhanden.
- Sehr komplexes Format.
- Mindestens 48bit Speicherbedarf pro Bildpunkt.

¹<http://www.openexr.com/>

²<http://www.anyhere.com/gward/pixformat/tiffluv.html>

³http://en.wikipedia.org/wiki/BMP_file_format

6. Integration der ausgewählten Verfahren

- Unterstützt verschiedenste Komprimierungen.
- Speichert viele, gegebenenfalls unnötige Zusatzinformationen wenn nur ein normales 2D HDR Bild gespeichert werden soll.

LogLuv Tiff:

- Baut auf das gängige TIFF Format auf.
- Implementierungen in freien Bibliotheken ist teilweise vorhanden.
- 24bit Speicherbedarf pro Bildpunkt.
- Sehr grobe Abstufung zwischen den einzelnen gespeicherten Helligkeitswerten (Aliasing)

HDR-JPEG:

- Unterkategorie des bekannten JPEG Bild Standards.
- Sehr gering verbreitet und wird nur selten unterstützt.
- Kaum freie Implementierungen und Bibliotheken vorhanden.
- Speichert ein direkt darstellbares LDR Bild mit.
- Sehr starke Kompression möglich.
- Benötigt sehr komplexe Rückumrechnung um auf Pixelebene bearbeitet werden zu können.

Bitmap-HDR:

- Enorm simple Dateistruktur.
- Ist nach außen hin nicht direkt als HDR Bild zu erkennen.
- Besitzt viele vorgefertigte Implementierungen und Bibliotheken.
- Enorm hoher Speicherbedarf (mind. 96bit pro Bildpunkt).

Radiance RGBE:

- Einfache Dateistruktur.
- Unterstützt rudimentäres Run-Length-Encoding (ca. 30% Speicherplatz Ersparnis)
- Sehr einfach zu implementieren.
- Am weitesten verbreitetes Format zum Speichern von 2D HDR Bildern.
- Geringer Speicherbedarf (32bit pro Bildpunkt)
- Helligkeitswerte sind trotz geringem Speicherbedarf sehr fein abgestuft.

Für die Speicherung unserer erzeugten HDR Bilder wurde sich letztendlich für das Radiance HDR Format entschieden, da es wenig Speicherplatz benötigt, einfach zu implementieren ist und eines der gängigsten HDR Bildformate ist. LogLuv-Tiff, Bitmap-HDR und HDR-JPEG sind wegen ihrer jeweiligen Mängel für unseren Anwendungsfall nicht geeignet. HDR-JPEG wird zu wenig unterstützt und würde damit einem Nutzer den Umgang mit unseren erzeugten HDR Bildern unnötig erschweren, während LogLuv-TIFF für unseren Anwendungsfall nicht ausreichend genau ist. HDR Bitmaps wären in unserem Fall recht gut geeignet, benötigen aber enorm viel Speicherplatz. Letztendlich würde OpenEXR eine lohnenswerte Alternative darstellen, wurde aber wegen seiner Komplexität als zu überdimensioniert für unseren Anwendungsfall erachtet.

7. Zusammenfassung und Ausblick

Bei der Analyse des momentan bestehenden Farbaufnahmeprozesses des FARO Focus 3D konnten einige Mängel festgestellt werden, die die Qualität der resultierenden Farbaufnahmen negativ beeinflussen. Unter anderem kommt eine Messung für die Beleuchtungssituation des umgebenen Scanbereichs zum Einsatz, die technisch bedingte Mängel aufweist, die dazu führen, dass einzelne Bereiche des Raums mehr anteiliges Gewicht an dem Messresultat haben als andere. Dies geschieht speziell dann, wenn sich helle Lichtquellen direkt oberhalb des Scanners befinden. Deshalb wurde eine Messmethodik implementiert, die nahezu überdeckungsfrei einen relevanten Bereich innerhalb des Scanbereichs einzeln vermisst und anhand dieser Vermessung eine möglichst optimale Belichtungszeit ermittelt. Dabei bleibt es dem Nutzer überlassen auszuwählen, wo sich dieser „relevante Bereich“ befindet. Hierzu wurden 3 Standardeinstellungen zur Verfügung gestellt, die sich grob an den 3 relevanten Zonen richten: Wände, Decke und gleichmäßige Verteilung. Durch die gezielte Abtastung dieser relevanten Bereiche konnte speziell in vorher kritischen Beleuchtungssituationen, wie sie vor allem verstärkt in Industriehallen vorkommen, eine deutliche Steigerung der durchschnittlichen Bildqualität erzielt werden.



Abbildung 7.1.: Resultat der alten Methode zur Bestimmung der optimalen Belichtungszeit. Wegen der Lampe direkt oberhalb des Scanners wurde eine zu kurze Belichtungszeit bestimmt.



Abbildung 7.2.: Resultat der neuen Methode zur Bestimmung der optimalen Belichtungszeit. Trotz der Lampe direkt oberhalb des Scanners wurde eine passende Belichtungszeit bestimmt.

Diese Auswahl der relevanten Bereiche geht allerdings nur davon aus, dass sich solche Bereiche immer in einem räumlich gegebenen Zusammenhang befinden. Sollten sich mehrere relevante Bereiche innerhalb der Scankuppel befinden, die in keinem räumlichen Zusammenhang stehen, beispielweise bei Fresken in einer Kirche (Wände + Decke), ist es noch nicht möglich, auf diese einzelnen „Hotspots“ einzugehen. Eventuell wäre eine Auswahl solcher Bereiche anhand eines schnellen Preview-Scans denkbar.

Gänzlich versagt dieses Verfahren allerdings dann, wenn sich innerhalb des relevanten Bereichs sowohl sehr helle, als auch sehr dunkle Bereiche befinden. In solchen Fällen reicht der Dynamikumfang der eingesetzten Kamera nicht aus um sowohl die hellen, als auch die dunklen Bereiche mit nur einer Belichtung gleichermaßen gut abzubilden. Aus diesem Grund wurde eine Methodik zur Aufnahme von HDR Bildern in den Scanprozess integriert. Hierbei können, mit den Werkzeugen normaler LDR Aufnahmen, Bilder aufgenommen bzw. erzeugt werden, die über einen deutlich höheren Dynamikumfang verfügen als die eingesetzte Kamera normalerweise umfasst. Es konnte gezeigt werden, dass solche HDR Aufnahmen in stark unterschiedlichen Beleuchtungssituationen eine sichtbare Steigerung der Bildqualität erzielen konnten und die entsprechenden Mechanismen wurden für den Nutzer einfach zugänglich und unkompliziert in den Scanprozess integriert. Die Abbildungen 7.3 und 7.4 zeigen zwei Resultate dieses Verfahrens.

Die Aufnahme und Erzeugung solcher HDR Bilder erfordert allerdings zusätzliche Aufnahme- und Berechnungszeit. Dies könnte in zukünftigen Versionen des Scanners verhindert werden, indem eine entsprechende HDR Kamera in den Scanner integriert werden würde. Ebenso muss weiterhin untersucht werden, welche Tonemapping verfahren ein möglichst gutes Resultat in Bezug auf den jeweiligen Anwendungsfall liefern können. Der momentan eingesetzte Tonemapper (Reinhardo2) liefert in den meisten Fällen ein brauchbares Resultat, besitzt jedoch die typischen Mängel eines globalen Tonemappers.

Zu den bereits genannten Optimierungsmöglichkeiten konnten weitere Möglichkeiten zur Verbesserung der Farbqualität festgestellt werden, die in der momentanen Verwendung des Kamerasensors begründet sind. Viele kameraspezifische Einstellungen werden momentan durch die Automatismen des Kamerachips selbst bestimmt. Die Mechanismen, anhand derer diese Berechnungen ausgeführt werden, sind jedoch für uns nicht klar ersichtlich und führen daher oft zu unerwarteten Ergebnissen. So konnte zum Beispiel beobachtet werden, dass der automatische Weißabgleich der Kamera unter bestimmten Bedingungen stark schwanken kann. Eine manuelle Bestimmung des Weißabgleichs würde mehr Kontrolle ermöglichen und könnte somit zur Verbesserung der Bildqualität beitragen.

Ebenfalls wäre das Anbringen eines Blitzes oder einer andersartigen aktiven Beleuchtung am Scanner selbst denkbar, um die Bildqualität speziell in dunklen Bedingungen weiter zu verbessern.

Ein weiteres Problem, das sich bei der Aufnahme mehrerer unabhängiger Scans zeigt, sind unterschiedliche Einfärbungen von gleichen Objekten auf mehreren Scans. Während wir bei der Aufnahme der Einzelbilder darauf eingehen, dass gleiche Objekte auf mehreren Bildern einen gleichen Farbeindruck besitzen, können wir diese Forderung noch nicht auf mehrere unabhängige Scans ausweiten. Beispielsweise könnte sich zwischen zwei hintereinander aufgenommenen Scans die Menge an Licht, das auf eine Wand strahlt, ändern, weil eine Wolke vor der Sonne vorbeizieht. In diesem Fall würde diese Wand, sofern sie in beiden Scans zu sehen ist, jeweils anders in ihrem Farbeindruck ausfallen und damit in der Zusammenführung beider Scans einen deutlichen Farbverlauf erzeugen. Der bereits erwähnte schwankende Weißabgleich der Kamera führt ebenfalls zu diesen Problemen. Versuche mit nachträglichen Farbabgleichen haben sich hierbei als äußerst effektiv erwiesen.

Abschließend kann also gefolgert werden, dass die umgesetzten Änderungen am Farbaufnahmeprozess des Focus 3D zu einer teils deutlichen Steigerung der Bildqualität innerhalb eines einzelnen Scans beitragen konnten. Im Hinblick auf möglichst gleichbleibenden Farbeindruck über mehrere Scans hinweg ist jedoch immer noch viel Verbesserungspotential vorhanden.

7. Zusammenfassung und Ausblick



Abbildung 7.3.: Gegenüberstellung von optimaler Einzelbelichtung (oben) zur fertig gemappten HDR Aufnahme (unten). Speziell im Himmel sind deutlich mehr Details zu erkennen. Der subjektive Farbeindruck des Gesamtbildes blieb nahezu erhalten.



Abbildung 7.4.: Gegenüberstellung von optimaler Einzelbelichtung (oben) zur fertig gemappten HDR Aufnahme (unten). Beim Tonemapping sind einige Rauschartefakte im Himmel entstanden. Die restliche Szene konnte allerdings vom Tonemapping profitieren.

A. Anhang

A.1. Anhang 1

A.1.1. Umrechnung von sRGB nach CIE xyY

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.4124564 & 0.3575761 & 0.1804375 \\ 0.2126729 & 0.7151522 & 0.0721750 \\ 0.0193339 & 0.1191920 & 0.9503041 \end{pmatrix} \times \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

$$\begin{aligned} x &= \frac{X}{X+Y+Z} \\ y &= \frac{Y}{X+Y+Z} \\ Y &= Y \end{aligned}$$

A.1.2. Rückumrechnung von CIE xyY nach sRGB

$$\begin{aligned} X &= \frac{Y}{y}x \\ Y &= Y \\ Z &= \frac{Y}{y}(1 - x - y) \end{aligned}$$

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 3.2404542 & -1.5371385 & -0.4985314 \\ -0.9692660 & 1.8760108 & 0.0415560 \\ 0.0556434 & -0.2040259 & 1.0572252 \end{pmatrix} \times \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

A.2. Anhang 2

A.2.1. Analytische Näherung der Parameter für Reinhard's Tonemapper

L_{max} und L_{min} stehen hierbei für den jeweiligen vorkommenden Maximal- und Minimalwert innerhalb der Radiance Map. \bar{L}_w steht für den Logarithmischen Mittelwert.

Key-Value:

$$\alpha = 0.18 * 4 \left(\frac{2 * \log_2 \bar{L}_w - \log_2 L_{min} - \log_2 L_{max}}{\log_2 L_{max} - \log_2 L_{min}} \right)$$

Weißpunkt: Key-Value:

$$L_{white} = 1.5 * 2^{(\log_2 L_{max} - \log_2 L_{min} - 5)}$$

Literaturverzeichnis

- [DD00] F. Durand, J. Dorsey. Interactive Tone Mapping. In *Proceedings of the Eurographics Workshop on Rendering Techniques 2000*, S. 219–230. Springer-Verlag, London, UK, UK, 2000. URL <http://dl.acm.org/citation.cfm?id=647652.732137>. (Zitiert auf den Seiten 37 und 41)
- [DM97] P. E. Debevec, J. Malik. Recovering high dynamic range radiance maps from photographs. In *Proceedings of the 24th annual conference on Computer graphics and interactive techniques, SIGGRAPH '97*, S. 369–378. ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 1997. doi:10.1145/258734.258884. URL <http://dx.doi.org/10.1145/258734.258884>. (Zitiert auf den Seiten 29 und 40)
- [DMMS03] F. Drago, W. L. Martens, K. Myszkowski, H.-P. Seidel. Perceptual evaluation of tone mapping operators. In *ACM SIGGRAPH 2003 Sketches & Applications, SIGGRAPH '03*, S. 1–1. ACM, New York, NY, USA, 2003. doi:10.1145/965400.965487. URL <http://doi.acm.org/10.1145/965400.965487>. (Zitiert auf Seite 41)
- [FLW02] R. Fattal, D. Lischinski, M. Werman. Gradient domain high dynamic range compression. *ACM Trans. Graph.*, 21(3):249–256, 2002. doi:10.1145/566654.566573. URL <http://doi.acm.org/10.1145/566654.566573>. (Zitiert auf den Seiten 35, 36 und 41)
- [LCTS05] P. Ledda, A. Chalmers, T. Troscianko, H. Seetzen. Evaluation of tone mapping operators using a High Dynamic Range display. *ACM Trans. Graph.*, 24(3):640–648, 2005. doi:10.1145/1073204.1073242. URL <http://doi.acm.org/10.1145/1073204.1073242>. (Zitiert auf Seite 41)
- [MMS06] R. Mantiuk, K. Myszkowski, H.-P. Seidel. A perceptual framework for contrast processing of high dynamic range images. *ACM Trans. Appl. Percept.*, 3(3):286–308, 2006. doi:10.1145/1166087.1166095. URL <http://doi.acm.org/10.1145/1166087.1166095>. (Zitiert auf Seite 41)
- [PFFG98] S. N. Pattanaik, J. A. Ferwerda, M. D. Fairchild, D. P. Greenberg. A multiscale model of adaptation and spatial vision for realistic image display. In *Proceedings of the 25th annual conference on Computer graphics and interactive techniques, SIGGRAPH '98*, S. 287–298. ACM, New York, NY, USA, 1998. doi:10.1145/280814.280922. URL <http://doi.acm.org/10.1145/280814.280922>. (Zitiert auf Seite 41)
- [RBS99] M. A. Robertson, S. Borman, R. L. Stevenson. Estimation-Theoretic Approach to Dynamic Range Enhancement Using Multiple Exposures. *Journal of Electronic Imaging*, 12:2003, 1999. (Zitiert auf den Seiten 29 und 40)

- [Reio2] E. Reinhard. Parameter estimation for photographic tone reproduction. *J. Graph. Tools*, 7(1):45–52, 2002. doi:10.1080/10867651.2002.10487554. URL <http://dx.doi.org/10.1080/10867651.2002.10487554>. (Zitiert auf Seite 45)
- [RSSF02] E. Reinhard, M. Stark, P. Shirley, J. Ferwerda. Photographic tone reproduction for digital images. *ACM Trans. Graph.*, 21(3):267–276, 2002. doi:10.1145/566654.566575. URL <http://doi.acm.org/10.1145/566654.566575>. (Zitiert auf den Seiten 43 und 45)
- [War91] G. Ward. *Graphic Gems II*, Kapitel "Real Pixels", S. 80–83. Academic Press, 1991. (Zitiert auf Seite 49)
- [WS05] G. Ward, M. Simmons. JPEG-HDR: a backwards-compatible, high dynamic range extension to JPEG. In *ACM SIGGRAPH 2005 Courses*, SIGGRAPH '05. ACM, New York, NY, USA, 2005. doi:10.1145/1198555.1198708. URL <http://doi.acm.org/10.1145/1198555.1198708>. (Zitiert auf Seite 49)

Alle URLs wurden zuletzt am 07.12.2012 geprüft.

Erklärung

Ich versichere, diese Arbeit selbstständig verfasst zu haben.

Ich habe keine anderen als die angegebenen Quellen benutzt und alle wörtlich oder sinngemäß aus anderen Werken übernommenen Aussagen als solche gekennzeichnet.

Weder diese Arbeit noch wesentliche Teile daraus waren bisher Gegenstand eines anderen Prüfungsverfahrens.

Ich habe diese Arbeit bisher weder teilweise noch vollständig veröffentlicht.

Das elektronische Exemplar stimmt mit allen eingereichten Exemplaren überein.

(Timo Dritschler)