

Schattenentfernung unter Verwendung des Retinex-Algorithmus

Sarah Steinmetz, Dietrich Paulus und Wolfram Hans

Universität Koblenz-Landau

Aktives Sehen

{sarah.steinmetz,paulus,hans}@uni-koblenz.de,

<http://www.uni-koblenz.de/agas>

Zusammenfassung. In [1] wird ein Verfahren zur Schattenentfernung in Farbbildern mit dem Retinex-Algorithmus vorgestellt. Das Verfahren geht davon aus, dass bekannt ist, wo sich die Schattenkanten im Bild befinden. Zur Erzeugung des Schattenkantenbilds wird auf [2], [3] verwiesen. Dort wird eine beleuchtungsinvariante Darstellung berechnet, um zu bestimmen, welche Kanten im Bild Schattenkanten darstellen. In diesem Beitrag stellen wir die Verfahren kurz vor und gehen auf die eigene Umsetzung ein. Außerdem wird ein erster Ansatz vorgestellt, wie Verfahren zur Schattenentfernung quantitativ bewertet werden können. Hierzu wird auch auf die Erstellung der benötigten Ground Truth Daten mit Hilfe von Laboraufnahmen eingegangen.

1 Einleitung

In diesem Beitrag stellen wir das Verfahren zur Schattenentfernung mit dem Retinex [1], sowie die dazu benötigte Schattenkantendetektion [2] vor und gehen auf die eigene Umsetzung ein. Ursprünglich war geplant, die Verfahren zur Schattenentfernung in Campusbildsequenzen einzusetzen. In Abschnitt 2 wird erläutert, welche Probleme sich dabei im Zusammenhang mit dem Verfahren zur Schattenkantendetektion ergeben. In diesem Abschnitt wird auch auf die erforderliche Kamerakalibrierung eingegangen. In [2] werden dazu mehrere Außenaufnahmen des ColorCheckers verwendet, hier wird die Kalibrierung mit künstlichen Lichtquellen vorgestellt. In Abschnitt 3 folgt die Erläuterung des ursprünglichen Retinex-Algorithmus [4], der Modifikation in [1] zur Schattenentfernung und unserer Anpassungen des Verfahrens. In Abschnitt 4 wird ein erster Ansatz zur quantitativen Bewertung von Schattenentfernungsverfahren, sowie die Erstellung der benötigten Ground Truth Daten mit Hilfe von Laboraufnahmen vorgestellt.

Da die in dieser Arbeit untersuchten Verfahren von einem stark vereinfachten Modell der Bildentstehung ausgehen, wird hier zunächst kurz darauf eingegangen: $f^{(k)}(\mathbf{x}) = E(\lambda_k, \mathbf{x}^w)\rho(\lambda_k, \mathbf{x}^w)$. Der Wert $f^{(k)}(\mathbf{x})$ an Pixelposition \mathbf{x} im Kanal k , mit $k \in \{r, g, b\}$, ergibt sich aus dem Produkt von Bestrahlungsstärke $E(\lambda_k, \mathbf{x}^w)$ und Reflektivität $\rho(\lambda_k, \mathbf{x}^w)$ am Oberflächenpunkt \mathbf{x}^w . Dabei wird angenommen, dass die Szene nur aus Lambert'schen Flächen besteht und dass ein

linearer Zusammenhang zwischen der Intensität des Lichts, das auf den Sensor trifft, und den Pixelwerten besteht. Die meisten CCD-Sensoren liefern zwar lineare Werte, aber oft finden in der Kamera nichtlineare Verarbeitungsschritte, wie zum Beispiel eine Gamma-Korrektur statt. Außerdem wird angenommen, dass die Sensorempfindlichkeitskurven der Kamera perfekt schmalbandig sind. Der k -te Sensor reagiert also nur auf eine bestimmte Wellenlänge λ_k .

2 Schattenkantendetektion

Zur Detektion der Schattenkante wird in [5], [2] zunächst ein beleuchtungsinvariantes Bild berechnet. Die invarianten Werte sind nur noch abhängig von der Reflektanz des abgebildeten Oberflächenpunktes. Die gesuchten Schattenkanten sind dann die Kanten, die im Ausgangsfarbbild auftreten, aber nicht im invarianten Bild.

2.1 Beleuchtungsinvariante Darstellung

Das Verfahren geht von der Annahme aus, dass jede Lichtquelle in der Szene als Planck'scher Strahler, der durch seine Farbtemperatur definiert ist, approximiert werden kann. In [3] wird mathematisch erläutert, dass unter diesen Annahmen zwischen logarithmischen Chromatizitätswerten, die Oberflächenpunkte der selben Reflektanz unter verschiedener Beleuchtung darstellen, ein linearer Zusammenhang besteht. Abb. 1 veranschaulicht idealisiert diesen Zusammenhang. Dargestellt sind die logarithmischen Kanalverhältnisse $\ln(\frac{R}{G})$ und $\ln(\frac{B}{G})$, die sich aus der Abbildung von 5 verschieden farbigen Punkten unter 6 verschiedenen Beleuchtungen ergeben. Oberflächenpunkte der selben Reflektanz werden im Log-Chromatizitäts-Raum auf eine bestimmte Linie abgebildet¹. Der Ort dieser Linie ist durch die Reflektanz bestimmt und die jeweilige Position auf dieser Linie ist nur abhängig von der Farbtemperatur der Beleuchtung an diesem Punkt. Die Linien sind parallel und ihre Richtung, die sogenannte beleuchtungsinvariante Richtung, ist nur abhängig von den spektralen Eigenschaften der verwendeten Kamera. Die gesuchte beleuchtungsinvariante Darstellung ergibt sich aus der Projektion der logarithmischen Kanalverhältnisse entlang der beleuchtungsinvarianten Richtung. Die invarianten Werte sind nur abhängig von der Reflektanz des abgebildeten Oberflächenpunktes und unabhängig von Intensität und Farbe der Beleuchtung.

Berechnung der invarianten Richtung durch Kamerakalibrierung Die invariante Richtung einer bestimmten Kamera kann durch eine Art Kalibrierung bestimmt werden. Dazu werden mehrere Aufnahmen des Macbeth ColorChecker

¹ Achromatische Oberflächenpunkte, sowie allgemein Farben, die sich nur in der Intensität unterscheiden werden jeweils auf eine bestimmte Linie abgebildet.

bei verschiedenen Beleuchtungssituationen (Farbtemperaturen) benötigt. Konkret werden in [2] im Freien 14 Bilder des ColorCheckers zu verschiedenen Tageszeiten aufgenommen. Tageslicht kann im Allgemeinen als Planck'scher Strahler approximiert werden [3], da sich die Chromatizitätswerte verschiedener Tageslichtphasen nahe der Planck'schen Ortskurve im CIE-xy-Chromatizitätsdiagramm befinden [6]. Für jedes Bild, und für jedes Patch, werden dann der Mittelwert und die zugehörigen Log-Chromatizitäts-Koordinaten berechnet. Für jedes Patch wird der Mittelwert der Log-Chromatizitätswerte über alle Bilder berechnet und von den Datenpunkten dieses Patches subtrahiert, was der Translation der Linien, entlang derer sich die Patch-Datenpunkte orientieren, in den Ursprung entspricht. Die beleuchtungsinvariante Richtung ergibt sich dann aus der Richtung der Geraden, die den mittleren quadratischen Abstand zu den Punkten minimiert.

Wir haben die Kalibrierung unter Verwendung künstlicher Lichtquellen durchgeführt. Eine Lichtquelle kann als Planck'scher Strahler beschrieben werden, wenn ihr Spektrum proportional zu dem eines Planck'schen Strahlers ist. Beide Spektren liefern in diesem Fall die selben xy-Chromatizitätskoordinaten. Das kontinuierliche Spektrum bestimmter glühender Materialien, wie z.B. Wolfram, ähnelt dem eines Planck'schen Strahlers. In Standardglühbirnen wird als Material für den Glühdraht meist Wolfram verwendet. Wolfram-Glühfadenlampen oder Halogenglühlampen können somit als Planck'scher Strahler beschrieben werden [6]. Für andere künstliche Lichtquellen gilt dies in der Regel nicht. In der Praxis wird im Allgemeinen aber angenommen, dass Lichtquellen als Planck'scher Strahler approximiert werden können, wenn sich ihre xy-Chromatizitätskoordinaten nahe der Planck'schen Ortslinie befinden [3]. Wir haben zunächst folgende künstliche Lichtquellen verwendet: Aus der Pantone Lichtbox die Standardlichtarten A (entspricht einer Wolframglühfadenlampe korrelierter Farbtemperatur 2856 K), D65 (repräsentiert eine Tageslichtphase korrelierter Farbtemperatur von ca. 6504 K) und TL84 (Leuchtstoffröhre), einen Halogenstrahler von 100 und 400 Watt, und eine Kohlebogenlampe. Außerdem haben wir eine 100 Watt Glühbirne mit einem Dimmer geregelt, um so verschiedene Farbtemperaturen zu erzeugen. An der Position des ColorCheckers wurde ein Beleuchtungsstärkemessgerät platziert, welches neben der Beleuchtungsstärke auch die xy-Chromatizitätskoordinaten ausgibt. Abb. 1 zeigt, dass die ermittelten Koordinaten für alle Lichtquellen nahe der Planckschen Ortskurve liegen. Erstaunlich ist dies bezüglich der Leuchtstoffröhre, da diese ein Linienspektrum besitzt und somit nicht mit dem kontinuierlichen Spektrum eines Planck'schen Strahlers vergleichbar ist [6]. Diese Lichtquelle wurde daher bei der Kalibrierung nicht verwendet. Die Chromatizitätskoordinaten der Glühbirne zeigen deutlich, dass durch das Dimmen verschiedene Farbtemperaturen erreicht werden können.

Ergebnisse Die Kalibrierung wurde zunächst mit einer Netzwerkkamera durchgeführt, die für die Campusbildsequenzen verwendet wurde. Die Sensorantworten werden durch automatischen Weißabgleich und JPEG-Kompression kameraintern verarbeitet, sodass keine linearen Bildwerte vorliegen. Abb. 2 zeigt links das

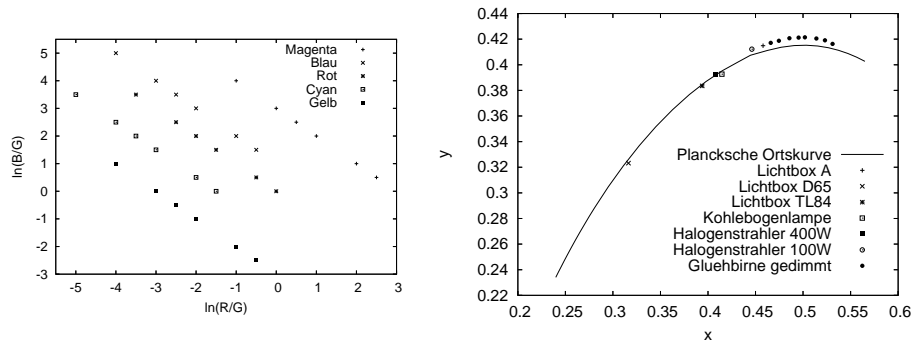


Abb. 1. Links: Log-Chromatizitäts-Koordinaten, die sich aus der Abbildung von 5 verschiedenen farbigen Punkten unter 6 verschiedenen Beleuchtungen ergeben, rechts: CIE xy-Chromatizitätsdiagramm mit Planck'scher Ortskurve und xy-Koordinaten der Lichtquellen

Ergebnis der Kalibrierung. Außer der Leuchtstoffröhre wurden dabei alle Lichtquellen verwendet. Von den Aufnahmen unter Einsatz der gedimmten Glühbirne konnten nur 3 verwendet werden, da das Bildrauschen bei niedrigen Dimmstufen zu groß wurde. Trotz der genannten Kameraeigenschaften zeigt das Ergebnis den gesuchten linearen Zusammenhang der Log-Chromatizitätswerte. Erwartungsgemäß liegen die achromatischen Patches (18 bis 23) annähernd auf einer Linie. Dies gilt aber auch für einige Datenpunkte, die unterschiedlich farbige Patches repräsentieren, sodass unterschiedliche Reflektanzen in der invarianten Darstellung nicht mehr unterschieden werden können. Abb. 3 zeigt ein beleuchtungsinvariantes Bild, das anhand der kalibrierten invarianten Richtung berechnet wurde. Die Schattenkante ist weiterhin erkennbar und die JPEG-Artefakte werden verstärkt. Außerdem verschwinden in der Darstellung auch Materialkanten, sodass eine Unterscheidung zwischen Schatten- und Materialkanten schwer möglich ist. Grundsätzlich ist das Verfahren für die Campusbilder eher nicht geeignet, da die Szenen viele achromatische Flächen besitzen.

Für die Durchführung der Laborexperimente wurde eine Firewirekamera (The Imaging Source, Modell DFK 31BF03) kalibriert, wobei nichtlineare Verarbeitung deaktiviert wurde. Abb. 2 zeigt das Ergebnis. Da für die Laborexperimente eine 100 Watt Glühbirne sowie der 400 Watt Halogenstrahler verwendet wurden, wurde die Kalibrierung unter Einsatz des Halogenstrahlers, 6 Dimmstufen der Glühbirne, sowie der Standardlichtart A durchgeführt. Die achromatischen Patches werden korrekt auf eine Linie abgebildet, die übrigen Linien zeigen geringeres Rauschen als bei dem Ergebnis der Netzwerkkamera. Der mittlere quadratische Anpassungsfehler, beträgt hier 0.00297, bei der Netzwerkkamera 0.012076.

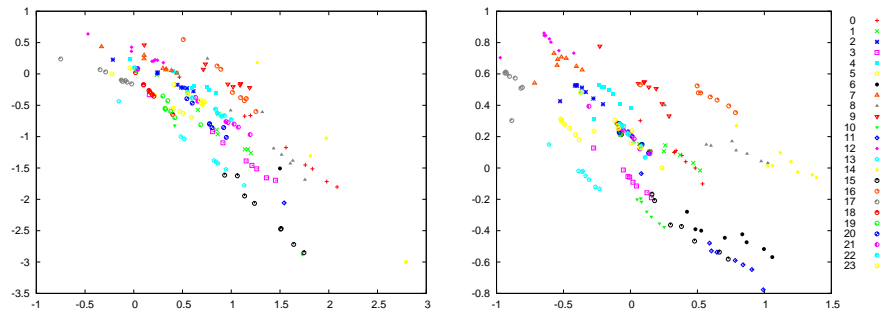


Abb. 2. Ergebnisse der Kalibrierung, links: Netzwerkkamera, rechts: Firewirekamera



Abb. 3. Farbbild und berechnetes beleuchtungsinvariantes Bild

2.2 Berechnung des Schattenkantenbilds

Ein Pixel wird nach [2] dann als Schattenkantenpixel klassifiziert, wenn es in einem der drei Kanäle des Ausgangsfarbbildes ein Kantenpixel darstellt, aber nicht im Kantenbild des invarianten Bilds. Wir haben das Verfahren zunächst wie in [2] erläutert implementiert, beim Vergleich der Schattenkantenbilder waren aber Anpassungen erforderlich: Da sich die Kantenpixel, die den selben Kantenpunkt in der Szene repräsentieren im invarianten Kantenbild und in den Kantenbildern der Farbkanäle nicht exakt an der gleichen Stelle befinden (versursacht durch unterschiedliche Parameter bei der Glättung und beim Canny-Operator), ergab der einfache Vergleich der Kantenpixel zahlreiche, vereinzelte, falsch klassifizierte Schattenkantenpixel. Um dieses Problem zu umgehen, führen wir auf dem invarianten Kantenbild vor dem Vergleich eine Dilatation durch. Wie in Abschnitt 2.1 und in [2], [5] angesprochen, werden Materialkanten im invarianten Bild nicht erkannt, wenn die Farbwerte benachbarter Flächen auf ähnliche, oder sogar die selben invarianten Werte abgebildet werden. In diesem Fall wird eine Materialkante fälschlicherweise als Schattenkante klassifiziert. Daher nehmen wir zusätzlich einen Vergleich mit dem Kantenbild des H-Kanals der HSV-Darstellung des Farbbildes vor. Nach [7] ändert sich die Farbinformation bei Übergang in den Schattenbereich nur gering. Hier wird daher angenommen, dass Kanten mit großer Änderung des Hue-Werts Materialkanten und keine Schattenkanten dar-

stellen. Die Parameter des Canny-Operators, der auf den Hue-Kanal angewendet wird, werden so gewählt, dass nur sehr starke Kanten detektiert werden. Detektierte Schattenkantenpixel, an deren Stelle sich im Hue-Kantenbild ein Kantenpixel befindet, werden verworfen. Problematisch ist dieses Vorgehen, wenn sich die Lichtquellen in der Szene stark in der Farbe unterscheiden.

3 Schattenentfernung mit dem Retinex

Ursprünglich stellt die Retinex-Theorie ein Modell der menschlichen Farbwahrnehmung nach Edwin Land [4] dar. Der Algorithmus wurde in der Bildverarbeitung für verschiedene Anwendungszwecke, wie z.B. Farbkorrektur oder Dynamikreduktion [8], angepasst und erweitert. In [1] wird das Verfahren so angepasst, dass damit Schatten entfernt oder reduziert werden können.

3.1 Pfadbasierter Retinex-Algorithmus

Nach Land findet die Farbwahrnehmung durch 3 unabhängige Retinex-Prozesse statt, wobei in jedem visuellen Kanal die relative Reflektanz, oder auch Lightness, der Punkte im Sichtfeld bestimmt wird. Um die relative Reflektanz eines Punktes zu bestimmen, wird seine Reflektanz relativ zu anderen Punkten im Sichtfeld bestimmt. Für direkt benachbarte Punkte entspricht die relative Reflektanz dem Verhältnis der gemessenen Intensitäten, unter der Annahme, dass die Beleuchtung E annähernd gleich ist: $\frac{\rho_2 E}{\rho_1 E} = \frac{\rho_2}{\rho_1}$. Für entfernte Punkte berechnet sich die relative Reflektanz als Produkt der Intensitätsverhältnisse entlang eines Verbindungspfades (Pfadprodukt), da die Beleuchtung E an jeweils benachbarten Punkten als annähernd gleich angenommen wird: $\frac{E_1 \rho_2}{E_1 \rho_1} \cdot \frac{E_2 \rho_3}{E_2 \rho_2} \cdot \frac{E_3 \rho_4}{E_3 \rho_3} = \frac{\rho_4}{\rho_1}$. Wenn der Startpunkt des Pfades die maximale Reflektanz von 1 besitzt, entspricht das Pfadprodukt dem gesuchten Lightness-Wert. Um den Lightness-Wert eines Punktes zu berechnen, werden daher mehrere zufällige Pfade zu diesem Punkt verfolgt und das Ergebnis gemittelt. Für den Startpunkt eines Pfades wird dabei zunächst der maximale Reflektanzwert angenommen. Das Pfadprodukt entspricht also dem Lightness-Wert am jeweiligen Punkt. Wenn das Pfadprodukt > 1 wird, so muss die Reflektanz des aktuellen Punktes größer als die des Startpunktes sein, sodass ein Reset erfolgt, d.h. ab diesem Punkt wird die Berechnung des Pfadprodukts neu gestartet. Wenn tatsächlich ein Punkt mit der maximalen Reflektanz gekreuzt wird, entsprechen alle folgenden Pfadprodukte dem Reflektanzwert am aktuellen Punkt. Um den Lightness-Wert eines Punktes zu schätzen, wird daher eine Mittelung mehrerer Pfadprodukte an diesem Punkt berechnet.

Der Retinex arbeitet auf logarithmierten Pixelwerten, die Division wird also zur Differenz, die Multiplikation zur Summe. Der Lightnesswert $l_k(\mathbf{x})$ des Pixel \mathbf{x} im Kanal $k \in \{r, g, b\}$ ergibt sich aus der Mittelung von n sequentiellen Pfadsummen

$$l_k(\mathbf{x}) = \frac{1}{n} \sum_{i=1}^n s_{k, \mathbf{x}_{m_i}}, \quad (1)$$

n ist die Anzahl der Pfade, die \mathbf{x} kreuzen. Pfad i startet an der zufälligen Position \mathbf{x}_0^i und führt über m_i Pixel zur Pixelposition $\mathbf{x}_{m_i} = \mathbf{x}$. Beim Start wird die Pfadsumme mit 0 initialisiert und entlang des Pfades immer um das logarithmierte Verhältnis des aktuellen Pixelwerts zum vorigen Pixelwert erhöht:

$$s_{k, \mathbf{x}_0^i} = 0 \quad (2)$$

$$s_{k, \mathbf{x}_j^i} = \delta(s_{k, \mathbf{x}_{j-1}^i} + \Theta(\log(f^{(k)}(\mathbf{x}_j^i)) - \log(f^{(k)}(\mathbf{x}_{j-1}^i)))), \quad (3)$$

mit $j = 1, \dots, m_i - 1$. Die Funktion δ realisiert den Reset Schritt:

$$\delta(s) = \begin{cases} 0 & \text{für } s > 0 \\ s & \text{sonst} \end{cases} \quad (4)$$

Durch den Reset Schritt wird eine Farbkorrektur der Pixelwerte erreicht, da im Mittel versucht wird, eine Normierung mit dem hellsten Pixel durchzuführen. Die Schwellwertfunktion $\Theta(d)$ eliminiert den Einfluss nicht uniformer Beleuchtung:

$$\Theta(d) = \begin{cases} 0 & \text{für } d \leq \theta \\ d & \text{sonst} \end{cases} \quad (5)$$

Differenzen d , die annähernd Null ergeben, werden somit auf Null gesetzt, weil angenommen wird, dass leichte Unterschiede benachbarter Pixelwerte durch Schwankungen in der Beleuchtung verursacht werden.

3.2 Schattenentfernung

Das Retinex-Verfahren eliminiert leichte Beleuchtungsschwankungen. In [1] wird die Schwellwertfunktion so angepasst, dass Schatten eliminiert werden. Angenommen ein binäres Schattenkantenbild $S(\mathbf{x})$, in dem Schattenkantenpixel den Wert 1 haben, ist gegeben. Die neue Schwellwertfunktion lautet dann:

$$\Theta_S(d_j) = \begin{cases} 0 & \text{für } S(\mathbf{x}_j) = 1 \\ d_j & \text{sonst} \end{cases} \quad (6)$$

An einer Schattenkante wird also die Log-Differenz benachbarter Pixel auf 0 (ihr Verhältnis auf 1) gesetzt.

Implementierung Die Pfade werden in [1] nicht rein zufällig erzeugt, sondern, wie in [9] als eine Approximation Brownscher Bewegung, was die Anzahl der benötigten Pfade nach [9] um den Faktor 10 reduziert. Bei [1] wird nicht an jeder Pixelposition eine Menge von Pfaden erzeugt, sondern wie folgt verfahren: Die Pfade werden nur einmal von Startpixel $(0, 0)$ aus erzeugt. An jeder Pixelposition werden die Pfadpunkte um den aktuellen Vektor verschoben, wenn der Pfad aus dem Bild tritt, wird er ab dieser Position gestartet, der Rest wird verworfen. Die Pfade werden von den Endpunkten bis zum aktuellen Pixel verfolgt, wo die Pfadprodukte gemittelt werden.

Wir haben den Algorithmus zunächst wie in [1] implementiert. Um mehr Werte pro Pixel mitteln zu können, haben wir aber die ursprüngliche Idee von Land [10] aufgegriffen, jedes Pixel, welches ein Pfad kreuzt, zu aktualisieren. In einem Akkumulatorarray wird für jedes Pixel, das von einem Pfad gekreuzt wird, das aktuelle Pfadprodukt aufaddiert. In einem Counterarray wird gezählt, wie oft ein Pixel gekreuzt wurde, um den Wert im Akkumulatorarray am Ende zu mitteln.

Um nur eine Schattenentfernung und keine Korrektur der Beleuchtung durchzuführen, wird in [1] das Eingabebild mit einem weißen Rand versehen. Die Pfadlänge wird so gewählt, dass die meisten Pfade den weißen Rand kreuzen. Dieser Effekt lässt sich aber auch erreichen, indem das Pfadprodukt zu Beginn mit dem (logarithmierten) Wert des Startpixels und nicht mit 0 initialisiert wird. Die Reset-Funktion wird entsprechend angepasst: Wenn das Pfadprodukt größer wird als der Logarithmus des Maximalwerts, wird der Pfad neu gestartet². Zur Darstellung müssen die Werte nur exponentiert und nicht auf den darstellbaren Bereich skaliert werden. Wenn der Retinex nicht zur Farbkorrektur eingesetzt wird, werden theoretisch nicht mehrere Pfade benötigt, doch dies ist weiterhin erforderlich um Artefakte zu reduzieren.

Problemfälle Es ergeben sich folgende Problemfälle, auf die auch in [11] hingewiesen wird: Lokale Fehler an der Schattenkante wirken sich global aus, da sie entlang der Pfade propagiert werden. Lokale Fehler entstehen zum einen, wenn das Schattenkantenbild unvollständig ist und der Pfad an bekannter Schattenkante in den Schatten eintritt, aber an unbekannter Schattenkante wieder austritt. Auch wenn sich direkt an der Schattenkante verrauschte Pixelwerte befinden, werden fehlerhafte Werte entlang des Pfades weitergereicht, was im Endergebnis auch zu Falschfarben führen kann. Ein Problem besteht auch, wenn der Pfad im Schatten startet, da die Pixelwerte nach Austritt aus dem Schatten abgedunkelt werden. Falsch klassifizierte Schattenkanten führen zu starken Artefakten, indem Materialkanten verwischt werden. Um diese Fehler zu minimieren ist die Mittelung mehrerer Pfade bedeutsam.

4 Ansatz zur quantitativen Bewertung

Für Algorithmen zur Schattendetektion liegen bereits Evaluationsmetriken [12] vor. Verfahren zur Schattenentfernung wurden bisher nur qualitativ bewertet. Unser Ansatz zur quantitativen Bewertung von Verfahren zur Schattenentfernung orientiert sich zunächst an [11]. Dort wird ein Maß vorgeschlagen, um bei der Schattenentfernung entstandene Fehler im Bereich außerhalb des Schattens messen zu können: Bildbereiche außerhalb des Schattens sollen durch die Schattenentfernung nicht beeinflusst werden, daher werden die Gradienten in diesen

² Ohne Farbkorrektur ist eigentlich auch kein Reset mehr nötig, aber durch das Eingreifen in die Pixelwerte durch die Schattenschwellwertfunktion können Werte entstehen, die größer sind als der Maximalwert.

Bereichen verglichen. Sei $\nabla I = \left(\frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right)$ der Gradient von Pixeln außerhalb des Schattens im logarithmierten Originalbild I und $\nabla I'$ der Gradient dieser Pixel im berechneten logarithmierten schattenfreien Bild I' . Die Distanz d zwischen diesen Gradienten wird berechnet durch $d = \frac{|\nabla I - \nabla I'|}{\|\nabla I\|}$. Dies stellt ein Maß für die Abweichung der Gradienten der rückgerechneten Pixelwerte von den Gradienten der originalen Pixelwerte außerhalb des Schattenbereichs dar. Um die Pixel außerhalb des Schattenbereichs bestimmen zu können, werden in [11] synthetische Bilder verwendet.

Unter Berücksichtigung der typischen Probleme, die sich bei Verfahren zur Schattenentfernung ergeben, werden im Folgenden Anforderungen an ein solches Verfahren genannt:

1. Der Schatten wird entfernt
2. Keine Erzeugung von Artefakten außerhalb des Schattens. Schattenentfernungsergebnisse zeigen oft unerwünschte Artefakte. Dies stellt nach [11] ein inhärentes Problem der Verfahren dar, die wie [5] und indirekt auch [1] auf der Nullsetzung der Ableitungen im Bereich der Schattenkante und der anschließenden Reintegration basieren. Ein lokaler Fehler wirkt sich global aus, sodass Artefakte auch außerhalb des Schattenbereichs entstehen.
3. Keine Erzeugung von Artefakten innerhalb des Schattens
4. Erhaltung der Struktur im Bereich der Schattenkante. Bei gradientenbasierten Verfahren zur Schattenentfernung gehen im Bereich der Schattenkante Strukturinformationen verloren, da die Gradienten in diesem Bereich auf Null gesetzt werden.

Um Ergebnisse hinsichtlich dieser Anforderungen zu bewerten, soll ein Gradientenvergleich in verschiedenen Bereichen vorgenommen werden. Gegeben sei als Ground Truth die Gradientenkarte einer Aufnahme der Szene ohne Schatten. Der Betrag der Differenz der Gradienten der Ground Truth Gradientenkarte und der Gradientenkarte des schattenentfernten Bildes soll im Bereich der Schattengrenzen, außerhalb der Schattenregionen, sowie innerhalb der Schattengrenzen möglichst klein sein. Das Maß aus [11] ergibt immer den Wert 1, wenn der Gradient des schattenbefreiten Bildes Null ist, sodass in diesem Fall für verschiedene Gradienten im Vergleichsbild der selbe Wert berechnet wird. Hier wird daher der euklidische Abstand ohne Normierung verwendet. Der mittlere Abstand der Gradienten des Ground Truth Bildes f und des schattenbefreiten Bildes f' an Pixelposition (x, y) berechnet sich aus der Mittelung der Gradientendifferenzen in jedem Kanal: $d(f, f')(x, y) = \frac{1}{3} \sum_{k=1}^3 (|\nabla f^{(k)}(x, y) - \nabla f'^{(k)}(x, y)|)$. Der mittlere Gradientenabstand $d(f, f')$ ergibt sich aus der Mittelung der Gradientenabstände über alle Pixel.

Gegeben seien zwei Binärbilder als Ground Truth Schatteninformation: $S_e(x, y)$ hat im Bereich der Schattenkante und $S_r(x, y)$ für alle Pixel innerhalb von Schattengrenzen den Wert 1. Die Anzahl der Schattenkantenpixel sei n_e , der Pixel innerhalb der Schattengrenzen n_i und die der Pixel außerhalb des Schattens n_o . Es ergeben sich 3 Maße, die mittleren Gradientenabstände im Bereich der Schattenkante d_e , im Bereich innerhalb der Schattengrenzen d_i und außerhalb

der Schattengrenzen d_o :

$$d_e(f, f') = \frac{1}{n_e} \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} d(f, f')(x, y) S_e(x, y) \quad (7)$$

$$d_i(f, f') = \frac{1}{n_i} \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} d(f, f')(x, y) S_r(x, y) \quad (8)$$

$$d_o(f, f') = \frac{1}{n_o} \sum_{y=0}^{N-1} \sum_{x=0}^{M-1} d(f, f')(x, y) (1 - S_e(x, y)) (1 - S_r(x, y)) \quad (9)$$

Dabei ist MN die Anzahl der Pixel.

5 Experimente und Ergebnisse

Um Testbilder sowie die Ground Truth Daten zu erzeugen, wurde folgender Laboraufbau realisiert. Zur Aufnahme wurde die kalibrierte Firewire-Kamera verwendet. Als Szene, in die ein Schatten geworfen wird, dient der ColorChecker auf homogenem Hintergrund. Die Szene wird indirekt durch einen 400 Watt Halogenstrahler und direkt durch eine 100 Watt Glühbirne beleuchtet. Der Schatten wird durch eine Holzleiste verursacht, die außerhalb der sichtbaren Szene befestigt ist. Zunächst werden beide Lichtquellen fest positioniert, dann wird die Position des ColorCheckers im sichtbaren Bereich gewählt. Aus der Aufnahme dieser Szene ohne Schatten (Abb. 4 a) wird die Ground Truth Gradientenkarte ∇f berechnet. Der ColorChecker wird mit einer homogenen Fläche abgedeckt, dieses Bild f_h wird zur Berechnung von $S_e(x, y)$ und $S_r(x, y)$ benötigt. Das Schattenursachenobjekt wird so positioniert, dass der Schatten im Kamerabild sichtbar ist. Das Bild f_{hs} (Abb. 4 b) wird auch zur Berechnung von $S_e(x, y)$ und $S_r(x, y)$ verwendet. Die homogene Abdeckung wird entfernt, das Bild des ColorCheckers mit Schatten f_s (Abb. 4 c) dient als Eingabe für das Verfahren zur Schattenentfernung.

Zur Erzeugung von $S_e(x, y)$ und $S_r(x, y)$ (Abb. 4 d, e) wird zunächst eine Kantendetektion auf f_{hs} durchgeführt. Eine dreifache Dilatation gefolgt von einer Schließung liefert $S_e(x, y)$. Die Binarisierung der Differenz $f_h - f_{hs}$, Veroderung mit $S_e(x, y)$ und anschließende Schließung liefert ein Binärbild $S_s(x, y)$, das alle Schattenpixel repräsentiert. Die Exklusiv-Veroderung von $S_s(x, y)$ und $S_e(x, y)$ liefert die Schattenpixel innerhalb der Schattengrenzen $S_r(x, y)$.

Abb. 4 zeigt den Vergleich der Ergebnisse mit und ohne Einbeziehung des Hue-Kanals bei der Schattenkantendetektion (2.2). In Tabelle 1 sind die zugehörigen mittleren Gradientenabstände aufgelistet. Dass innerhalb und besonders außerhalb des Schattenbereichs bei Verwendung der fehlerhaften Schattenkanteninformation (Abb. 4 i) starke Artefakte entstehen, zeigt sich jeweils deutlich im Vergleich der mittleren Gradientenabstände. Im Bereich der Schattenkante unterscheiden sich die Ergebnisse kaum, der etwas höhere Wert beim

Ergebnis ohne Einbeziehung des Hue-Kanals ist wohl auch auf die Artefakte zurückzuführen.

Das Maß des mittleren Gradientenabstands im Schattenkantenbereichs liefert im

Tabelle 1. Mittlere Gradientenabstände der Ergebnisse mit und ohne Einbeziehen des Hue Kanals bei der Schattenkantendetektion

Gradientenabstand	Verwendung des Hue-Kanals	Keine Verwendung des Hue-Kanals
d_e	6.306	6.601
d_i	8.040	9.163
d_o	3.814	5.074

Beispiel also keine Aussage über die Güte der eigentlichen Schattenentfernung, sondern schließt auch die Bewertung der Entstehung von Artefakten mit ein. Um nur Schattenentfernung zu bewerten, ist z.B. denkbar, als Eingabe ein Bild mit einem Schatten auf homogenem Grund zu verwenden (Abb. 4 b). Allerdings hängt bei den Verfahren zur Schattenentfernung meist von der Komplexität der Szene ab, wie gut ein Schatten entfernt werden kann.

Die vorgestellten Maße stellen einen ersten Ansatz zur quantitativen Bewertung von Verfahren zur Schattenentfernung dar, die Auswertung aller Laboraufnahmen, um die Zuverlässigkeit der Maße zu bestimmen, steht noch aus.

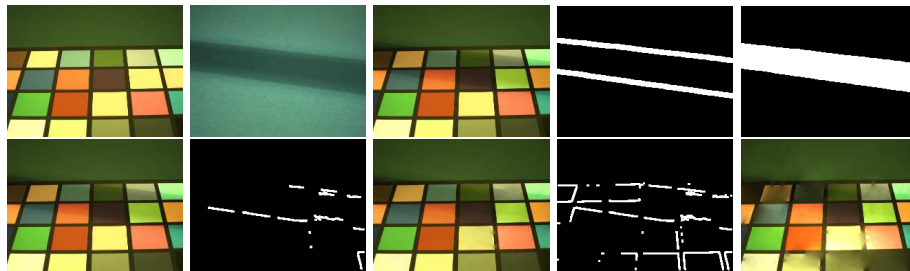


Abb. 4. a: Ground Truth Szene, b: Schatten auf homogenem Grund, c: Szene mit Schatten, d: Ground Truth Schattenkante, e: Ground Truth Schattenregion, f: Eingabebild, g: Schattenkantendetektion mit Einbeziehung des Hue-Kanals, h: Retinex-Schattenentfernung unter Eingabe von f und g, i: Schattenkantendetektion ohne Einbeziehung des Hue-Kanals, j: Retinex-Schattenentfernung unter Eingabe von f und i

6 Zusammenfassung

Bei der Umsetzung der vorgestellten Verfahren haben sich verschiedene Schwierigkeiten ergeben. Bei der Schattenkantendetektion besteht vor allem das Problem, dass in der invarianten Darstellung Materialkanten fehlen, wenn die Farben

benachbarter Flächen auf ähnliche invariante Werte abgebildet werden. Durch Einbeziehen des Hue-Kanals können einige falsch klassifizierte Schattenpixel verworfen werden, unter der Annahme, dass sich die Farben der Lichtquellen nicht stark unterscheiden. Experimente haben ergeben, dass die benötigte Kamera-Kalibrierung auch unter Einsatz künstlicher Lichtquellen durchgeführt werden kann. Auch die nichtlineare Kamera lieferte die erwarteten Ergebnisse, aufgrund der starken JPEG-Artefakte waren die invarianten Bilder allerdings nicht verwendbar. Die Schattenentfernung mit dem Retinex wurde für den Einsatz ohne Farbkorrektur angepasst, die Probleme, dass sich fehlerhafte Operationen an der Schattenkante global auswirken, bestehen weiterhin. Selbst wenn die Schattenkanteninformation korrekt ist, können Fehler entstehen, wenn Pfade im Schatten starten. Zur Bewertung von Verfahren zur Schattenentfernung wurde ein erster Ansatz vorgestellt. Die benötigten Ground Truth Daten lassen sich im Labor erzeugen, allerdings funktioniert die automatische Erzeugung der Daten nur, wenn ausreichend scharfe Schattenkanten vorliegen. Von der Auswertung weiterer Experimente erwarten wir eine Bestätigung der Ergebnisse und der Zuverlässigkeit der Mae.

Literatur

1. Finlayson, G.D., Hordley, S., Drew, M.S.: Removing shadows from images using retinex. In: Color Imaging Conference, Scottsdale, Arizona (2002) 73–79
2. Finlayson, G.D., Hordley, S., Drew, M.S.: Removing shadows from images. In: European Conference on Computer Vision (ECCV 2002), Copenhagen (2002) IV 823–836 Lecture Notes in Computer Science Vol. 2353.
3. Finlayson, G.D., Hordley, S.: Color constancy at a pixel. *Optical Society of America Journal A* **18** (2001) 253–264
4. Land, E., McCann, J.: Lightness and retinex theory. *Journal of the Optical Society of America* **61** (1971) 1–11
5. Finlayson, G.D., Hordley, S., Lu, C., Drew, M.S.: On the removal of shadows from images. *Pattern Analysis and Machine Intelligence* **28** (2006)
6. Wyszecki, G., Stiles, W.S.: *Color Science: Concepts and Methods, Quantitative Data and Formulas*. second edn. Wiley (1982)
7. Cucchiara, R., Grana, C., Prati, A., Piccardi, M.: Detecting objects, shadows and ghosts in video streams by exploiting color and motion information. In: *Image Analysis and Processing*, 2001. Proceedings. 11th International Conference on. (2001) 360–365
8. Jobson, D., Rahman, Z., Woodell, G.: A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image Processing* **6** (1997) 965–976
9. Marini, D., Rizzi, A.: A computational approach to color adaptation effects. *Image and Vision Computing* **18** (2000) 1005–1014
10. Land, E.: Recent advances in retinex theory and some implications for cortical computations: color vision and the natural image. *Proceedings of the National Academy of Sciences of the United States of America* **80** (1983) 5163–5169
11. Fredembach, C., Finlayson, G.D.: Hamiltonian path based shadow removal. In: *Proceedings of 16th British Machine Vision Conference (BMVC)*. (2005) 970–980

12. Prati, A., Mikic, I., Trivedi, M., Cucchiara, R.: Detecting moving shadows: Algorithms and evaluation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25** (2003) 918–923