



Modellierung und Anpassung von Bildeigenschaften zwischen Aufnahmen durch unterschiedliche Kameras

Studienarbeit

HUMBOLDT-UNIVERSITÄT ZU BERLIN
MATHEMATISCH-NATURWISSENSCHAFTLICHE FAKULTÄT
INSTITUT FÜR INFORMATIK

eingereicht von: Clemens Peter Seibold

geboren am: 01.05.1989

in: Berlin

Betreuer: M.A. Markus Kettern

Gutachter: Prof. Dr.-Ing. Peter Eisert

eingereicht am:

Abstract

The adaption and modeling of image properties is an important factor in image editing and image processing. For example, color adaption can be used to smooth the intersecting regions of a panoramic photography. This study presents and compares adaption methods for color and sharpness alignment and models to describe the sensor noise. The color alignment is based on the adaption of probability distributions. The difference in sharpness is described by a lowpass filter. The sensor noise is approximated by a polynomial approach and an exponential damped function. The best color adaption results could be obtained by a nonlinear luminance and an affine chrominance transformation. In case of missing data of dark and bright regions, the noise characteristic estimation was more reliable using an exponential damped function.

Zusammenfassung

Die Anpassung und Abschätzung von Bildeigenschaften ist ein wichtiges Werkzeug in der Bildbearbeitung und Bildverarbeitung. Als Beispiel sei die Anpassung der Farben für einen glatten Übergang bei Panoramas genannt. In dieser Arbeit werden Methoden zur Anpassung von Farben, Schärfe und zur Modellierung des Rauschverhaltens von Kameraaufnahmen vorgestellt und verglichen. Zur Farbanpassung wird auf Ansätze fokussiert, die auf Wahrscheinlichkeitsfunktionen basieren. Die Anpassung der Schärfe wird über ein Tiefpassfilter realisiert. Zur Modellierung des Rauschverhaltens werden sowohl Polynome als auch eine exponentiell gedämpfte Funktion verwendet. Die besten Ergebnisse bezüglich der Farbanpassung konnten mit einer nichtlinearen Luminanz- und einer affinen Chrominanztransformation erzielt werden. Das Rauschverhalten konnte bei fehlenden Messwerten in den Extrembereichen der Lichtintensitäten besser mit einer exponentiell gedämpften Funktion beschrieben werden.

Inhaltsverzeichnis

1	Einleitung	2
2	Problembeschreibung	4
2.1	Farben	4
2.2	Unschärfe	6
2.3	Farbrauschen	7
3	Farbanpassung	8
3.1	Übersicht	8
3.2	Verfahren	12
3.2.1	Anpassung per Hauptkomponentenanalyse	12
3.2.2	Lineares Transportproblem nach Monge und Kantorowitsch	12
3.2.3	Nichtlineare Luminanz- und affine Chrominanzanpassung (CIELAB-Farbraum)	14
3.2.4	Nichtlineare Luminanz- und affine Chrominanzanpassung nach PCA	15
3.3	Entfernen von Ausreißern	16
3.4	Iterative Anpassung	18
3.5	Vergleich	20
3.6	Fazit	23
4	Anpassung der Unschärfe	27
5	Modellierung des Farbrauschens	33
5.1	Farbvarianzbestimmung	33
5.2	Approximationsmodelle	36
6	Ergebnisse	40
7	Fazit	48
A	Anhang	49
A.1	Kurzanleitung Software	49
A.2	Übersicht zugehöriger Datenträger	49

Abbildungsverzeichnis

2.1	Farbtrennung vor dem Sensor	5
3.1	Feature-/Block-Matching	9
3.2	Schrittweise Anpassung der Hauptkomponenten	13
3.3	Vergleich zwischen LumHist(PCA) und LumHist(Lab)	16
3.4	Iterative Anpassung	19
3.5	Vergleich zwischen MK-Methode und PCA-Methode (Aufwand)	22
3.6	Farbverschiebung bei der MK-Methode und PCA-Methode	23
3.7	Vergleich zwischen MK-Methode und LumHist(PCA)-Methode	24
3.8	Auswirkung der Ausreißerentfernung	25
4.1	Geglättete Quotienten der Amplituden	29
4.2	Anpassung der Schärfe	32
5.1	Auswahl der Varianzdaten	35
5.2	20 gemessene Camera Response Functions	38
5.3	Approximation der Standardabweichung über den Mittelwert	39
6.1	Vergleich verschiedener Masken; getrennte Anpassung einzelner Regionen	41
6.2	MK-Methode und LumHist(Lab)-Methode, angewendet auf Kopf und Oberkörper	42
6.3	Anpassung des Farbrauschens	44
6.4	Anpassung anderer Motive (Parlamentsgebäude Grenoble)	45
6.5	Anpassung anderer Motive (Kirche Grenoble)	47
A.1	Graphische Oberfläche	50

Kapitel 1

Einleitung

Eine häufige Aufgabe in der Bildbearbeitung ist die Anpassung der Eigenschaften, wie Farbe oder Rauschverhalten. Auch in der Bildverarbeitung werden Anpassungen verschiedener Bilder oder Modelle derer Eigenschaften benötigt, um beispielsweise optimale Parameter eines Verfahrens zu schätzen bzw. für diese Verfahren benötigte Anpassungen der Aufnahmen vorzunehmen. Als Beispiel innerhalb der Bildbearbeitung sei die Farbanpassung vor partiellen Bildüberblendungen und innerhalb der Bildverarbeitung die Bestimmung optimaler Parameter zur Kantenerkennung mit dem Canny-Algorithmus genannt [LFSK06]. Eine Sonderform der Modellierung ist die Kamerakalibrierung. Diese hat die Bestimmung von internen Kameraparametern als primäres Ziel, um somit eine Anpassung oder bessere Vergleichbarkeit zu ermöglichen.

Ziel dieser Arbeit ist es, Modelle für die Eigenschaften eines Bildes bzw. globale Transformationen zur Anpassung zweier Bilder zu testen und anzupassen. Die internen Eigenschaften der Kameras werden nur indirekt zur Modellfindung berücksichtigt. Als globale Transformation ist eine Transformationsvorschrift zu verstehen, die unabhängig von der Position eines Pixels oder einer Region ist und somit nur von deren Farbwerten abhängt. Dies hat den Vorteil, dass eine Transformation durch einen Teil des Bildes bestimmt, aber auf das ganze Bild angewendet werden kann. Adressiert wird in dieser Arbeit die Anpassung der Farben und Schärfe sowie die Modellierung des Farbrauschens.

Betrachtet werden in dieser Arbeit primär Bilder von Gesichtern vor verschiedenen, homogenen Hintergründen. Diese wurden sowohl mit einer Digitalkamera in einer

Auflösung von 2592x1728 als auch mit einer Videokamera in HD 1080 aufgenommen.

Kapitel 2

Problembeschreibung

2.1 Farben

Die Sensoren einer Kamera zählen einen Teil der Photonen, die während der Belichtungszeit auf die Sensoren treffen. Um nicht nur Schwarz-Weiß-Bilder erzeugen zu können, müssen die Farben zuvor gefiltert oder getrennt werden. Die gängigste Methode hierzu ist ein Farbfilter, welcher über die Sensoren gesetzt wird. Somit misst jeder Sensor nur die durch den Filter definierten Wellenlängen. Die am weitesten verbreitete Anordnung der Farbfilter ist die Bayer-Matrix (Abb. 2.1), welche zu 50% aus grün-durchlässigen und zu je 25% aus rot- und blau-durchlässigen Filtern besteht. Eine andere weit verbreitete Form der Farbtrennung ist das 3-Chip-Modell. Hierbei werden die Farben durch ein Prisma getrennt, wodurch jede Grundfarbe auf ein anderes Sensor-Array gebrochen wird. Damit im Fall der Verwendung eines Bayer-Sensors pro Pixel nicht nur die vom Filter durchlässige Farbe dominiert, werden die einzelnen Sensorwerte miteinander verrechnet. Für diesen Prozess, der als Demosaicking bezeichnet wird, existieren mehrere unterschiedliche Verfahren und dazugehörige Patente [RSBI02].

Um durch die Abbildung von den Sensorwerten in einen RGB-Farbraum keine vom Hersteller willkürliche oder kameraabhängige Kodierung zu erhalten, verwenden die meisten Kameras den geräteunabhängigen und standardisierten sRGB-Farbraum. Dennoch führen Aufnahmen derselben Szene, die zwar unter identischen Bedingungen, jedoch nicht mit derselben Kamera aufgenommen wurden, in der Regel zu verschiedenen Ergebnissen. Dies ist sogar bei verschiedenen Kameras desselben Mo-

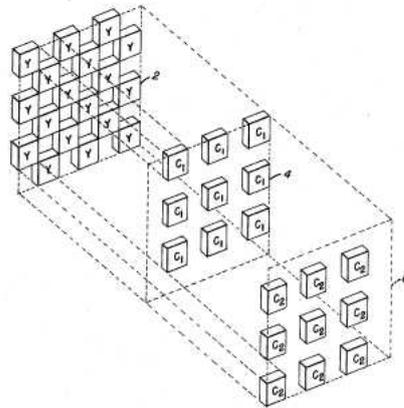


Abbildung 2.1: Farbtrennung vor dem Sensor

In der gezeigten Patentskizze [Bay76] des Bayer-Filters steht jeder Quader für einen Farbfilter. Die Rechtecke Y stehen für Luminanz-sensitive Filter (in der Praxis grüne Filter) und C_1 bzw. C_2 für Chrominanz-sensitive Filter (in der Praxis rote und blaue Filter).

dells zu beobachten. Die Differenzen hierbei sind auf Fabrikationsschwankungen bei den Sensoren bzw. den Farbfiltern und gegebenenfalls das Sensorrauschen zurückzuführen. In einer Studie von Joshi et al. wurden bei Aufnahmen von identischen bzw. sich überschneidenden Szenen mit demselben Kamertyp und nach Kalibrierung der Kameras Farbdifferenzen von bis zu 4% und im Mittel (MSE) ca. 0.5% gemessen. [JWV⁺05]. Die Unterschiede bei verschiedenen Fabrikaten lassen sich durch stark unterschiedliche Filterkennlinien [SDP⁺09] und die interne Verrechnung der einzelnen Sensoren erklären. Noch größer werden die Farbunterschiede identischer Objekte, sowohl bei gleichen als auch unterschiedlichen Kameras, unter verschiedenen Beleuchtungsbedingungen. Eine Ursache hiervon ist die Anpassung der gemessenen Werte an die jeweilige Farbtemperatur. Je nach Farbtemperatur nimmt das menschliche Auge einen anderen Farbton, bzw. spektrale Zusammensetzungen, als weiß wahr. Dies führt dazu, dass unter verschiedenen Lichtquellen wie einer 100W-Glühbirne mit einer Farbtemperatur mit 2800K als auch Mittagssonne mit 5500 – 5800K [Wag14] ein weißes Blatt als weiß erscheint. Diesen Prozess, der in der Physiologie als chromatische Adaption bzw. technisch als Weißabgleich bezeichnet wird, versuchen Kamerahersteller nachzubilden. Dazu muss die Farbtemperatur bzw. der Weißpunkt geschätzt werden [RSYD05]. Je nach Güte diese Schätzung bleiben die Farben jedoch verschieden stark verfälscht. Ein weiterer Faktor bei der Entstehung von Farbdifferenzen sind die Oberflächen der abzubildenden Objekte selbst, welche je nach Lichtstärke und Einfallswinkel verschieden reflektieren. In dieser Arbeit werden

Modelle zur Beschreibung der Farbverteilungen und Möglichkeiten zur Minimierung der Farbdifferenzen vorgestellt.

2.2 Unschärfe

Die Schärfe eines Bildes hängt von vielen Faktoren ab, die ihre Ursache sowohl in der Kamera als auch in der aufgenommenen Szene haben.

Der wohl bekannteste Faktor ist die Schärfentiefe, welche sowohl als von der Kamera als auch von der Szene abhängig betrachtet werden kann. Vor der Aufnahme eines Bildes wird entweder manuell oder automatisch die Distanz zwischen der Linse und den Sensoren eingestellt. Dieser Vorgang, das Fokussieren oder auch Scharfstellen, bestimmt die Entfernung, in der Objekte scharf abgebildet werden. Die Schärfentiefe beschreibt diese Entfernung und den umliegenden Bereich, in dem Objekte scharf abgebildet werden können. Eine Berechnung dieses Bereichs ist mit Hilfe der eingestellten Brennweite, Blendenzahl und Sensorgröße möglich [Jen14]. Die dadurch entstandene Unschärfe kann durch einen Gauß-Filter, dessen Standardabweichung abhängig von der Entfernung zur Kamera ist, modelliert werden [TF05].

Diese allgemein bekannte Tatsache kann bei jeder Kamera bedingt beeinflusst werden. Dazu kommen noch durch den Aufbau und die interne Verarbeitung der Kamera und Einstellungen bezüglich Brennweite und Blendenzahl bedingte Faktoren hinzu, welche sich gegenseitig beeinflussen. Die Beschreibung des Zusammenspiels dieser Faktoren erfolgt durch Modulationsübertragungsfunktionen (MTF). Die MTF beschreiben das Verhältnis des Kontrastes des aufgenommenen Bildes zu dem des realen Objektes in Abhängigkeit der Frequenz [Kor14]. Die Frequenz bezieht sich hierbei auf das auf den Sensor projizierte Bild und wird meist in Linienpaaren pro mm oder Breite der Linien pro Bildhöhe angegeben. Für Kameras gilt prinzipiell, dass mit höherer Frequenz auch der gemessene Kontrast im Verhältnis zum realen Kontrast abnimmt.

Ein weiterer Faktor, der die Schärfe beeinflusst, ist die Bewegungsunschärfe. Diese kann durch Bewegung der Kamera oder der abgebildeten Objekte während der Belichtungszeit entstehen. Zur Beschreibung dieses Effektes werden in der Bildsynthese unter anderem gerichtete Filter, ähnlich einem Mittelwertfilter, verwendet. Die Länge des Filters hängt hierbei von der Entfernung des Objektes zur fiktiven Kamera

und der Bewegung der Objekte bzw. der Kamera ab [Pix14].

Zur Anpassung der Unschärfe wird in dieser Arbeit ein richtungsunabhängiger Tiefpass verwendet, welcher unabhängig von der Position des Pixels bzw. der Entfernung des Objektes ist. Dies adressiert primär die durch Defokussierung planarer Objekte und durch verschiedene MTF bedingte Unschärfe. Als Tiefpass wird ein Gauß-Filter verwendet.

2.3 Farbrauschen

Als Farbrauschen bezeichnet man zufällige und von der Szene unabhängige Veränderungen des Farbwertes eines Pixels. Da dieses Rauschen unmittelbar von den einzelnen Sensoren abhängt, ist es von der kamerainternen Verarbeitung der Sensorwerte abhängig, inwieweit das Rauschen innerhalb eines Bildes bezüglich der Farbkanäle bzw. der Pixelumgebung korreliert. Die Hauptursache des Farbrauschens ist das thermische Rauschen im Sensor. Dieses entsteht durch die wärmebedingte Erzeugung freier Ladungsträger im Sensor und nimmt mit der Temperatur und Belichtungszeit zu. Die meisten Kameras schätzen das thermische Rauschen und ziehen diese Schätzung vom Bild ab [RSYD05].

Weitere, in der Alltagsfotografie vernachlässigbare Rauscheffekte sind das Photonenrauschen, das Quantisierungsrauschen und das Ausleserauschen [Erh08]. Als Photonenrauschen bezeichnet man die ungleichmäßige Verteilung der Photonen auf die einzelnen Sensoren bei gleicher Lichtintensität. Dieser Effekt lässt sich mit einer von der mittleren Lichtintensität abhängigen Poisson-Verteilung beschreiben [Erh08]. Das Quantisierungsrauschen entsteht durch die Digitalisierung der Sensorwerte und die digitale interne Umrechnung der Farbwerte. Das Ausleserauschen entsteht bei der analogen Aufintegration der durch die Photonen freigesetzten Elektronen und liegt bei modernen Digitalkameras bei einigen Elektronen [Aus14]. Dies entspricht bei Isowerten bis zu 800 weniger als 0.5% der maximalen Sättigung.

Das Verhältnis zwischen der Lichtintensität am Sensor und dem dieser zugeordneten Farbwert beschreibt die Camera Reponse Function (CRF). Diese ist in der Regel nichtlinear, weshalb die Stärke des thermischen Rauschens im endgültigen Bild zu einer Funktion in Abhängigkeit der Lichtintensität am Sensor wird.

Kapitel 3

Farbanpassung

3.1 Übersicht

Der Farbwert aus dem anzupassenden Bild bzw. das Bild selbst wird im Folgenden als Quellfarbe bzw. Quellbild bezeichnet. Das Bild, an dessen Farben angepasst werden soll bzw. die dort der Quellfarbe entsprechende Farbe, wird als Zielbild bzw. Zielfarbe bezeichnet. Soweit nicht anders erwähnt, wird im sRGB-Farbraum und mit Farbwerten von 0 bis 255 gearbeitet.

Zur Farbanpassung werden nur globale Transformationen betrachtet, d.h. jeder Pixel wird unabhängig von seiner Position innerhalb der Szene und im Bild transformiert. Schatten oder starke Reflexionen, die in nur einem Bild vorkommen, werden folglich nur indirekt angepasst. Die Ortsunabhängigkeit ermöglicht eine Berechnung der Transformation auf einem unabhängigen Teilbild oder einem anderem als dem anzupassendem Bild. Da sich in dem zur Verfügung stehenden Beispieldatensatz die Hintergründe von Quell- und Zielbild stark unterscheiden und primär die Gesichter angepasst werden sollen, werden nur die Gesichtsregionen zur Bestimmung der Farbtransformation verwendet. Die Auswahl der Gesichtsregionen erfolgte durch eine per Hand erstellte, binäre Maske.

Die Berechnung der Transformationen zur Farbanpassung lassen sich in die mittlere quadratische Abweichung (MSE) minimierende Methoden und die Farbverteilungsfunktion anpassende Methoden einteilen.

Zur Anpassung über die Minimierung der mittleren quadratischen Abweichung wird

eine partielle Zuordnung der Pixel des Quellbildes auf die des Zielbildes benötigt. Diese kann durch die Suche von ähnlichen Blöcken oder Features [Low04] mit anschließender Registrierung erstellt werden. Die auf diese Art entstandenen Farbpaare werden mit einer Funktion approximiert, die als Farbtransformation für das gesamte Bild verwendet werden kann. Die Approximation kann sowohl einzeln für jeden Farbkanal erfolgen als auch eine Funktion der Form $[0, 255]^3 \rightarrow [0, 255]^3$ sein, also jedes RGB-Tupel auf ein anderes abbilden. Als Approximationsfunktionen sind sowohl lineare Funktionen, Polynome als auch Exponentialfunktionen üblich. Da in allen Fällen nur Koeffizienten geschätzt werden, handelt es sich hierbei um ein lineares Optimierungsproblem.



Abbildung 3.1: Feature-/Block-Matching

(a) zeigt die Ergebnisse eines Feature-Matching Verfahrens. Andere Verfahren zur Feature-Bestimmung bzw. zum Feature-Matching und Veränderungen der Parameter führten zu ähnlichen Ergebnissen. Die in (b) dargestellten Ergebnisse eines Block-Matchings konnten auch durch andere Minimierungskriterien oder Parameteränderungen nicht signifikant verbessert werden.

Auf MSE basierende Verfahren werden meist zur Kalibrierung von Kameras benutzt, da diese ohne Annahmen bezüglich der Farbverteilung arbeiten und somit flexibler eingesetzt werden können.

Auf Verfahren, die Pixel-zu-Pixel-Beziehungen benötigen, wird in dieser Arbeit nicht weiter eingegangen, sondern es wird auf Verfahren fokussiert, die auch zu ansprechenden Resultaten führen, sollte keine korrekten Pixel-zu-Pixel-Beziehung bestimmt werden können. Abbildung 3.1 zeigt an einem Beispiel den Versuch eines Korrespondenzaufbaus mit Hilfe von Feature- und Block-Matching.

Die Verfahren bezüglich der Farbverteilung lassen sich wiederum in zwei Klassen unterteilen. Die Verfahren einer Klasse basieren auf der Anpassung statistischer Werte und gehen von bestimmten Verteilungsfunktionen aus; in der anderen Klasse werden Histogramme angepasst und keine Modellannahmen getroffen.

Häufig wird bei auf Verteilungsfunktionen basierenden Verfahren die Annahme getroffen, dass die Verteilung der Farbwerte in Quell- und Zielbild multivariaten Normalverteilungen entsprechen:

$$\mathcal{N}(\mu, \Sigma) = \exp(-0.5(x - \mu)^T \Sigma^{-1} (x - \mu))$$

mit $\mu \in \mathbb{R}^3$ als Mittelwertsvektor und $\Sigma \in \mathbb{R}^{3 \times 3}$ als Kovarianzmatrix. Diese Annahme ist meist erfüllt, wenn das Bild aus mehreren ähnlichen oder einer großen homogenen Fläche besteht. Alternativ lassen sich die Verfahren auch auf Superpositionen von Normalverteilungen anwenden, wenn diese zuvor im Farbraum getrennt werden können.

Zur Anpassung der Bilder werden die Farbwerte des Quellbildes affin transformiert, sodass dessen Normalverteilung bezüglich der Mittelwerte, Varianzen und Kovarianzen mit der des Zielbildes übereinstimmt. Geht man von kontinuierlichen Normalverteilungen aus, lassen sich die Bedingungen, die diese Transformation erfüllen muss, wie folgt bestimmen. Dies zeigt gleichzeitig, dass eine affine Transformation zwei verschiedene Normalverteilungen optimal anpassen kann:

Sei f die Verteilungsfunktion des Quellbildes, g die des Zielbildes, u die Zufallsvariable der Verteilung des Quellbildes und v die des Zielbildes. Ziel ist es, eine affine Transformation $u \rightarrow t(u)$ zu finden, sodass sich $t(u)$ entsprechend der Verteilungsfunktion g verhält. Es gilt:

$$\begin{aligned} f(u)du &= g(v)dv \\ \Leftrightarrow f(u)du &= g(t(u))dt(u) \\ \Leftrightarrow f(u) &= g(t(u))|\det J_t(u)| \end{aligned}$$

Da $t(u)$ die Form $Tu + t_0$ mit T als 3×3 -Matrix und t_0 als 3×1 -Vektor hat, sind die Jacobi-Matrix und ihre Determinante konstant. Es gilt also:

$$f(u) \propto g(t(u))$$

Angenommen u und v sind normalverteilt mit den Mittelwerten μ_u, μ_v und den Kovarianzmatrizen Σ_u, Σ_v , so gilt mit $t_0 = -T\mu_u + \mu_v$:

$$\begin{aligned}
& (t(u) - \mu_v)^T \Sigma_v^{-1} (t(u) - \mu_v) = (u - \mu_u)^T \Sigma_u^{-1} (u - \mu_u) \\
\Leftrightarrow & (T(u - \mu_u))^T \Sigma_v^{-1} (T(u - \mu_u)) = (u - \mu_u)^T \Sigma_u^{-1} (u - \mu_u) \\
\Leftrightarrow & (u - \mu_u)^T T^T \Sigma_v^{-1} T (u - \mu_u) = (u - \mu_u)^T \Sigma_u^{-1} (u - \mu_u) \\
& \Leftrightarrow T^T \Sigma_v^{-1} T = \Sigma_u^{-1} \quad (3.1)
\end{aligned}$$

Jede Transformation, die die Bedingung (3.1) bzw. $T\Sigma_u T^T = \Sigma_v$ erfüllt und t_0 als $-T\mu_u + \mu_v$ definiert, passt folglich die Farbverteilung des Quellbildes der des Zielbildes an, angenommen die Farben im Quell- und Zielbild sind normalverteilt. Diese Transformation ist jedoch nicht eindeutig, ein Beweis dazu findet sich in [PK07].

Eine weitere Methode, die auf der Verteilung der Farbwerte basiert, ist die Anpassung der Histogramme. Um dies zu realisieren, werden von beiden Bildern kanalweise kumulative Histogramme erzeugt. Als Transformationsfunktion dient eine Tabelle, die jedem Farbwert x im Quellbild einen Wert y im Zielbild zuordnet, so dass $H_V(H_U(x))^{-1} = y$ mit H_U als Histogramm des Quellbildes und H_V als Histogramm des Zielbildes. Dies ist jedoch nur kanalweise möglich, da zur Erstellung des kumulativen Histogramms eine totale Ordnung der beschriebenen Menge benötigt wird. Um eine deterministische Histogrammanpassung in \mathbb{R}^3 zu realisieren, würde man zusätzliche Optimierungskriterien benötigen. Ein Vorteil der Histogrammanpassung gegenüber der Anpassung mit Hilfe von Parametern einer Wahrscheinlichkeitsfunktion ist die Unabhängigkeit von den zugrundeliegenden Verteilungsfunktionen. Häufig auftretende Verteilungen wie Superpositionen von Normalverteilungen, Badewannenkurven-förmige Verteilungen und Heavy-Tail-Verteilungen können per Histogrammanpassung ineinander transformiert werden. Ein großer Nachteil ist, dass die einfache Histogrammanpassung nur die Kanäle einzeln betrachtet und somit stark von der Wahl des Farbraumes abhängig ist, während die meisten auf Verteilungsfunktionen basierenden Verfahren unabhängig gegenüber vorherigen orthogonalen Transformationen sind [PK07].

Um starke Sprünge in der Transformationsfunktion der Histogramme zu vermeiden, werden diese in der Praxis mit einem Gauß-Filter geglättet.

3.2 Verfahren

3.2.1 Anpassung per Hauptkomponentenanalyse

Eine Möglichkeit zur Berechnung einer Transformation, die die Bedingung (3.1) erfüllt, ist die Anpassung der Varianzen und Kovarianzen mit Hilfe der Eigenwerte der Kovarianzmatrix Σ_u des Quellbildes und Σ_v des Zielbildes. Hierbei wird zunächst die Kovarianzmatrix der Verteilung des Quellbildes auf eine Einheitsmatrix gesetzt. Mit P_u als Matrix mit den Eigenvektoren von Σ_u als Zeilen und der Diagonalmatrix D_u mit den zugehörigen Eigenwerten ist dies wie folgt möglich:

$$T_1 = P_u^T D_u^{-0.5} P_u$$

Geometrisch betrachtet findet eine Rotation der Normalverteilung statt, sodass die Hauptkomponenten auf den R-/G-/B-Achsen liegen, anschließend eine Skalierung der Varianzen der einzelnen Achsen auf 1 und eine Rotation der R-/G-/B-Achsen zurück auf die Hauptkomponenten, vgl. Abb. 3.2.

Im zweiten Schritt werden durch ein ähnliches Vorgehen die Hauptkomponenten auf die des Zielbildes gesetzt. Der einzige Unterschied besteht in der Skalierung. Die Varianzen der R-/G-/B-Achsen werden auf die Eigenwerte von Σ_v statt auf 1 skaliert:

$$T_2 = P_v^T D_v^{0.5} P_v$$

Als Transformationsvorschrift ergibt sich somit:

$$T_{PCA} = P_v^T D_v^{0.5} P_v P_u^T D_u^{-0.5} P_u \quad (3.2)$$

Eine Anpassung nach dieser Vorschrift wird im folgenden als PCA-Methode bezeichnet.

3.2.2 Lineares Transportproblem nach Monge und Kantorowitsch

Die Anpassung der Normalverteilungen kann als lineares Transportproblem nach Monge und Kantorowitsch betrachtet werden [PK07]. Die geschätzte Normalverteilung des Quellbildes wird hierbei als Menge betrachtet, die unter minimalem Aufwand (3.3) der geschätzten Normalverteilung des Zielbildes angepasst werden

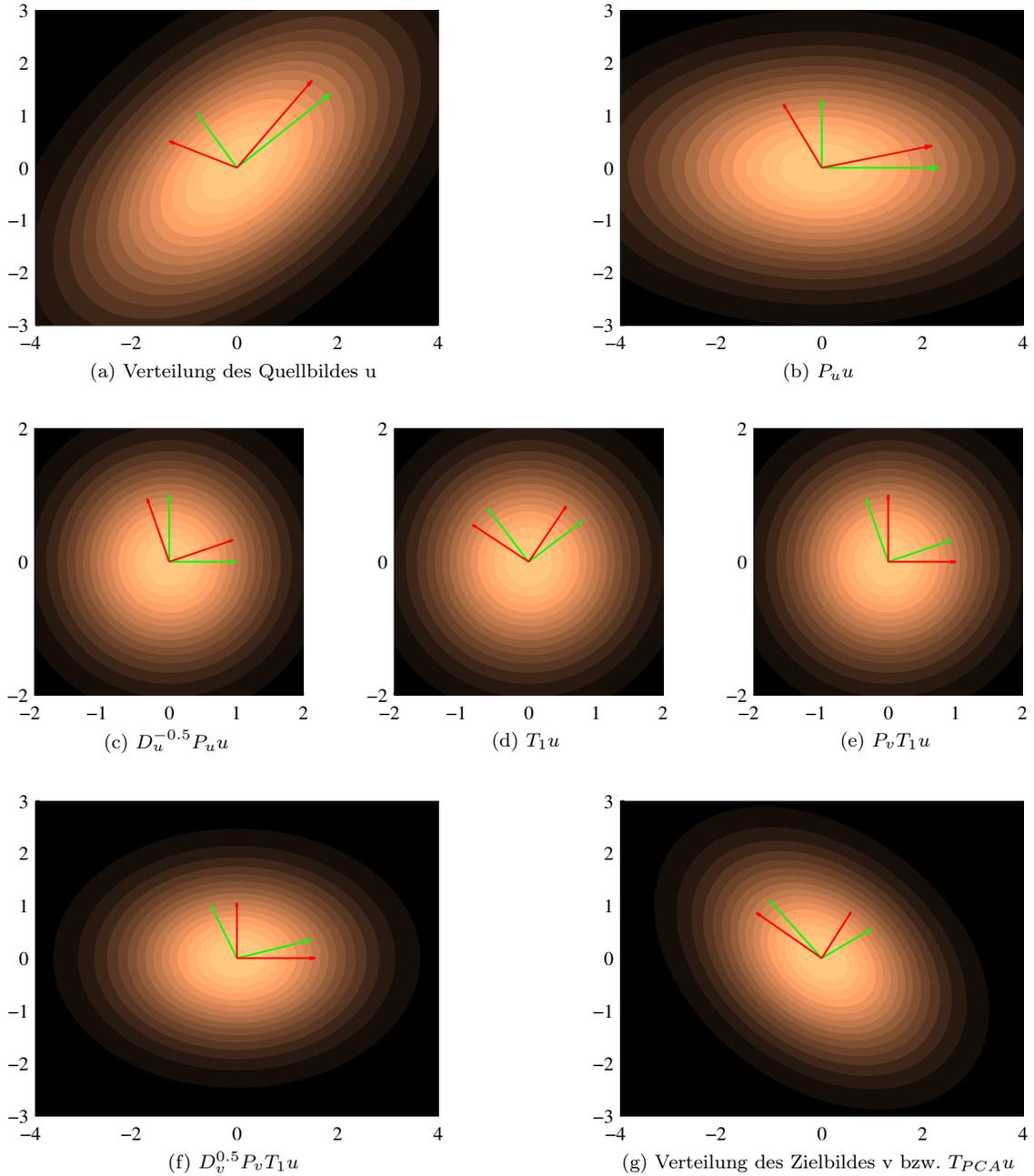


Abbildung 3.2: Schrittweise Anpassung der Hauptkomponenten

Die beiden grünen Vektoren in Abbildung 3.2a stellen die mit den entsprechenden Standardabweichungen skalierten Eigenvektoren der Kovarianzmatrix des Quellbildes und die roten Vektoren die des Zielbildes nach einer Transformation mit T_{PCA}^{-1} dar. Die Abbildungen 3.2b bis 3.2g zeigen die Normalverteilung und die Vektoren aus 3.2a nach der jeweiligen Transformation.

soll. Als Aufwand wird hierbei die Verschiebung eines Farbwertes, im Sinne der Euklidischen Distanz in \mathbb{R}^3 , gewichtet mit seiner Häufigkeit im Quellbild betrachtet:

$$I[t] = \int_u \|t(u) - u\|^2 f(u) du \quad (3.3)$$

Die Lösung dieses Optimierungsproblems existiert immer und ist eindeutig und symmetrisch. Ein ausführlicher Beweis hierzu ist in [OP82] zu finden. Als Transformationsvorschrift ergibt sich:

$$T = \Sigma_u^{-0.5} (\Sigma_u^{0.5} \Sigma_v \Sigma_u^{0.5})^{0.5} \Sigma_u^{-0.5} \quad (3.4)$$

Eine Anpassung nach dieser Vorschrift wird im folgenden als MK-Methode bezeichnet.

3.2.3 Nichtlineare Luminanz- und affine Chrominanzanpassung (CIELAB-Farbraum)

In ihrer Veröffentlichung „Example-Based Video Color Grading“ [BSPP13] stellen Bonnel et al. eine Methode zur Anpassung der Farben vor, welche getrennt die Luminanz und die Chrominanz anpasst. Hierzu wird das Bild in den CIELAB-Farbraum [Hof14] mit D65 als Weißpunkt transformiert. Das folgende Verfahren ist stark an die Methode von Bonnel et al. angelehnt.

Vor jeglicher Anpassung werden sowohl das Quellbild als auch das Zielbild in den CIELAB-Farbraum transformiert. Die Anpassung der Luminanz findet mit Hilfe einer Histogrammanpassung wie in Abschnitt 3.1 beschrieben statt. Die Übertragungsfunktion wird mit einem Gauß-Filter mit einer Standardabweichung von 10% der maximal vorkommenden Luminanzdifferenz des Quellbildes geglättet. Zur Anpassung der Chrominanz werden die Farbwerte in verschiedene, in Bezug auf die Anzahl der Pixel gleich große Bänder in Abhängigkeit ihrer Luminanz eingeteilt. Dies soll ermöglichen, dass Regionen mit verschiedenen Helligkeiten auch farblich verschieden angepasst werden können. Um Sprünge innerhalb eines Farbverlaufes zu vermeiden, überschneiden sich die Bänder zu 10%. Für jedes Band wird eine Transformation nach Formel (3.4) für die Chrominanzkanäle bestimmt. Zur Berechnung der Chrominanz, welche sich durch die Überschneidung der Bänder in zwei Bändern befinden, werden die Chrominanz für beide Bändern berechnet und ge-

wichtet addiert. Als Gewichtungsfunktion wird eine lineare Funktion, welche zum Punkt des Eintritts in das benachbarte Band den Wert 1 hat und bis zum Verlassen des eigenen Bandes auf 0 abfällt, verwendet.

Um optimale Resultate zu erhalten, sind die Größe der Überschneidungen, Anzahl der Bänder und Art/Reichweite der Gewichtungsfunktion entsprechend zu wählen. Bonnel et al. verwenden 3 Bänder mit je 10% Überschneidung und eine Gewichtungsfunktion mit kubischen Abfall, welcher sich weit in die benachbarten Bänder erstreckt. Die Aufteilung in drei Bänder wird mit der in vielen Bildbearbeitungsprogrammen üblichen Trennung von dunklen Stellen wie Schatten, normalen Farbtönen und Glanzeffekten begründet.

Durch eine vorherige Analyse der dem Quell- und Zielbild zugrundeliegenden Verteilungsfunktionen können diese Parameter begründet und das Modell ggf. noch um weitere Parameter, wie einzelne Farbabschnitte mit verschiedenen Größen für die einzelnen Bänder, erweitert werden. Diese Methode wird im weiteren Teil der Arbeit mit LumHist(Lab) abgekürzt.

3.2.4 Nichtlineare Luminanz- und affine Chrominanzanpassung nach PCA

Statt der Verwendung des CIELAB-Farbraumes, um die Luminanz getrennt verarbeiten zu können, lässt sich auch eine Trennung mit Hilfe der Hauptkomponenten der Farbwerte des Quellbildes realisieren. Hierzu werden die Eigenvektoren der Kovarianzmatrix des Quellbildes als Zeilen einer Transformationsmatrix in einen neuen Farbraum verwendet. Der Eigenvektor mit dem größten Eigenwert spiegelt hierbei den Luminanzkanal wieder. Da in nicht künstlichen Bildern die RGB-Kanäle stark korreliert sind [JLW⁺12], ist in der Regel der dadurch berechnete Luminanzkanal ähnlich dem im CIELAB-Farbraum.

Wesentliche Unterschiede zur Verwendung des CIELAB-Farbraumes werden erst bemerkbar, wenn die RGB-Kanäle des Quellbildes nur schwach oder gar nicht korreliert sind.

Abbildung 3.3 zeigt einen solchen Fall. Eine rote Hauptkomponente dominiert sowohl das Quellbild als auch das Zielbild. Diese Komponente definiert auch die Glanzeffekte im Quellbild. Durch die Anpassung mit der LumHist(Lab)-Methode werden diese lediglich skaliert, da sie nicht als Luminanz betrachtet werden, und können so-

mit nur schlecht entfernt werden. Die Anpassung über die LumHist(PCA)-Methode betrachtet die rote Komponente als Luminanzkanal und entfernt durch Histogrammanpassung die Glanzeffekte.

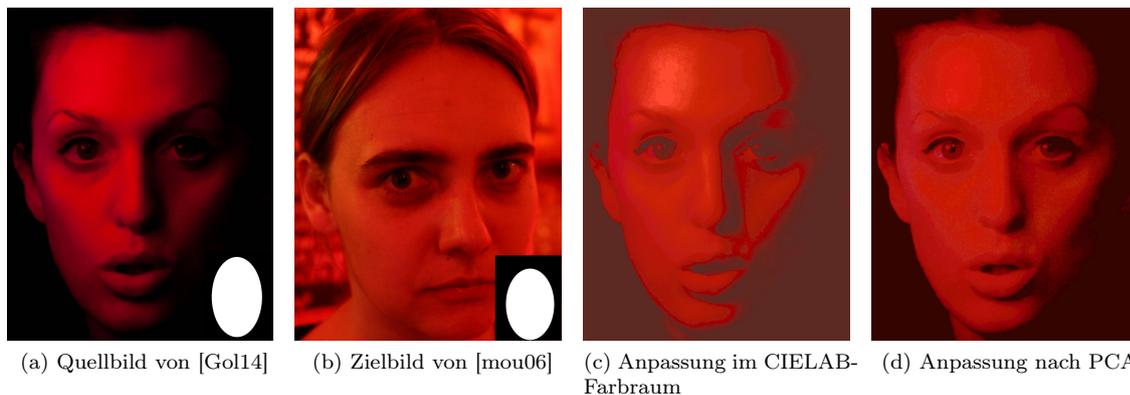


Abbildung 3.3: Vergleich zwischen LumHist(PCA) und LumHist(Lab)

Die jeweils verwendete Maske ist im unteren rechten Teil abgebildet. In beiden Fällen wurde im Chrominanzraum eine Ausreißerentfernung nach 3.3 vorgenommen.

3.3 Entfernen von Ausreißern

Klassische Verfahren zum Erkennen und Entfernen von Ausreißern benötigen entweder eine 1:1 Zuweisung der Werte oder sind nur auf einzelne Datenreihen anwendbar. Bei der Anpassung zweier Datenreihen, wie beispielsweise den Farbwerten zweier Bilder, ist beides nicht gegeben. Eine mögliche Definition von Ausreißern wären Farbwerte im Quell-/Zielbild, in deren RGB-Umgebung kein oder nur ein geringeres Vorkommen der Farbwerte im Ziel-/Quellbild zu beobachten ist. Durch diese Definition werden beispielsweise Heavy-Tails entfernt, welche nur in einem Bild vorhanden sind und bei Farbanpassungen, wie zum Beispiel in Abschnitten 3.2.2 und 3.2.1 beschrieben, zu falsch geschätzten Varianzen und Mittelwerten zugrunde liegender Normalverteilungen führen können. Andere Arten von Farbwerten, die dadurch entfernt werden, sind Sensorfehler wie defekte Pixel oder durch die Szene bedingte Farben wie Reflexionen, die nur in einem der Bilder vorkommen. Bei der Anpassung von Gesichtsbildern können diese auch durch verschiedene Augenfarben, Leberflecke oder Narben bedingt sein.

Um Ausreißer nach diesem Kriterium zu entfernen, bietet sich folgende Möglichkeit

an:

Zuerst wird ein normalisiertes 3D-Histogramm (H) des Quellbildes und (G) des Zielbildes erstellt. In dieser Arbeit wurden die Histogramme mit 64, die gleiche Anzahl an möglichen Farbwerten umfassenden, Klassen pro Kanal erstellt. Dies entspricht bei einem im sRGB-Farbraum gespeicherten Bild mit ganzzahligen Farbwerten von 0 bis 255 pro Kanal 4 und pro Histogrammeintrag 64 verschiedenen Farbwerten. Für jeden möglichen Farbwert x wird nun eine Distanz d_h für das Quellbild nach folgendem, im Rahmen dieser Arbeit entworfenem, Algorithmus berechnet:

Require: $Dist$ Farbwerteliste, geordnet nach Distanz zu x

```

remaining  $\leftarrow H(x)$ 
d  $\leftarrow 0$ 
for  $i = 0; i < |Dist|; i++$  do
  if  $remaining \geq G(Dist(i))$  then
     $d \leftarrow d + ||Dist(i)|| \cdot G(Dist(i))/H(x)$ 
  else
     $d \leftarrow d + ||Dist(i)|| \cdot remaining$ 
  end if
   $remaining \leftarrow remaining - G(Dist(i))$ 
  if  $remaining \leq 0$  then
    return  $d$ 
  end if
end for

```

Entsprechend werden auch die Distanzen für das Zielbild berechnet. Im folgenden Schritt werden alle Distanzen mit einem R-, G- oder B-Anteil von dem kleinst- bzw. größtmöglichen Wert (i.d.R. je nach Format 0 und 1 oder 0 und 255) auf 0 gesetzt und aus dem Quell- und Zielbild entfernt. Abschließend wird eine prozentuale Anzahl an Pixeln entfernt, deren entsprechende Farbwerte die größten Distanzen aufweisen. Wie viele Pixel entfernt werden sollten, ist von den jeweiligen Szenen und erwarteten Unterschieden abhängig. Sofern nicht anders angegeben, werden in dieser Arbeit 20% entfernt. Die Distanzen mit einem R-, G- oder B-Anteil von dem kleinst- bzw. größtmöglichen Wert werden bei einer Ausreißerentfernung im Chrominanzraum nicht entfernt.

Die Grundidee dieses Verfahrens zur Ausreißereliminierung ist es für jede einzelne Farbklasse im Quell- bzw. Zielbild zu überprüfen, ob deren relatives Vorkommen durch die Farbverteilung des Ziel- bzw. Quellbild begründet werden kann. Dazu wird für jede Farbklasse ihre relative Häufigkeit im einen Bild, mit der relativen Häufigkeit der entsprechenden Farbklasse und deren Farbraumumgebung im anderen Bild verglichen. Treten Farbwerte einer Farbklasse in einem Bild wesentlich häufiger auf, als diese Farben und im verwendeten Farbraum nahe liegende Farben im anderen Bild, führt dies zu einer hohen Distanz und somit zu einer größeren Wahrscheinlichkeit, als Ausreißer entfernt zu werden.

Eine andere Möglichkeit zur Entfernung von Ausreißern ist der RANSAC-Algorithmus [FB81]. Hierbei werden zufällig Werte als Ausreißer definiert. Es wird eine Transformation berechnet und überprüft, wie gut die transformierten Ursprungsdaten die Zieldaten beschreiben. Dieses Verfahren wird mehrmals wiederholt und das beste Ergebnis verwendet. Als Gütemaß kann beispielsweise die Differenz des Histogramms des transformierten Bildes zu dem Histogramm des Zielbildes verwendet werden.

3.4 Iterative Anpassung

Durch die beschriebenen Verfahren zur Farbanpassung werden die Mittelwerte, Varianzen und Kovarianzen der Farben des Quellbildes an die des Zielbildes angepasst. Eine weitere Anpassung mit den Werten des transformierten Quellbildes führt demnach zu keiner weiteren Änderung. Anders ist dies jedoch, wenn durch das Entfernen von Ausreißern unterschiedliche Pixel als Grundlage zur Anpassung dienen. Bei einer iterativen Anpassung ist jedoch darauf zu achten, dass durch die Anpassung Werte kleiner als 0 bzw. größer als 255 entstehen können. Diese können zwar nicht dargestellt werden, sind für die Entfernung der Ausreißer jedoch relevant. Die Entfernung der Pixel mit Intensitätswerten von 0 bzw. 255 auf einem beliebigen Kanal sollte demnach nur vor der ersten Iteration durchgeführt werden, da durch die Anpassung der zulässige Wertebereich vergrößert wird. Um bei der Verwendung der MK-Methode weiterhin einen minimalen Aufwand nach (3.3) für die Anpassung zu erhalten, werden die transformierten Werte nur zur Bestimmung der Ausreißer-Pixel verwendet. Die eigentliche Abschätzung findet weiterhin mit den Werten des Quellbildes statt. Abbildung 3.4 zeigt eine iterative Anpassung mit der MK-Methode und einer Ausreißerentfernung von sowohl 5% als auch 20% aller Pixel. Während

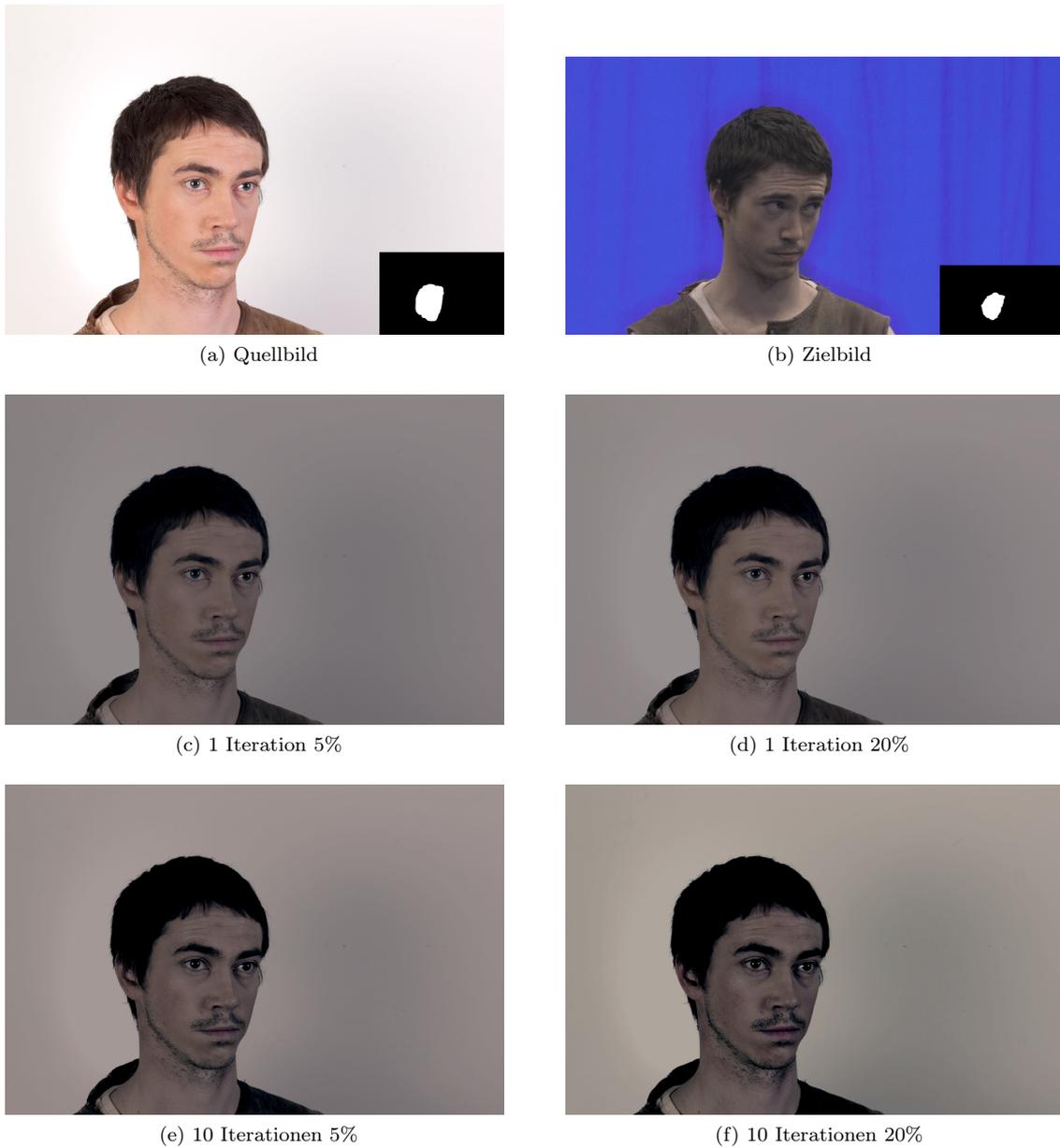


Abbildung 3.4: Iterative Anpassung
Die jeweils verwendete Maske ist im unteren rechten Teil des Bildes abgebildet.

nach der ersten Iteration beide Anpassungen noch zu nahezu identischen Ergebnisse führen, sind nach 10 Iterationen starke Unterschiede erkennbar. Diese sind auf das verschieden starke Entfernen der Heavy-Tails zurückzuführen, welcher im Quellbild in die Richtung der dunklen Farbwerte und im Zielbild in die der hellen Farbwerte geht. Durch weitere Iterationen entstanden keine nennenswerten Veränderungen.

3.5 Vergleich

Die PCA-Methode und MK-Methode führen unter bestimmten Bedingungen zu identischen Transformationen. Dies ist beispielsweise der Fall, wenn die Farbkanäle eines Bildes vollständig dekorreliert sind und identische Varianzen besitzen oder in beiden Bildern die Farbkanäle keinerlei Korrelation aufweisen. Diese Bedingungen erscheinen auf den ersten Blick ziemlich unrealistisch, treten jedoch nach einer vorherigen Transformation in den CIELAB-Farbraum, wie sie in der LumHist(Lab)-Methode verwendet wird, bedingt auf.

In den meisten Fällen verschiebt die MK-Methode das Zentrum der geschätzten Normalverteilung weniger als die PCA-Methode. Dies liegt daran, dass Verschiebungen an dieser Stelle einen hohen Aufwand nach (3.3) zur Folge haben. An von dem Mittelpunkt der Normalverteilung weiter entfernten Stellen befinden sich dafür vereinzelt stärkere Verschiebungen als sie bei der PCA-Methode entstehen. Auf Grund des seltenen Vorkommens dieser Farbwerte führen diese jedoch insgesamt nur zu einem geringen Aufwand.

Abbildung 3.5 zeigt an einem Beispiel in 2D mögliche Unterschiede zwischen der PCA- und der MK-Methode. Abbildung 3.5a zeigt eine 2D-Normalverteilung der Farbwerte eines Quellbildes mit einer Varianz von $4 \cdot 10^{-2}$ auf der x-Achse, $3 \cdot 10^{-3}$ auf der y-Achse und einer Kovarianz von $1.5 \cdot 10^{-2}$. Abbildung 3.5b zeigt die Normalverteilung des Zielbildes. Diese ist unkorreliert und hat als Varianz in der x-Dimension $2 \cdot 10^{-3}$ und in der y-Dimension $5 \cdot 10^{-3}$. Abbildung 3.6 zeigt die Verschiebungen der einzelnen Farbwerte durch die Transformation nach der MK-Methode in grün und nach der PCA-Methode in rot. Die im Hintergrund angedeutete Normalverteilung bezieht sich auf das Quellbild. Zu erkennen ist, dass die Verschiebung um den Mittelpunkt der Normalverteilung bei Verwendung der PCA-Methode größer ist als bei der MK-Methode. Diese Differenz nimmt jedoch mit dem Abstand zum Mittelpunkt

ab.

Der Aufwand nach (3.3) nimmt bei Verwendung der PCA-Methode wie in Abbildung 3.5c sichtbar anfangs stark zu und nur langsam ab. Die diesen Aufwand minimierende MK-Methode hat hingegen zwei weit vom Zentrum entfernte Regionen mit stärkerem maximalem Aufwand als er in der PCA-Methode existiert. Abbildung 3.5e visualisiert die Aufwandsdifferenzen. Diese sind definiert als der Aufwand nach (3.3) durch die MK-Methode abzüglich dem durch die PCA-Methode. Trotz der oben genannten Regionen ist der maximale Wert der Aufwandsdifferenzen wesentlich kleiner als der Betrag des kleinsten negativen Wertes. Die beschriebenen Effekte konnten auch bei den meisten anderen Werten für Varianzen und Korrelationen der Normalverteilungen des Quell- und Zielbildes beobachtet werden. Eine Ausnahme sind natürlich die Fälle, in denen die PCA-Methode und die MK-Methode identische Transformationen berechnen.

Deutliche Unterschiede zwischen der MK- bzw. PCA-Methode und der LumHist(Lab)-Methode bzw. LumHist(PCA)-Methode sind bei verschiedenen Luminanzverteilungen im Quell- und im Zielbild erkennbar. Diese können sich, je nach den zugrunde liegenden Verteilungsfunktionen, durch das Abschwächen bzw. Verstärken von Schatten und Glanzeffekten äußern. Abbildung 3.7 zeigt eine Farbanpassung sowohl mit der MK-Methode als auch mit der LumHist(PCA)-Methode. In beiden Fällen wurden keine Ausreißer entfernt und nur ein Chrominanzband verwendet. Das Quellbild weist im Luminanzkanal einen starken Heavy-Tail in Richtung der dunklen Werte auf, welcher im Zielbild nicht vorhanden ist. Diese asymmetrische Verteilung führt zu einer erhöhten Varianz im Quellbild und somit zu einer starken Stauchung bei der Anpassung an das Zielbild. Diese starke Stauchung wäre, um eine optimale Anpassung der Verteilungen zu erreichen, jedoch nur für Farbwerte kleiner dem unteren Quartil notwendig, siehe Abb. 3.7. Durch eine globale Skalierung, wie sie die PCA-/MK-Methode vornimmt, wird auch die Verteilung der Luminanz der helleren Farbwerte stark gestaucht. Im Bild ist dies an der Abschwächung der Glanzeffekte zu erkennen. Bei einer Histogrammanpassung werden diese jedoch, da sie im Zielbild verstärkt auftreten, durch die Transformation verstärkt.

Die LumHist-Methoden haben den Vorteil, dass eine Entfernung der Ausreißer auch bandweise im Chrominanzbereich möglich ist. Abbildung 3.8 zeigt eine Transforma-

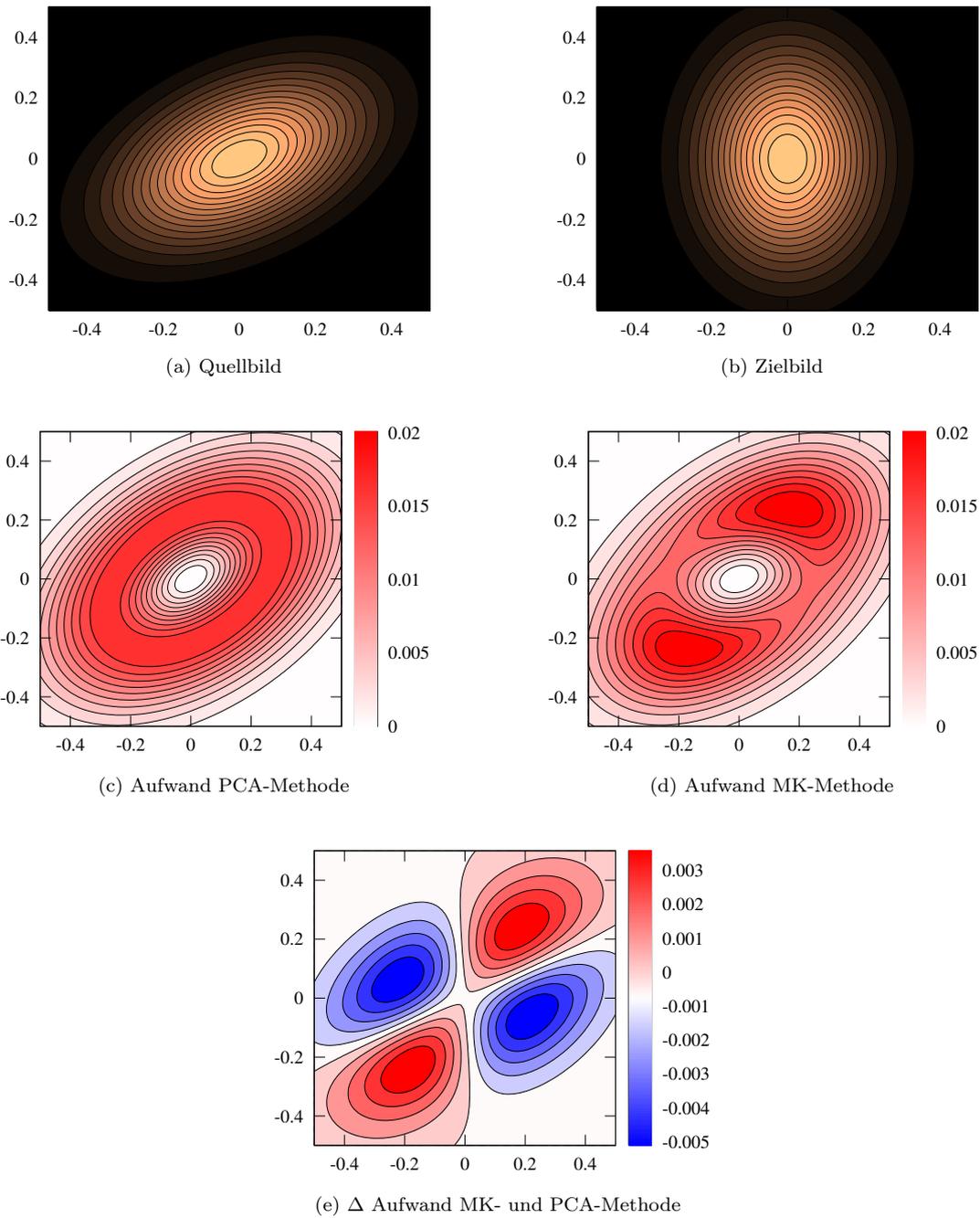
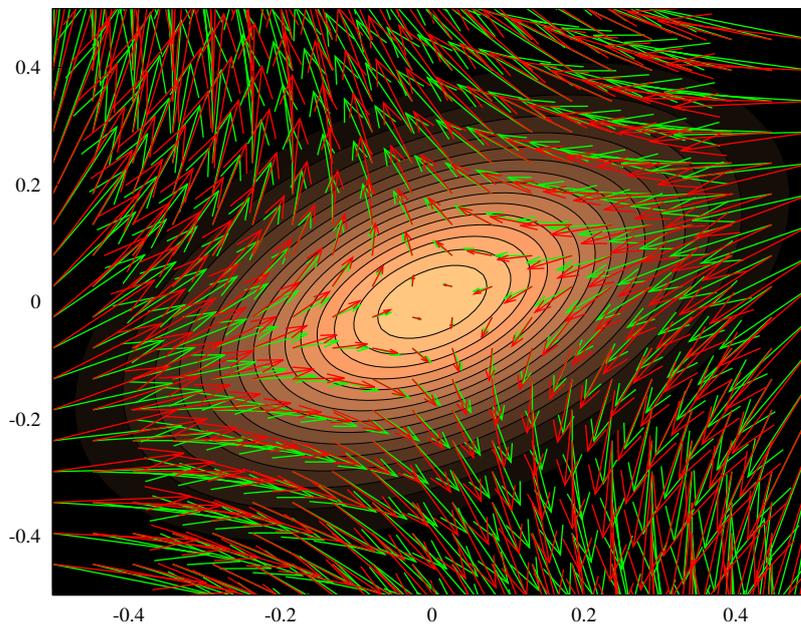


Abbildung 3.5: Vergleich zwischen MK-Methode und PCA-Methode (Aufwand)



Die Farbverschiebungen der MK-Methode wurden in grün und die der PCA-Methode in rot dargestellt.

Abbildung 3.6: Farbverschiebung bei der MK-Methode und PCA-Methode

tion mit der LumHist(Lab)-Methode sowohl mit als auch ohne Entfernung der Ausreißer. Die Ausreißer wurden hierbei, wie in Abschnitt 3.3 beschrieben, pro Band in den ab-Kanälen entfernt. Für die Luminanzanpassung wurden keine Werte entfernt. Farbliche Unterschiede sind hier in dem dunklen und mittleren Band zu erkennen. Im dunklen Band äußern sich diese in einem bläustichigen Farbton um die Augen bzw. in einer blauen statt grün-blauen Iris, sofern keine Ausreißerentfernung durchgeführt wurde. Das mittlere Band, welches primär schattige Hautregionen beinhaltet, weist unter Einbezug der Ausreißer einen gelbstichigen Farbton auf, der bei vorheriger Entfernung der Ausreißer nicht vorhanden ist.

3.6 Fazit

Ein Problem bei der Bewertung der Farbanpassungen verschiedener Szenen ist, dass keine eindeutige objektive Möglichkeit zum Bewerten der Verfahren existiert. Zur Bewertung von Kamerakalibrierungs-Methoden wird meist der MSE von korrespondierenden Pixeln bzw. Farbwerten verwendet. Da bei der Anpassung verschiedener



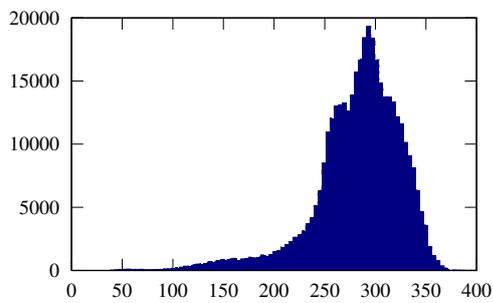
(a) Quellbild



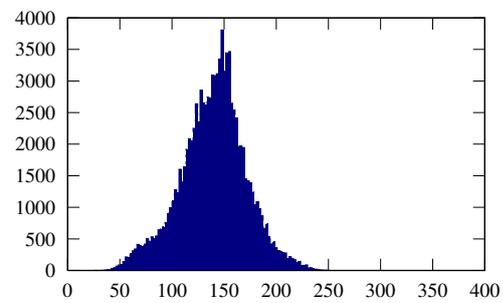
(b) Zielbild



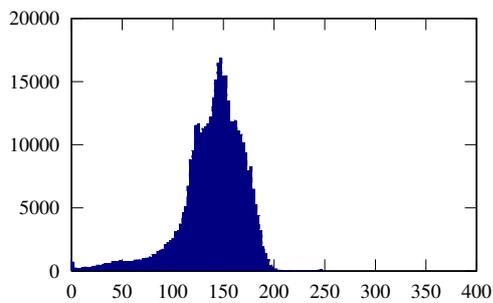
(c) Ergebnis der Farbanpassung (links Lum-Hist(PCA), rechts MK)



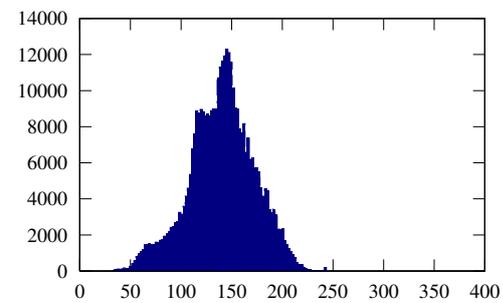
(d) Luminanzhistogramm Quellbild



(e) Luminanzhistogramm Zielbild



(f) Luminanzhistogramm nach Anpassung mit MK-Methode



(g) Luminanzhistogramm nach Anpassung mit LumHist(PCA)-Methode

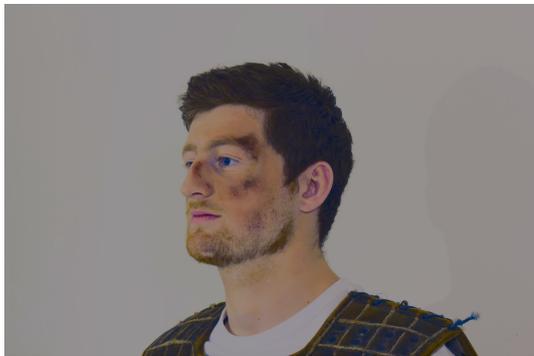
Abbildung 3.7: Vergleich zwischen MK-Methode und LumHist(PCA)-Methode
Die jeweils verwendete Maske ist im unteren rechten Teil abgebildet.



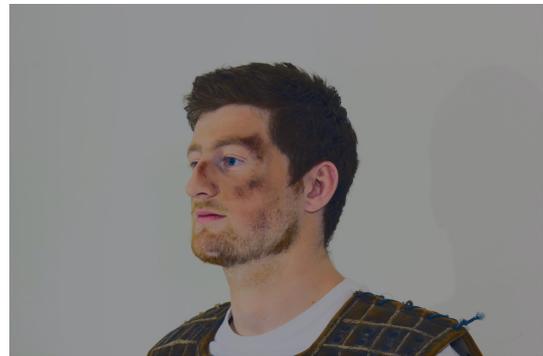
(a) Quellbild



(b) Zielbild



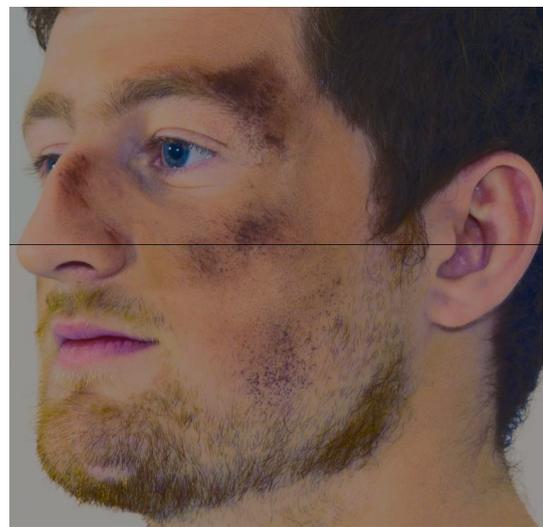
(c) Anpassung ohne Ausreißerentfernung



(d) Anpassung mit Ausreißerentfernung



(e) Unterteilung Chrominanzbänder



(f) Direkter Vergleich (c) zu (d)

Abbildung 3.8: Auswirkung der Ausreißerentfernung
 Abbildung (f) zeigt in der oberen Hälfte der Ergebnis mit Entfernung der Ausreißer und in der unteren ohne. Die jeweils verwendete Maske ist im unteren rechten Teil abgebildet.

Szenen diese Korrespondenz jedoch nicht eindeutig hergestellt werden kann, entfällt diese Option. Ob eine auf deskriptiver Statistik basierende Methode eine sinnvolle Bewertung der verschiedenen Transformationen darstellt, ist fraglich, da die Methoden auf der Anpassung der hierzu gängigsten Vergleichswerte wie dem arithmetischen Mittel oder der Varianz beruhen, bzw. wie die Histogrammanpassung diese unmittelbar zur Folge haben. In Publikationen werden deshalb meist nur die vorgestellten Verfahren anhand von Beispielen präsentiert und mit anderen Verfahren verglichen.

Die ansprechendsten Ergebnisse konnten mit den LumHist-Methoden erzielt werden. Dies liegt unter anderem daran, dass die Luminanzanpassung frei von Annahmen über zugrunde liegende Verteilungsfunktionen ist, während bei der Chrominanz lediglich mit einer affinen Transformation die Varianzen, die Kovarianz und der Mittelwert pro Band angepasst werden.

Die beiden vorgestellten LumHist-Methoden führen zwar meist zu ähnlichen Ergebnissen. Sie unterscheiden sich jedoch stark, wenn die Hauptkomponenten der Bilder nicht der Luminanz im Sinne vom CIELAB-Farbraum entsprechen und die Luminanz verschiedene Verteilungsfunktionen vorweist. Dies ist meist bei Bildern unter künstlicher, farbiger Beleuchtung der Fall. Hierbei empfiehlt es sich, die Luminanz über die Hauptkomponenten zu bestimmen, da bei einer Anpassung über den CIELAB-Farbraum diese nur durch eine Skalierung und Addition angepasst werden kann.

Kapitel 4

Anpassung der Unschärfe

Die Schärfe einer Aufnahme ist sowohl von den Kameraeigenschaften und Kameraeinstellungen als auch von der Szene abhängig. Unter den von der Szene abhängigen Faktoren sind hierbei die Bewegung der Kamera bzw. abgebildeten Objekte und relative Position der Objekte zu der Kamera zu verstehen. Daher ist auch hierbei die Auswahl einer repräsentativen Region im Quellbild und im Zielbild von Vorteil. Als repräsentative Region wird das kleinste Rechteck, welches vollständig eine per Hand gezeichneten binäre Maske beinhaltet, gewählt. Ein Rechteck ist erforderlich, da zur Abschätzung des Unschärfeverhältnisses eine 2D-Fouriertransformation beider Bilder verwendet wird. Ein für das Verfahren wichtiger Punkt bei der Auswahl der repräsentativen Regionen ist, dass die Rechtecke denselben Bereich auf den betrachteten Objekten aus dem selben Blickwinkel beschreiben und folglich auch dasselbe Seitenverhältnis aufweisen.

Als vereinfachende Annahme wird im Folgenden davon ausgegangen, dass die einzelnen Sensoren der Kamera quadratisch sind. Sollte dies nicht der Fall sein, muss das Verfahren lediglich um Skalierungen in Richtung der Breite oder Höhe der Bilder erweitert werden.

Die Grundidee des Verfahrens besteht darin, die Unterschiede der Amplituden der Frequenzen zwischen Quellbild und Zielbild durch Multiplikation mit einer Gauß'schen Glockenkurve zu beschreiben. Dies entspricht im Ortsbereich der Anwendung eines Gauß-Filters [STI14].

Zuerst werden beide Bildabschnitte von dem sRGB-Raum in ein grau-stufiges Bild umgewandelt. Dies geschieht durch eine gewichtete Addition der einzelnen Farb-

kanäle. Als Gewicht für Rot wird 0.299, für Grün 0.587 und für Blau 0.144 verwendet. Dies entspricht auch der Luminanzdefinition des YUV-Farbraums. Da die diskrete Fouriertransformation das zu transformierende Signal als periodisch betrachtet [MH04], führen große Unterschiede zwischen gegenüberliegenden Bildrändern zu großen Sprüngen im Signal und somit zu hohen Frequenzen. Um dies zu Vermeiden, werden die Bilder so mit ihrer horizontal, vertikal und sowohl horizontal als auch vertikal gespiegelten Version erweitert, dass alle gegenüberliegenden Bildränder identisch sind. Als nächstes werden die Bilder in den Fourierbereich transformiert und alle redundanten Teile und die Gleichanteile der Spektren entfernt. Die Spektren werden dann durch Entfernen der höchsten Frequenzen so zugeschnitten, dass die Spektren gleich groß sind. Die Einträge der Koeffizienten (x,y) im Spektrum des Quellbildes und des Zielbildes spiegeln dieselbe Frequenz wieder, da die Frequenzdifferenzen zwischen zwei benachbarten Punkten im diskreten Spektrum nur von der beobachteten Fläche abhängen [MH04].

Die horizontale und vertikalen Frequenzen des Spektrums werden im Folgenden einheitslos als in den Bereich von 0 bis 1 normalisierte Frequenzen bezeichnet.

Im weiteren Teil dieser Arbeit werden die Ortsfrequenzspektren wie eindimensionale Spektren mit unterschiedlichen Frequenzabständen und möglichen mehrfachen Werten pro Frequenz behandelt. Die Umrechnung einer Ortsfrequenz in eine 1D-Frequenz erfolgt nach:

$$F1D(x, y) = \sqrt{(x^2 + y^2)} \quad (4.1)$$

Zur Notwendigkeit eines Mittelwertfilters:

Rotationen der Objekte oder der Kamera im 3D-Raum, leichte Skalierungen und Ungenauigkeiten der Maske führen zu Verschiebungen der Frequenzen und sind meist nicht vermeidbar. Um dennoch die Frequenzen des Quellbildes mit denen des Zielbildes direkt vergleichen zu können, wird ein Mittelwertfilter zur Aufteilung der Amplituden über mehrere Frequenzen verwendet. Die Reichweite des Mittelwertfilters hängt dabei von den normalisierten Frequenzen und nicht der Anzahl der Messwerte ab. In den hier verwendeten Beispieldaten, welche teilweise starke Unterschiede bezüglich des Aufnahmewinkels zu dem aufgenommenen Objekt aufweisen, hat sich ein Schwellwert von 0.01 als sinnvoll erwiesen.

Zur Notwendigkeit einer Frequenzauswahl:

Hohe Frequenzen sind teilweise nicht durch die Szene bedingt, sondern entstehen durch die Aufnahme- oder Speichertechnik der Kamera. Das JPG-Format zerteilt zur Speicherung die Bilder in einzelne 8x8 Pixel-Blöcke und komprimiert diese verlustbehaftet einzeln. Durch die blockweise Komprimierung können an den Übergängen zwischen den Blöcken hohe Frequenzen entstehen, während durch die verlustbehaftete Kompression hohe durch die Szene bedingte Frequenzen abgeschwächt werden können. Ein weiterer Störfaktor ist das thermische Rauschen der Sensoren, welches ein konstantes Leistungsdichtespektrum besitzt. Um durch diese Effekte keine falsche Schätzung für den Tiefpass zu erhalten, werden alle Frequenzen oberhalb eines gewissen Schwellwertes entfernt.

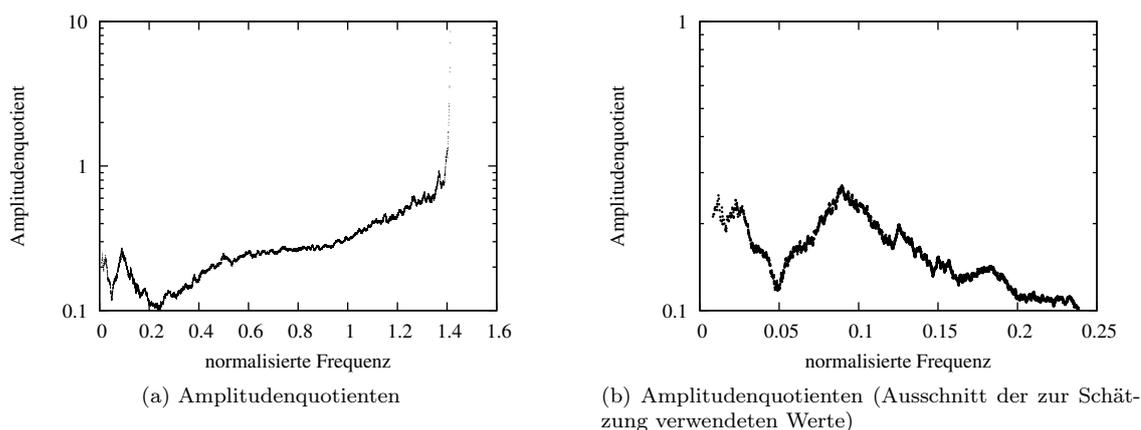


Abbildung 4.1: Geglättete Quotienten der Amplituden

Die Amplituden wurden mit einem Mittelwertfilter mit einer Reichweite von 0.01 gefiltert. Dargestellt wurde das Verhältnis der Quotienten Amplitude Zielbild / Amplitude Quellbild über die normalisierten Frequenzen des Quellbildes. Als Ziel- und Quellbild wurden die Bilder in 4.2a und 4.2b verwendet.

Dieser Schwellwert sollte so gewählt werden, dass die beschriebenen Störfaktoren nur noch Einfluss auf die Frequenzen oberhalb dieses Wertes haben. Da im Rahmen dieser Studienarbeit durch das JPG-Format komprimierte Quellbilder und Zielbilder im PNG-Format untersucht wurden, wurde die Bestimmung des Schwellwertes an diese Gegebenheiten angepasst.

Zur automatischen Bestimmung des Schwellwertes wurden die Quotienten aus den Amplituden des Fouriertransformierten Zielbildes und denen des Quellbildes verwen-

det. Als Schwellwert wurde das Minimum aus der Frequenz mit dem kleinsten Quotienten und einem absolutem Schwellwert, der dem unteren Quartil entspricht, gewählt. Die Verwendung des kleinsten Quotienten als Schwellwert kann durch die Annahme begründet werden, dass die Frequenzunterschiede bis zu einem bestimmten Wert durch die abzuschätzende, physikalisch bedingte Unschärfe entstanden sind und somit durch Multiplikation mit einer Gauß'sche Glockenkurve beschrieben werden können. Ab diesem Schwellwert ist ein starken Frequenzabfall im Quellbild vorhanden. Dieser ist vermutlich durch das Kompressionsverfahren des JPG-Formats entstanden und führt dazu, dass die Frequenzen im PNG-Zielbild überwiegen, vgl. Abb. 4.1. Dieser Effekt wurde nur empirisch in den gegebenen Beispieldaten beobachtet und nicht nachgewiesen.

Die somit übrig gebliebenen Ortsfrequenzen (F_{Est}) werden nun verwendet, um die Frequenzunterschiede mit Hilfe einer Gauß'schen Glockenkurve zu approximieren. Minimiert wird hierbei die Funktion

$$\min_{\alpha, \sigma} \sum_{f \in F_{Est}} \left(\alpha \cdot FT_Q(f) \exp\left(\frac{-F1D(f)^2}{2 \cdot \sigma^2}\right) - FT_Z(f) \right)^2 \quad (4.2)$$

mit FT_Q als Spektrum des Quellbildes und FT_Z als Spektrum des Zielbildes. Der Faktor α dient der Skalierung des Spektrums und soll vermeiden, dass der Gauß-Filter durch mögliche Varianzunterschiede, beispielsweise durch unterschiedliche ISO-Werte oder Sensorempfindlichkeiten, zwischen Quell- und Zielbild beeinflusst wird. Die Funktion (4.2) wurde mit Hilfe des Gauß-Newton-Verfahrens minimiert. Als Startwert für α wurde 1 und für σ $(2 \cdot \pi)^{-1}$ verwendet, was einer Standardabweichung des Gauß-Filters im Ortsbereich von einem Pixel entspricht. Zur Bestimmung der Standardabweichung des Gauß-Filters wurde nur σ verwendet. Die Umrechnung der Standardabweichung im Frequenzbereich in den Ortsbereich ist im diskreten 1D-Fall nach

$$\sigma_{Ort} = \frac{N}{2\pi\sigma} \quad (4.3)$$

mit N als Anzahl der Abtastwerte möglich. Um die Normalisierung der Frequenzen rückgängig zu machen, wird dieser Term mit $(0.5N)^{-1}$ multipliziert. Da im vorliegenden Fall die 1D-Frequenzen jedoch aus 2D-Ortsfrequenzen berechnet wurden, entspricht N nicht der Anzahl an Abtastwerten, sondern der Höhe bzw. Breite in Pixeln. Die somit bestimmte Standardabweichung für den Ortsbereich ist nun in

Pixeln des kleineren Bildes angegeben. Die Umrechnung in Pixel des größeren Bildes ist mit einer Multiplikation mit der Breite/Höhe der größeren Maske durch die Breite/Höhe der kleineren Maske möglich.

Als Formel zur Berechnung der Standardabweichung des Gauß-Filters in Pixeln ergibt sich somit

$$\sigma_{Ort} = \frac{1}{\pi\sigma} \quad (4.4)$$

für das kleinere Bild und

$$\sigma_{Ort} = \frac{m_{max}}{\pi\sigma m_{min}} \quad (4.5)$$

mit m_{max} als Breite der größeren Maske und m_{min} als Breite der kleineren Maske für das größere Bild.

Abbildung 4.2 zeigt die Ergebnisse einer Anpassung der Schärfe.

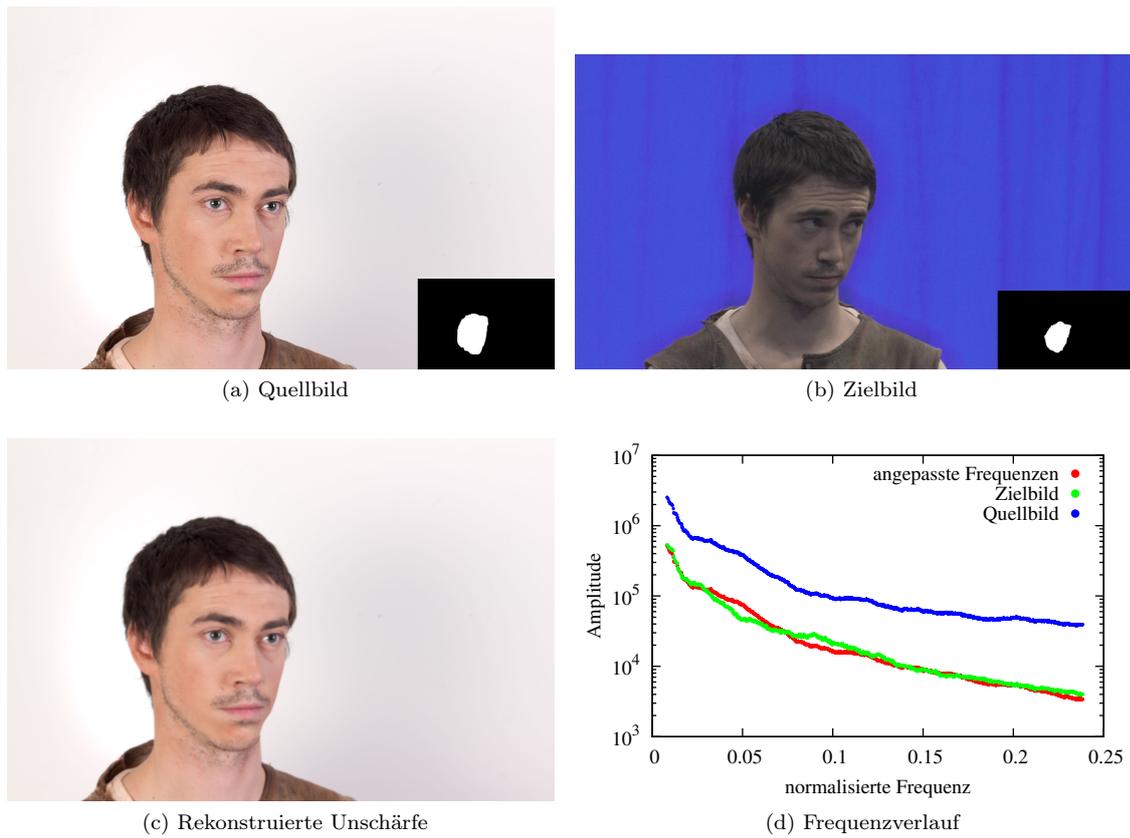


Abbildung 4.2: Anpassung der Schärfe

Die jeweils verwendete Maske ist im unteren rechten Teil abgebildet. Als Standardabweichung des Gauß-Filters des Quellbildes wurde der Wert von 2.81 Pixeln berechnet. Abbildung (d) zeigt die Amplituden der normierten Frequenzen der Bilder und die geschätzte Anpassung im Frequenzbereich.

Kapitel 5

Modellierung des Farbrauschens

5.1 Farbvarianzbestimmung

Das Farbrauschen ist eine nur von der Kamera und ihren Einstellungen abhängige Eigenschaft und kann somit im Gegensatz zur Farbverteilung anhand des vollständigen Bildes bestimmt werden. Da die hier beschriebene Abschätzung kanalweise geschieht, empfiehlt es sich auf den unverarbeiteten Rohdaten der Kamera zu arbeiten, um durch das Demosaicking oder andere interne Verarbeitungsschritte entstandene Korrelationen zu vermeiden. Eine auf den RGB-Kanälen basierende Abschätzung ist unter Beachtung der durch Demosaicking entstandenen Effekte auch üblich [LFSK06]. Alle im Folgenden genannten Werte beziehen sich auf einen Wertebereich von 0 bis 255.

Wie bereits in 2.3 beschrieben, ist die Varianz des Rauschens vom jeweiligen Intensitätswert abhängig und kann durch eine Normalverteilung, deren Standardabweichung abhängig von diesem ist, dargestellt werden. Da jedoch für jeden einzelnen Pixel nur der gemessene verrauschte Farbwert bekannt ist, muss abgeschätzt werden, welche Pixel denselben Farbwert darstellen sollten. Denkbar ist eine vorherige Messung mit einer Szene, in der die Farbwerte bekannt sind. Hierbei wäre jedoch zusätzlich zu beachten, dass das Rauschen sowohl von der ISO-Zahl als auch der Temperatur des Sensors und weiteren Faktoren abhängig ist. Dadurch wäre eine umfangreiche Kalibrierung notwendig. Um dies zu umgehen, wird in der Praxis [LFSK06] eine Segmentierung des Bildes in homogene Flächen vorgenommen.

Das hier verwendete Segmentierungsverfahren verfolgt den Top-Down-Ansatz. Das

Bild wird zu Anfang in 4 gleich große Rechtecke geteilt. Jedes Rechteck wird nun wieder in gleich große Rechtecke unterteilt. Dies wird für jedes neu entstandene Rechteck so lange wiederholt, bis die Varianz der Intensitätswerte innerhalb eines Rechtecks kleiner als ein vorgegebener Schwellwert ist oder durch die Unterteilung Rechtecke mit weniger als 50 Pixeln entstehen würden. Dieser Schwellwert sollte etwas über der erwarteten maximalen Varianz des Sensorrauschens liegen. Zusätzlich werden große Rechtecke immer geteilt, wenn die durch die Teilung entstehenden Rechtecke mehr als 1000 Pixel enthalten. Der Schwellwert von 1000 Pixeln soll vermeiden, dass große homogene Flächen zu nur einem Varianzmesswert führen. Alternativ wäre auch eine Gewichtung der Werte denkbar. Als Grundlage für die Approximation der Varianzen des Sensorrauschens werden segmentweise die Mittelwerte und die dazugehörige Varianzen der Intensitäten verwendet.

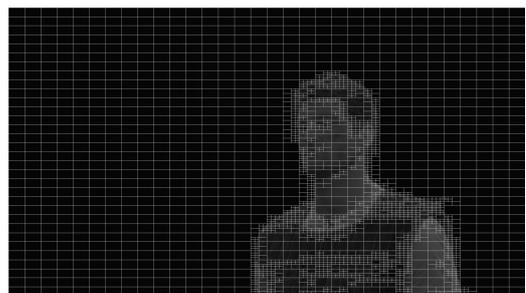
Ein großer Teil der Varianzen hat seinen Ursprung nicht nur in dem Sensorrauschen, sondern auch in der Szene. Um durch diese Messwerte das Sensorrauschen nicht zu hoch zu schätzen, werden sowohl über prozentuale als auch absolute Schwellwerte die gemessenen Varianzen aussortiert. Ein absoluter Schwellwert soll in erster Linie verhindern, dass Rechtecke, welche bei der Segmentierung nur auf Grund ihrer Größe nicht weiter unterteilt wurden, in die Schätzung mit eingehen. Da dieser Schwellwert oberhalb der maximalen Varianz des Sensorrauschens liegt, ist immer noch ein Großteil der gemessenen Varianzen durch die Szene bedingt. Um diese zu entfernen, werden die Messwerte in Abhängigkeit des Mittelwertes in 51 Klassen eingeteilt. Diese Aufteilung der Mittelwerte in Klassen, die je 5 mögliche Intensitätswerte umfassen, hat sich bei den gegebenen Daten für die Methode als geeignet erwiesen. Die i -te Klasse enthält alle Messdaten mit einem Mittelwert im Intervall $[5(i - 1), 5i)$. Die 51. Klasse enthält zusätzlich alle Messwerte mit einem Mittelwert von 255. Pro Klasse werden alle Varianzen entfernt, die über einem bestimmten Quantil der jeweiligen Klasse liegen. In den vorhandenen Gesichtsbildern wurde hierzu der Median verwendet. Ein pauschaler Schwellwert lässt sich nicht definieren, da dieser stark von der jeweiligen Szene abhängt. Die übriggebliebenen Messwerte dienen der in 5.2 beschriebenen Schätzung als Grundlage.

Abbildung 5.1 zeigt die einzelnen Schritte, nach denen die Messwerte berechnet und selektiert werden. Gewählt wurde hierzu der roten Farbkanal des in 5.1a dargestellten Bildes. Das Ergebnis der Segmentierung mit einem Schwellwert von 200 ist in

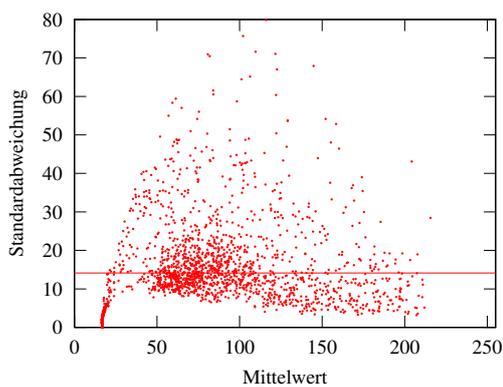
5.1b zu sehen. Abbildung 5.1c stellt die Standardabweichungen aller Segmente in Abhängigkeit ihrer Mittelwerte dar. Die horizontale Linie kennzeichnet den Schwellwert von 200. Es ist deutlich zu erkennen, dass ein Großteil der Varianzen unter dem Schwellwert liegt und dieser somit entsprechend den Anforderungen gewählt wurde. In Abbildung 5.1d sind die nach der klassenweisen Entfernung von je 50% der höchsten übriggebliebenen Varianzen dargestellt. Zu erkennen ist, dass die entfernten Werte größtenteils oberhalb von dichteren Punktwolken lagen. Unter der Annahme, dass viele homogene Flächen existieren und sich die Varianzen dieser bei identischem Mittelwert um einen Wert verteilen sollten, ist dies ein Indiz dafür, dass hauptsächlich durch die Szene bedingte Varianzen entfernt wurden.



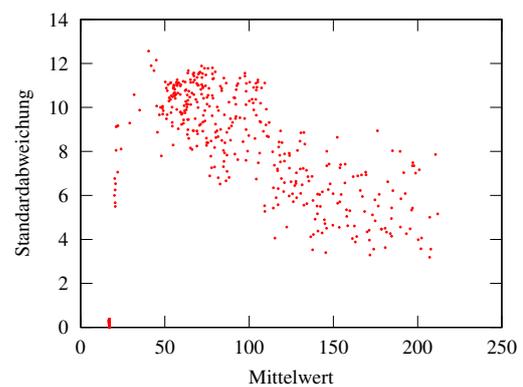
(a) Beispielbild



(b) Segmentierung (roter Kanal)



(c) Standardabweichungen der Segmente



(d) gefilterte Standardabweichungen der Segmente

Abbildung 5.1: Auswahl der Varianzdaten

5.2 Approximationsmodelle

Das Rauschen der Aufnahme verhält sich, mit einer Standardabweichung am Sensor von σ und der Intensität der zu messenden Farbe I , wie folgt:

$$CRF(I) - CRF(I + \mathcal{N}(0, \sigma^2)). \quad (5.1)$$

Dieses Verhalten weist auf eine starke Beziehung mit dem jeweiligen Differenzenquotienten hin. Bei Standardabweichungen des Rauschens am Sensor, die bei beliebigen Intensitäten nur zu geringen Änderungen der Ableitung der Camera Response Function (CRF) führen, verhält sich die Noise Level Function (NLF) folglich wie die Ableitung der CRF. Abbildung 5.2a zeigt 20 verschiedene CRFs über je 1024 verschiedene Intensitäten. Die Daten wurden zufällig aus einer Datenbank mit 201 gemessenen CRFs verschiedener digitaler und analoger Film- und Digitalkameras aus [GN03] entnommen. Abbildung 5.2b stellt die Vorwärtsdifferenzenquotienten dar. Eine numerische Simulation des Rauschens nach Formel (5.1) hat jedoch ergeben, dass aufgrund der teilweise starken Steigung der CRFs deren Ableitung sich nur bedingt wie die entsprechende NLF verhält, vgl. Abb. 5.2c. Schon Sensorrauschen mit einer geringen Standardabweichung führt zu stark unterschiedlichen Funktionsverläufen zwischen dem Differenzenquotienten und der simulierten NLF. Die Form der meisten simulierten NLF und die der Differenzenquotienten legen eine Approximation per Polynom oder mittels exponentiell gedämpfter Funktion nahe. Diese beiden Ansätzen werden im Folgenden vorgestellt und verglichen.

Der polynomielle Ansatz wurde mit Polynomen unterschiedlicher Grade getestet. Diese konnten bei einem niedrigen Grad die NLF der Beispieldaten nur schlecht approximieren. Eine Erhöhung des Grades hatte zur Folge, dass die geschätzte Standardabweichung von Intensitätswerten, zur denen keine Messdaten existieren, unrealistische und teilweise nicht mögliche Werte annimmt. Abbildung 5.3a visualisiert diese Problematik. Die verwendeten Schwellwerte entsprechen denen aus Abschnitt 5.1.

Der zweite Ansatz verwendet eine exponentiell gedämpfte Funktion der Form

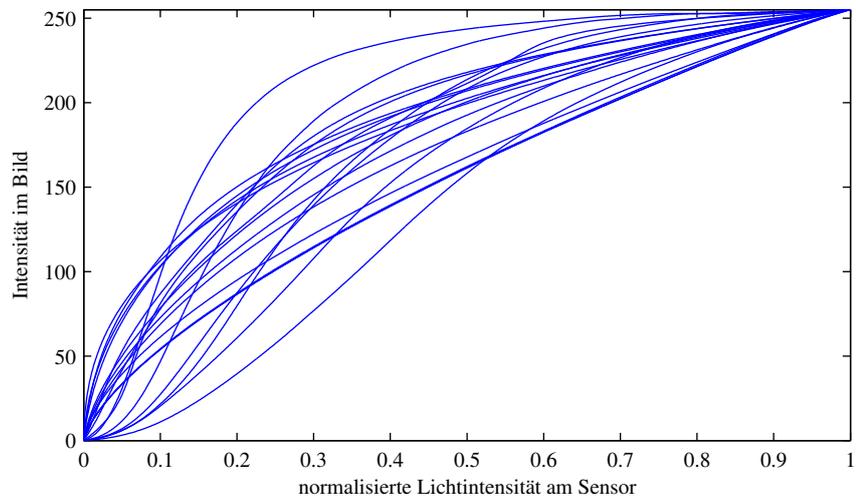
$$NLF(I) = a_0 I^{a_1} \exp(-a_2 I) \quad (5.2)$$

mit den zu schätzenden Parametern a_0 , a_1 und a_2 als Approximationsfunktion für die NLF. Ein Vorteil der exponentiell gedämpften Funktion gegenüber Polynomen ist, dass nur ein Wendepunkte (wp) existiert und dieser durch Ersetzen des Parameters a_1 durch $a_2 \cdot wp$ vorgegeben werden kann. Dies ist bei Polynomen, je nach Grad, nur bedingt möglich. Von besonderem Interesse kann das Festlegen eines Wendepunktes sein, wenn vor oder hinter diesem keine Messwerte existieren. Zur Berechnung der Parameter wurde das Gauß-Newton-Verfahren verwendet. Da hierbei wiederholt Konvergenzprobleme auftraten, wurde die Update-Regel um einen Schrittweitenparameter erweitert. Zur Bestimmung dieses Parameters und Sicherstellung der Konvergenz wurde die Armijo-Bedingung [Kan14] verwendet:

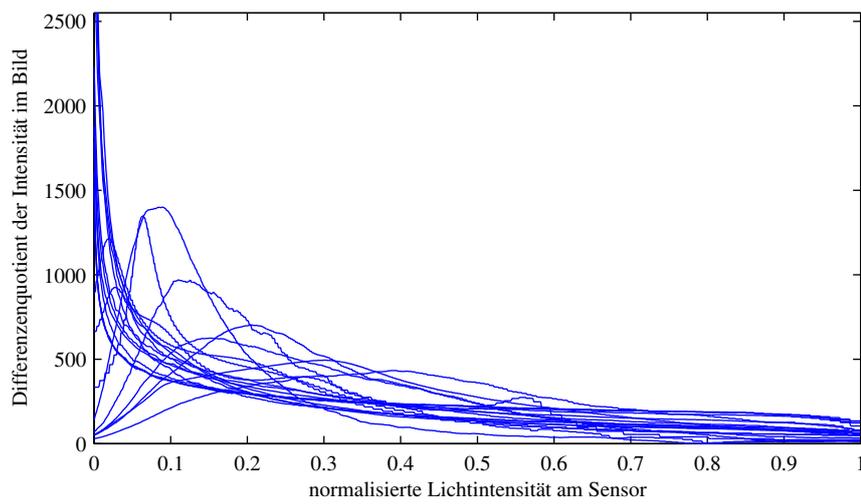
$$NLF(\alpha^k + t_k d^k) \leq NLF(\alpha^k) + t_k \sigma \nabla NLF(\alpha^k)^T d^k \quad (5.3)$$

mit $\sigma \in (0, 1)$ und $t_k \in \{\beta^i | i \in \mathbb{N}_0\}$ und $\beta \in (0, 1)$, wobei β und σ auf 0.5 gesetzt und der Schrittweitenparameter durch Erhöhen von i angepasst wurde.

Abbildung 5.3b vergleicht beide Ansätze mit den in Abschnitt 5.1 beschriebenen Messwerten. Zu erkennen ist, dass die exponentiell gedämpfte Funktion mit 3 Parametern die Messwerte besser approximieren kann als das entsprechende Polynom zweiten Grades. Polynome höheren Grades führen zwar zu geringeren mittleren quadratischen Differenzen, jedoch auch zu unrealistischen Werten an Stellen, für die keine Messwerte existieren.



(a) Camera Response Functions



(b) Differenzenquotient

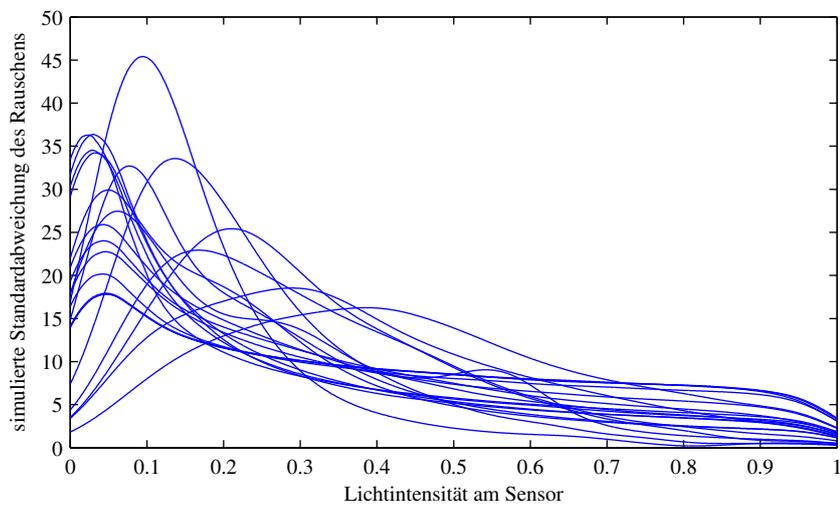
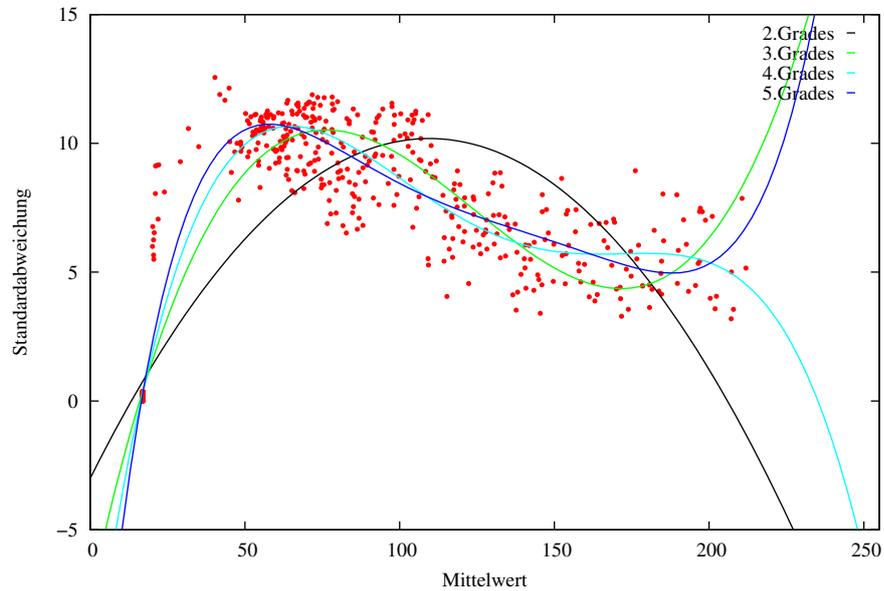
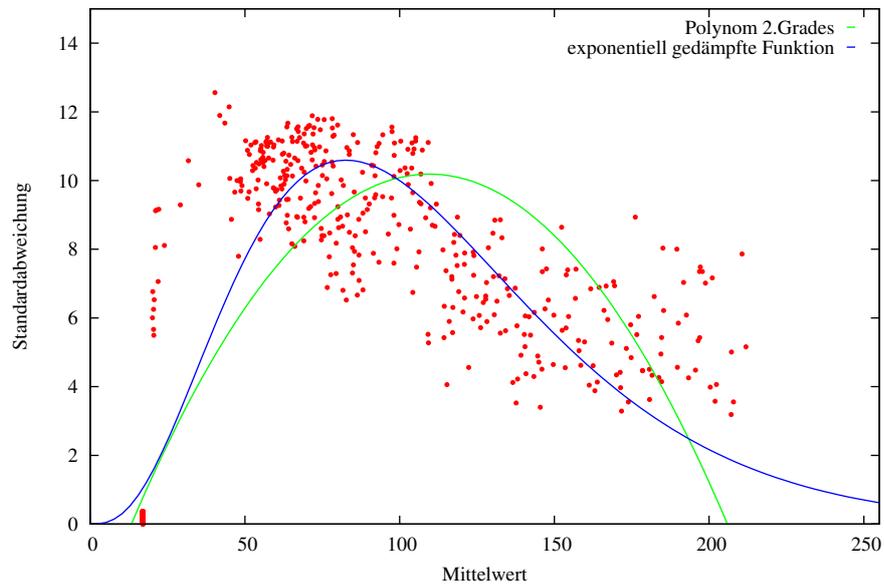
(c) simulierte NLF ($\sigma = 10$)

Abbildung 5.2: 20 gemessene Camera Response Functions



(a) Approximation mit verschiedenen Polynomen



(b) Approximation mit einer exponentiell gedämpften Funktion und einem Polynom zweiten Grades

Abbildung 5.3: Approximation der Standardabweichung über den Mittelwert
 Die Approximation durch eine exponentiell gedämpfte Funktion ist mit einem MSE von 2.31 möglich. Der polynomielle Ansatz führt für ein Polynom zweiten Grades zu einem MSE von 4.74. Polynome dritten, vierten und fünften Grades ergeben einen MSE von 1.38, 0.91 und 0.84.

Kapitel 6

Ergebnisse

Abbildung 6.1 zeigt die Resultate einer Farbanpassung durch die MK-Methode. Zur Auswahl der für die Transformation relevanten Pixel wurden unterschiedliche Masken verwendet. Die Maske zu der in Abbildung 6.1c berechneten Transformation beschreibt den ganzen Kopf und Teile des Oberkörpers, die zu Abbildung 6.1d nur das Gesicht. Das in Abbildung 6.1c gezeigte Ergebnis der Anpassung führt bei subjektiver Betrachtung zwar zu einem ansprechenderem Bild, jedoch stimmt weder die Helligkeit der Haare noch die des Gesichtes mit denen des Zielbildes überein. Dies liegt daran, dass im Quellbild die Haare dunkler als das Gesicht sind, im Zielbild jedoch das Gegenteil der Fall ist. Dies ist auf unterschiedliche Beleuchtungsverhältnisse zurückzuführen. Eine affine Transformation, wie sie die MK-Methode verwendet, kann solche Unterschiede nicht kompensieren. Die Anpassung, basierend auf den Farbwerten des Gesichtes, kann nur die Helligkeit des Gesichts anpassen, die der Haare bleibt weiterhin stark unterschiedlich zu der des Zielbildes. Die LumHist-Methoden führen in diesen Fällen auch zu keinen zufriedenstellenden Ergebnissen, wie aus Abbildung 6.1e ersichtlich ist. Durch eine getrennte Anpassung der Haare und Gesichter, z.B. mit der MK-Methode, können ansprechende Ergebnisse erzielt werden, vgl Abb. 6.1f.



(a) Quellbild



(b) Zielbild



(c) MK-Methode (Maske: Kopf und Oberkörper)



(d) MK-Methode (Maske: Gesicht)



(e) LumHist-Methode (Maske: Kopf und Oberkörper)



(f) MK-Methode (Haare getrennt angepasst)

Abbildung 6.1: Vergleich verschiedener Masken; getrennte Anpassung einzelner Regionen
Die verwendeten Masken sind am linken Rand des Quell- und Zielbildes abgebildet.

Sind die Haare in beiden Fällen dunkler als das Gesicht und die Helligkeitsdifferenzen zwischen dem Gesicht und den Haaren ist in einem Bild nur minimal, im anderen jedoch extern, führt dies bei Anwendung der MK-Methode zu demselben Effekt wie in Abbildung 6.1. Die LumHist-Methoden können in diesen Fällen jedoch durch die nichtlineare Luminanztransformation das Bild ansprechend anpassen, vgl. Abb. 6.2c.

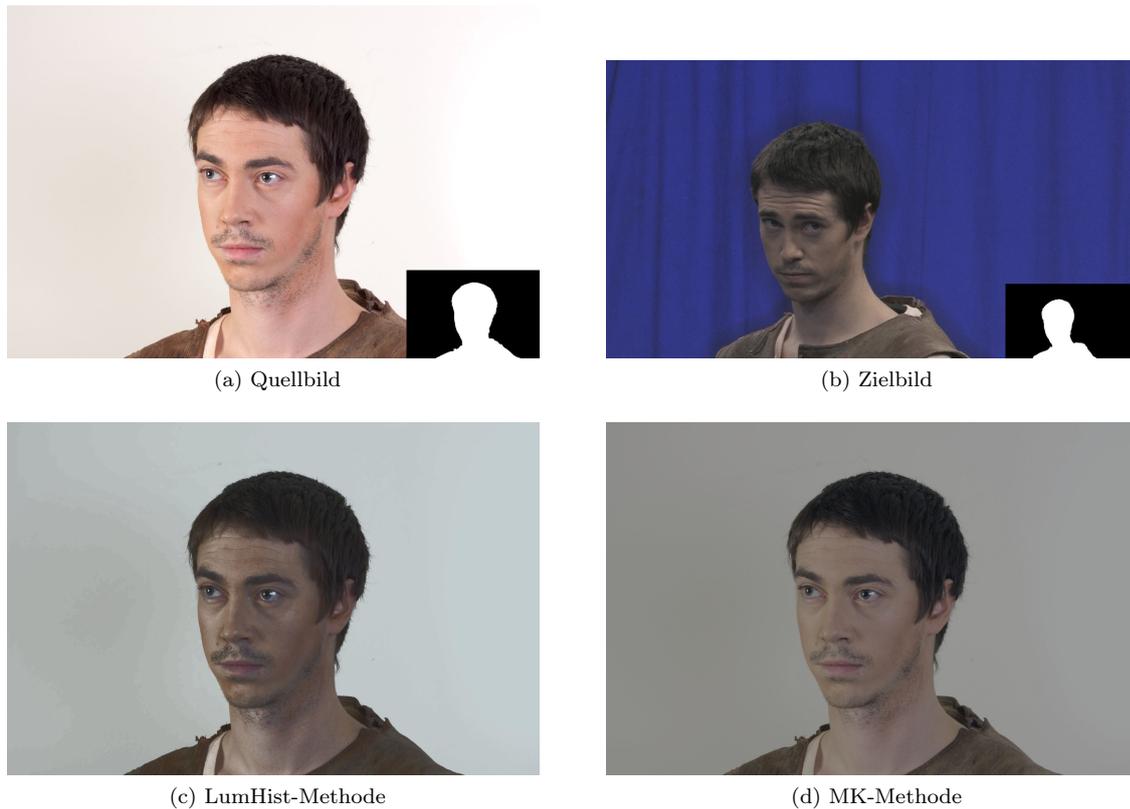


Abbildung 6.2: MK-Methode und LumHist(Lab)-Methode, angewendet auf Kopf und Oberkörper
Die jeweils verwendete Maske ist im unteren rechten Teil des Bildes abgebildet.

Abbildung 6.3 zeigt die Anpassung des Rauschverhaltens. Zusätzlich wurde zuerst die Unschärfe angepasst. Dies hat den Vorteil, dass das Farbrauschen des Quellbildes durch den hierbei verwendeten Gauß-Filter nahezu vollständig entfernt wurde. Eine Anpassung des Rauschverhaltens ist folglich durch die pixel- und kanalweise Addition einer normalverteilten Zufallsvariablen mit der zuvor geschätzten intensitätsabhängigen Standardabweichung denkbar. Zu einem besseren Vergleich des Ergebnisses mit dem Zielbild wurde vor dem Hinzufügen des simulierten Rauschens eine Farbanpassung mit der LumHist(Lab)-Methode mit 3 Chrominanzbändern mit einer Ausreißerentfernung für jedes Band im ab-Raum vorgenommen. Die Abbildungen 6.3f und 6.3g zeigen beide denselben Ausschnitt des bezüglich der Farben, der Schärfe und des Rauschverhaltens angepassten Quellbildes und des Zielbildes. Den Ausschnitt des Quellbildes und die einzelnen Zwischenergebnisse der Anpassung (Farbanpassung mit anschließender Anpassung der Unschärfe) zeigen die Abbildungen 6.3c bis 6.3d. Eine Bewertung der Anpassung ist bei unterschiedlicher Größe der Region, die von einem Pixel abgebildet wird, nur schwer möglich. Die in Abbildung 6.3h dargestellte Bildregion wurde vor der Anpassung des Rauschverhaltens durch Verringerung der Anzahl an Pixeln (downsampling) so verkleinert, dass ein Pixel in etwa dieselbe Fläche wie im Zielbild abbildet. Ein Vergleich des Ausschnitts aus dem Zielbild mit Abbildung 6.3h zeigt, dass die durch Demosaicking entstandene Korrelation zwischen benachbarten Pixeln bei der Simulation des Farbrauschens ebenfalls beachtet werden muss.

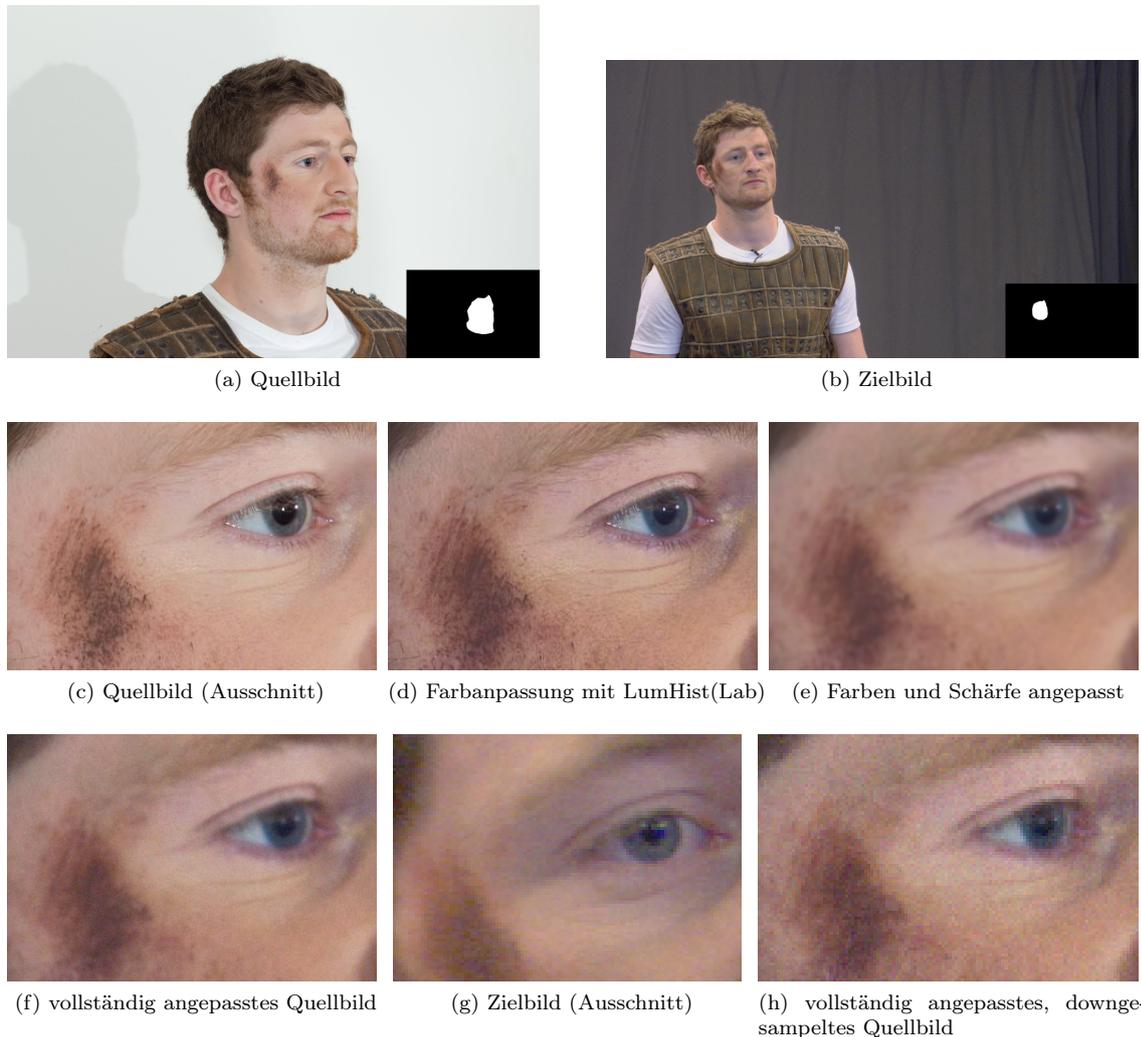


Abbildung 6.3: Anpassung des Farbrauschens

Die jeweils verwendete Maske ist im unteren rechten Teil des Bildes abgebildet.

Die vorgestellten Verfahren können auch bei anderen Motiven zu ansprechenden Resultaten führen. In Abbildung 6.4 wurden die Farben eines aktuellen Bildes des ehemaligen Parlamentsgebäudes in Grenoble an die eines eingescannten Bildes von 1968 angepasst. Die Anpassung des Himmels ist getrennt von der des restlichen Bildes erfolgt. Deutlich zu erkennen ist, dass für eine authentische Anpassung weitere Qualitätsmerkmale angepasst werden müssen. Diese betreffen unter anderem durch das Einscannen bzw. durch die Alterung der Aufnahme bedingte Effekte. Die Vignettierung im Quellbild sollte für eine optimale Anpassung ebenfalls kompensiert werden.



(a) Quellbild



(b) Zielbild



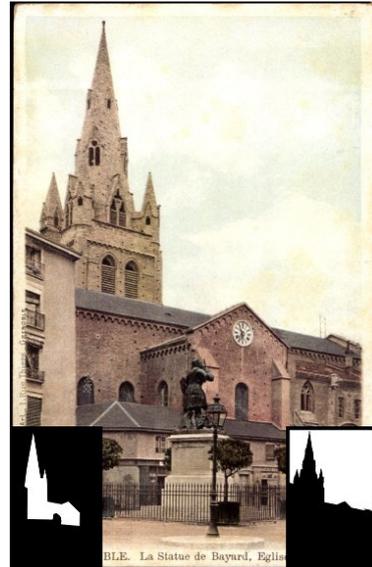
(c) angepasstes Quellbild

Abbildung 6.4: Anpassung anderer Motive (Parlamentsgebäude Grenoble)
 Die jeweils verwendete Maske ist im unteren linken bzw. rechten Teil des Bildes abgebildet.

Abbildung 6.5 zeigt als weiteres Beispiel die Anpassung einer aktuellen Aufnahme der Saint-André Kirche in Grenoble durch eine Digitalkamera an eine eingescannte Aufnahme von 1960. Auch hier wurde der Himmel getrennt von dem Rest des Bildes angepasst.



(a) Quellbild



(b) Zielbild



(c) angepasstes Quellbild

Abbildung 6.5: Anpassung anderer Motive (Kirche Grenoble)
Die jeweils verwendete Maske ist im unteren linken bzw. rechten Teil des Bildes abgebildet.

Kapitel 7

Fazit

Bei den in dieser Arbeiten vorgestellten Verfahren zur Farbanpassung konnten die besten Resultate mit einer getrennten Anpassung der Luminanz und der Chrominanz erreicht werden, wobei die Luminanz nichtlinear und die Chrominanz affin angepasst wurde. Die Anpassung der Schärfe der Aufnahmen wurde mit einem Gauß-Filter realisiert und unterliegt einigen Restriktionen, wie dem Vorhandensein identischer, annähernd planarer und zur Bildebene paralleler Flächen. Die Varianz des Farbrauschen konnte mit einer exponentiell gedämpften Funktion abhängig vom Intensitätswert beschrieben werden. Zur Nachbildung des Rauschens müssen jedoch zusätzlich die durch die interne Verarbeitung der Kamera bedingte Korrelationen der einzelnen Pixel, wie sie beispielsweise durch das Demosaicking entstehen kann, berücksichtigt werden.

Für die Anpassung entscheidende Schwellwerte oder Parameter der Verfahren können nicht pauschal bestimmt werden, sondern sind in der Regel stark von den jeweiligen Szenen abhängig. Ebenso ist eine Anpassung der vollständigen Aufnahme nicht immer möglich, sodass einzelne Bereiche getrennt angepasst werden müssen. Die ist meist der Fall bei Objekten vor verschiedenen Hintergründen oder bei Freiluftaufnahmen, bei denen eine getrennte Anpassung des Himmels meist nötig ist.

Anhang A

Anhang

A.1 Kurzanleitung Software

Im Rahmen dieser Arbeit wurde ein GUI-basiertes Programm zur Verwendung und Bewertung der vorgestellten Verfahren entwickelt. Als externe Bibliotheken wurden Qt für die graphische Oberfläche und OpenCV zur Implementierung der Verfahren verwendet. Abbildung A.1 zeigt einen Screenshot des Programms. Das Einlesen der Bilder ist unter dem Menüpunkt File und der Masken unter Preprocessing möglich. Die Methoden zur Farbanpassung sind unter dem Menüpunkt Colortrafos zu finden, die Anpassung der Unschärfe und des Rauschverhaltens unter Noise Estimation. Das obere linke Bild zeigt das Quellbild und das obere rechte Bild das Zielbild. Das Ergebnis der Anpassung ist unten links zu finden. Rechts davon können das Zielbild und das Transformationsergebnis direkt verglichen werden. Der horizontale linke Slider unterhalb der beiden Bilder ermöglicht das Abschneiden des Ergebnisses an der rechten Seite und der rechte Slider das Abschneiden des Zielbildes an der linken Seite. Mit dem linken vertikalen Slider ist eine Skalierung des Zielbildes möglich, mit dem Slider rechts von diesem kann das Zielbild vertikal verschoben werden. Eine ausführlichere Anleitung befindet sich in dem den Quellcode enthaltenden Ordner.

A.2 Übersicht zugehöriger Datenträger

Auf dem mit der Arbeit abgegebenen Datenträger sind neben diesem pdf-Dokument noch die Ordner *Program* und *Results* zu finden.

Der Ordner *Program* enthält den Quellcode des für diese Arbeit entwickelten Pro-

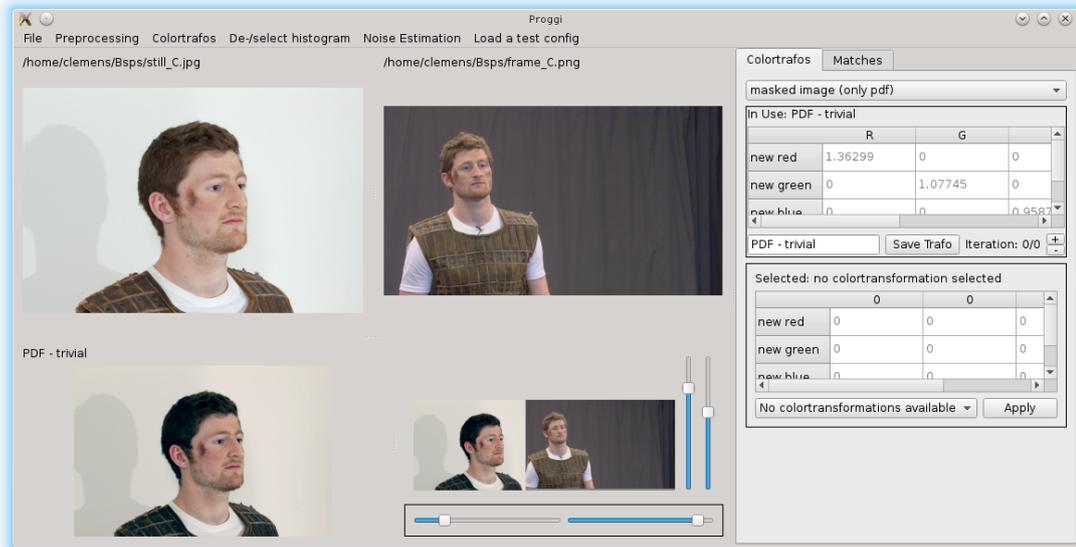


Abbildung A.1: Graphische Oberfläche

gramms. Zur Kompilierung wird ein C++-Compiler, der den c++11-Standard unterstützt, und Qt 5.0 oder höher benötigt. Alle weiteren benötigten Bibliotheken sind in der .pro-Datei des Qt-Projektes genannt.

Der Ordner *Results* enthält in Unterordnern weitere Beispiele bezüglich der Anpassung der Farbe, Schärfe und des Rauschverhaltens. Was bei den jeweiligen Bildern angepasst wurde, steht in der zugehörigen *info.txt*.

Literaturverzeichnis

- [Aus14] *Sensorgen - digital camera sensor data*. <http://www.sensorgen.info>, July 2014.
- [Bay76] B.E. Bayer. Color imaging array, July 20 1976. US Patent 3,971,065.
- [BSPP13] Nicolas Bonneel, Kalyan Sunkavalli, Sylvain Paris, and Hanspeter Pfister. Example-based Video Color Grading. *ACM Trans. Graph.*, 32(4):39:1–39:12, July 2013.
- [Erh08] Angelika Erhardt. *Einführung in Die Digitale Bildverarbeitung: Grundlagen, Systeme und Anwendungen*, pages 116–117. Vieweg+Teubner Verlag, 2008.
- [FB81] Martin A. Fischler and Robert C. Bolles. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Commun. ACM*, 24(6):381–395, June 1981.
- [GN03] M.D. Grossberg and S.K. Nayar. What is the Space of Camera Response Functions? In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, volume II, pages 602–609, Jun 2003.
- [Gol14] Chris Goldberg. Woman in Red Light - London 2014. <http://www.flickrriver.com/photos/chrisgold/14025789165>, April 2014. CC BY-NC 2.0 <http://creativecommons.org/licenses/by-nc/2.0/>, zugeschnitten und entsprechend Abbildungsbeschreibung modifiziert.
- [Hof14] Gernot Hoffmann. CIE Lab Color Space. <http://docs-hoffmann.de/cielab03022003.pdf>, July 2014.

- [Jen14] Oliver Jennrich. Ein Blick auf die Schärfentiefe. <http://www.traxel.de/foto/drf/schaerfentiefe.pdf>, July 2014.
- [JLW⁺12] Yang Jin, Zhen Liu, Peng Fei Wang, San Guo Liu, and Hong Jie Zhai. Study on Correlation of Color Components Image in Different Color Spaces. In *Applied Mechanics and Materials (Volume 262)*, pages 86–91, Dec 2012.
- [JWV⁺05] N. Joshi, B. Wilburn, V. Vaish, M. Levoy, and M. Howoritz. Automatic Color Calibration for Large Camera Arrays. *UCSD CSE Tech Report CS2005-0821*, May 2005.
- [Kan14] Prof. Dr. Christian Kanzow. Nichtlineare Gleichungssysteme - Kapitel 2 - Newton-Verfahren. http://www.mathematik.uni-wuerzburg.de/~kanzow/ne/NE_07_Kap2.pdf, July 2014.
- [Kor14] Norman Koren. Understanding resolution and MTF. <http://www.normankoren.com/Tutorials/MTF.html>, July 2014.
- [LFSK06] Ce Liu, W.T. Freeman, R. Szeliski, and Sing Bing Kang. Noise estimation from a single image. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 901–908, June 2006.
- [Low04] David G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vision*, 60(2):91–110, November 2004.
- [MH04] Beate Meffert and Olaf Hochmuth. *Werkzeuge der Signalverarbeitung: Grundlagen, Anwendungsbeispiele, Übungsaufgaben*. Pearson Studium, München, 2004.
- [mou06] mouton.rebelle. Emilie fait le dragon. <https://www.flickr.com/photos/mouton/150004414>, May 2006. CC BY-NC 2.0 <http://creativecommons.org/licenses/by-nc/2.0/>, zugeschnitten und entsprechend Abbildungsbeschreibung modifiziert.
- [OP82] I. Olkin and F. Pukelsheim. The distance between two random vectors with given dispersion matrices. *Linear Algebra and its Applications*, 48(0):257 – 263, 1982.

- [Pix14] Pixar. Motion Blur & Depth of Field. <http://renderman.pixar.com/view/motion-blur-and-depth-of-field>, July 2014.
- [PK07] F. Pitié and A Kokaram. The linear Monge-Kantorovitch linear colour mapping for example-based colour transfer. In *Visual Media Production, 2007. IETCVMP. 4th European Conference on*, pages 1–9, Nov 2007.
- [RSBI02] Rajeev Ramanath, Wesley E. Snyder, Griff L. Bilbro, and William A. Sander Iii. Demosaicking methods for Bayer color arrays. *Journal of Electronic Imaging*, 11:306–315, 2002.
- [RSYD05] R. Ramanath, W.E. Snyder, Y. Yoo, and M.S. Drew. Color image processing pipeline. *Signal Processing Magazine, IEEE*, 22(1):34–43, Jan 2005.
- [SDP⁺09] Fred Sigernes, Margit Dyrland, Nial Peters, Dag Arne Lorentzen, Trond Svenøe, Karsten Heia, Sergey Chernouss, Charles Sterling Deehr, and Mike Kosch. The absolute sensitivity of digital colour cameras. *Opt. Express*, 17(22):20211–20220, Oct 2009.
- [STI14] PROF. DR.-ING. C. STILLER. Tabelle zur Fourier-Transformation. <http://info.php-4.info/attachment.php?attachmentid=385&sid=1887badd70653b8350f0af6105ae9db5>, July 2014.
- [TF05] J.R.A Torreao and J.L. Fernandes. Single-Image Shape from Defocus. In *Computer Graphics and Image Processing, 2005. SIBGRAPI 2005. 18th Brazilian Symposium on*, pages 241–246, Oct 2005.
- [Wag14] Patrick Wagner. Farbtemperatur, Vollautomatischer halbautomatischer manueller Weißabgleich, Schwarzer Körper, Strahlung. <http://www.filmscanner.info/Farbtemperatur.html>, July 2014.

Selbständigkeitserklärung

Ich erkläre hiermit, dass ich die vorliegende Arbeit selbständig verfasst und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe. Weiterhin erkläre ich, eine ...arbeit in diesem Studienggebiet erstmalig einzureichen.

Berlin, den 6. November 2014

.....

Statement of authorship

I declare that I completed this thesis on my own and that information which has been directly or indirectly taken from other sources has been noted as such. Neither this nor a similar work has been presented to an examination committee.

Berlin, 6th November 2014

.....