

Methoden des automatischen Weißabgleichs und ihr Bezug zur Farbkonstanz

Rainer Jahn, Karl-Heinz Franke

Zentrum für Bild- und Signalverarbeitung e. V.

Gustav-Kirchhoff-Str. 5, D-98693 Ilmenau

E-Mail: rainer.jahn@zbs-ilmenau.de

Web: <http://www.zbs-ilmenau.de>

Zusammenfassung: Das menschliche visuelle System klassifiziert die Farbe eines Objektes im Wesentlichen nach seinen Remissionseigenschaften; der Einfluss der Beleuchtung wird eliminiert. Damit die Bilder der Objekte in den richtigen Farben erscheinen, muss der Beleuchtungseinfluss auch aus den Bildern eliminiert werden. Dies erfolgt in Digitalkameras durch den automatischen Weißabgleich (AWB). Ausgewählte AWB-Verfahren, die dem Stand der Technik entsprechen, werden kurz beschrieben. Die Ergebnisse eigener Arbeiten zum AWB werden an Beispielen demonstriert.

1 Farbkonstanz

Unter „Farbkonstanz“ versteht man die Fähigkeit, Objekte unabhängig von den Beleuchtungsbedingungen nach ihrer Farbe klassifizieren zu können. Das menschliche visuelle System ist also in der Lage, den Einfluss der Beleuchtung auf die Farbwahrnehmung weitgehend zu neutralisieren. Die wahrgenommene Farbe beruht fast nur auf dem spektralen Verhalten der remittierenden Pigmente.

Diese Fähigkeit lässt sich durch die so genannte Retinex-Theorie von E. Land u. a. erklären (siehe [2]). Die Theorie setzt die Existenz von drei visuellen Kanälen voraus, die auf unterschiedliche Spektralbereiche sensibilisiert sind. Sie postuliert, dass die Farbempfindung von den relativen Helligkeiten bestimmt wird, die über die drei Kanäle empfangen werden. Nach der Theorie wird die relative Helligkeit R in einem Punkt B durch die sensorischen Reize G_j entlang von Pfaden bestimmt, die in B enden (siehe Abb. 1). Die mathematische Formulierung lautet gemäß [1]:

$$\log R = \frac{1}{N} \sum_{k=1}^N r_k \quad (N \text{ Pfade}) \quad (1)$$

mit

$$r_k = \sum_{j=0}^n \delta \cdot \left| \log \frac{G_j}{G_{j-1}} \right|, \quad \delta = \begin{cases} 1 & \text{wenn } \left| \log \frac{G_j}{G_{j-1}} \right| > \gamma, \\ 0 & \text{sonst.} \end{cases} \quad (2)$$

Hierbei ist γ ein Schwellwert nahe 0. Die relative Helligkeit ergibt sich also durch schrittweise Multiplikation der Reizverhältnisse von jeweils benachbarten Punkten entlang eines Pfades. Verhältnisse nahe 1 werden im Produkt weggelassen, da sie laut [2] ausschließlich von Inhomogenitäten der Beleuchtung herrühren. Die Reizmaße G_j sind Relativwerte, und ein Pfad beginnt immer mit einem festen Bezugswert, z. B. mit $G_0 =$

100. Wenn beim Abtasten eines Pfades das bisherige Teilprodukt größer als 100 wird, setzt man den aktuellen Reizwert auf 100 zurück und verwirft den Anfang des Pfades bis zu dieser Stelle.¹ Dadurch wird der Bezugswert sukzessive Punkten mit immer größerem sensorischen Reiz zugeordnet, bis er mit hoher Wahrscheinlichkeit einem Punkt zugeordnet ist, der den Maximalreiz des visuellen Kanals liefert. Aufgrund der Mittelung über viele Pfade gemäß (1) wird die Zuverlässigkeit dieses Prozesses erhöht.

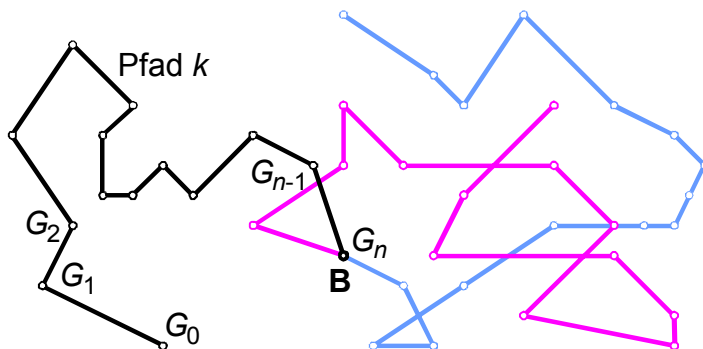


Abb. 1 Pfade in der Bildebene (Netzhaut), basierend auf der Retinex-Theorie.

Aus der Retinex-Theorie folgt, dass die Farbwahrnehmung von Inhomogenitäten der Beleuchtung unabhängig ist. Sie erklärt ferner, dass die wahrgenommene Farbe auf den relativen Reizen der sensorischen Kanäle beruht, die auf den Maximalreiz des Kanals normiert sind. Das macht die Unabhängigkeit der Farbempfindung vom Beleuchtungsspektrum in einem gewissen Maße plausibel.

Im Widerspruch dazu steht die Erfahrung mit metameren (bedingt-gleichen) Farben, die bei einer Beleuchtungsart gleich, bei einer anderen Beleuchtungsart jedoch verschieden erscheinen. Wenn man im Extremfall monochromatisch beleuchtet, erscheinen alle Objekte im gleichen Farbton, und die Dimension unterscheidbarer Farbklassen reduziert sich auf die Helligkeit.

Ob dieser Effekt auch bei alltäglichen Beleuchtungsverhältnissen relevant ist, wurde anhand eines standardisierten Satzes von Remissionsspektren (Munsell-Farbkatalog) untersucht. Für alle Proben des Satzes wurden die Farbvalenzen bei den Lichtarten D65 (Tageslicht) und F2 (Leuchtstofflampe) berechnet. Die Lichtarten wurden so ausgewählt, dass sich ihre Spektren und Eigenfarben möglichst stark unterschieden. Der paarweise Vergleich der Ergebnisse hat ergeben, dass fast alle Probenpaare entweder bei beiden Beleuchtungsarten farblich unterscheidbar oder bei beiden Beleuchtungsarten nicht unterscheidbar sind.

Die größte Abweichung von dieser Regel ist in Abb. 2 dargestellt. Man erkennt, dass sich die betreffenden Proben nur bei D65 unterscheiden. Der berechnete Farbabstand im CIELAB-Raum beträgt 1,7 bei F2 und 7,4 bei D65. Wenn man voraussetzt, dass die untersuchten Remissionsspektren repräsentativ und umfassend sind, kann man folgende Schlussfolgerung ziehen: Der Metamerie-Effekt ist in alltäglichen Situationen vernachlässigbar und beeinträchtigt die Fähigkeit der Farbkonstanz nicht wesentlich.²

¹ Das Rücksetzen und Verwerfen ist nicht in den Formeln (1) und (2) ersichtlich.

² Beim Ausbessern von Lackierungen können solche Farbnuancen jedoch entscheidend sein; das ist also in diesem Sinne keine alltägliche Situation.

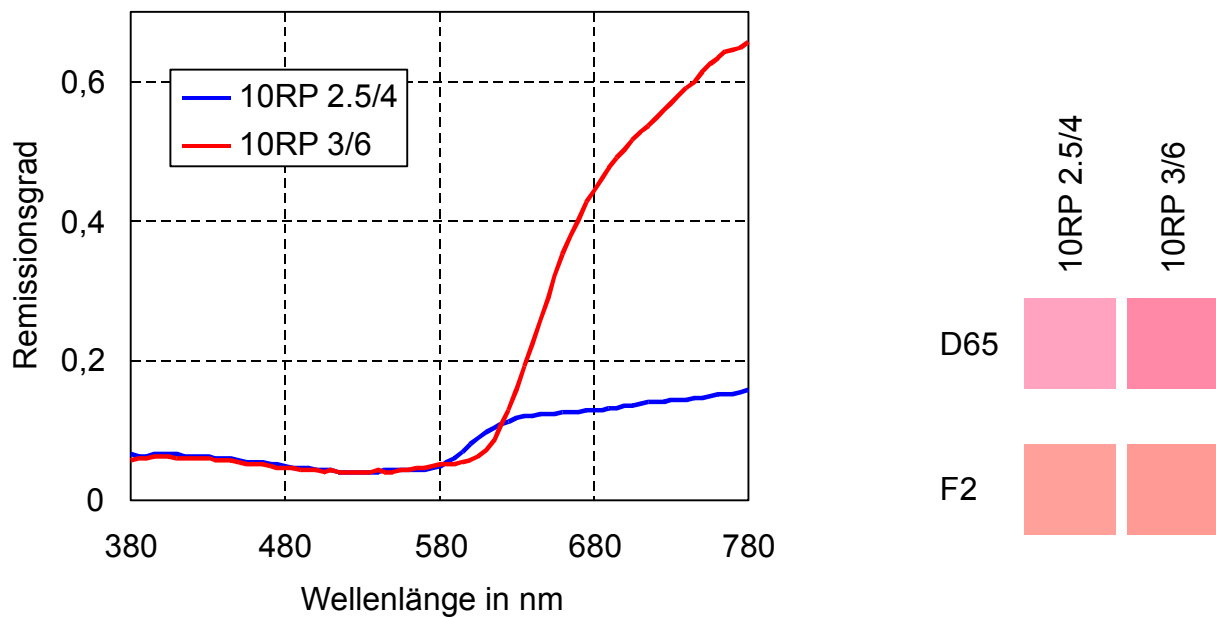


Abb. 2: Der größte Metamerie-Effekt, der anhand des Munsell-Farbkatalogs gefunden wurde, ergibt sich für das Probenpaar 10RP 2.5/4 + 10RP 3/6. Links ist das Remissionsspektrum der Proben dargestellt, rechts der annähernde Farbeindruck bei Beleuchtung mit D65 bzw. F2.

2 Automatischer Weißabgleich (AWB) im Überblick

Ein Fotoempfänger registriert die aktuellen, vom Beleuchtungsspektrum abhängigen Körperfarben. Ohne eine angemessene Korrektur wären die Farben des resultierenden Bildes in Richtung der Lichtquellenfarbe verschoben. Das heißt, die Remissionsspektren von Papierbild und Original-Objekt würden sich derart unterscheiden, dass der Mechanismus der Farbkonstanz für Objekt und Bild unterschiedliche Farben klassifizieren würde. Denn wie oben dargelegt, basiert die Farbklassifikation im Wesentlichen auf dem Remissionsspektrum. Gleiches gilt, wenn man die Remission des Papierbildes durch die Emission eines Monitorbildes ersetzt.

Die Aufgabe des Weißabgleichs besteht darin, das Spektrum des Bildes so zu verändern, dass der Betrachter im Bild die gleichen Farben klassifiziert, die er in der Original-Szene erwarten würde. Erfahrungsgemäß ist es nicht notwendig, das Spektrum des Objekts exakt nachzubilden. Stattdessen genügt der Abgleich in wenigen breitbandigen Kanälen, also letztlich eine Farbtransformation.

Bei herkömmlichem Fotomaterial erfolgt der Weißabgleich auf verschiedene Weise manuell (Fotolabor, Konversionsfilter, Kunstlichtfilm). Digitalkameras lassen mittels Zugriff auf das Rohdatenbild durch kamerainterne Prozessoren einen flexiblen Weißabgleich zu, der auch automatisierbar ist. Die Methoden des automatischen Weißabgleichs basieren u. a. auf folgenden Grundideen:

Graue-Welt-Hypothese. Nach dieser Annahme erscheint eine beobachtete Szene im Mittel unbunt. Im einfachsten Fall sind also die aufgenommenen Farbvalenzen so zu transformieren, dass das arithmetische Mittel einen Grauton ergibt. Bessere Verfahren verwenden gewichtete Mittelwerte, in die weitere Annahmen zur Farbzusammensetzung der Szene einfließen.

Existenz idealer Reflektoren. Es wird angenommen, dass sich die hellsten und am wenigsten gesättigten Bildbereiche wie wellenlängen-neutrale Reflektoren verhalten und vom Beobachter als grau klassifiziert werden. Die Retinex-Theorie basiert implizit auf einer Variante dieser Annahme. Dort werden jedoch die hellsten Stellen in jedem Kanal getrennt ermittelt und müssen nicht zwangsläufig zusammenfallen.

Gamut-Mapping beruht auf der Tatsache, dass alle Farbvalenzen nach dem Weißabgleich in einer wohldefinierten konvexen 3D-Punktmenge liegen müssen. Diese Punktmenge ist ein Farbkörper („Gamut“) und besteht aus allen Körperfarben, die bei einer farbneutralen Referenzbeleuchtung möglich sind. Beim Gamut-Mapping werden die Farbvalenzen des Bildes so transformiert, dass ihre Gesamtheit den erwähnten Farbkörper weitgehend ausfüllt. Ein konkretes Verfahren ist in [8] beschrieben.

Physiologische Aspekte. Einige Verfahren verwenden physiologische Erkenntnisse zur Farbverarbeitung im menschlichen Sehsystem. In [4] wird z. B. vorgeschlagen, zunächst die Farbvalenzen eines Bildes in einen physiologisch motivierten Farbraum zu transformieren. Dort erfolgt der Weißabgleich durch eine Transformation, bei der die Hauptachse der Punktmenge in die Unbuntgerade überführt wird.

Im allgemeinen Fall erfolgt der Weißabgleich in Digitalkameras durch lineare Farbtransformation des Rohdatenbildes. Durch die Wirkung der Bayer-Maske liefert jedes 2x2 Pixel große Feld einen roten, einen blauen und zwei grüne Farbwerte, die sich zur Farbvalenz $(R_0, G_0, B_0)^T$ eines „Makropixels“ zusammenfassen lassen. G_0 ist z. B. der Mittelwert beider Grünwerte. Die entsprechende Farbvalenz $(R_1, G_1, B_1)^T$ im abgeglichenen Bild ergibt sich allgemein nach der Vorschrift

$$\begin{pmatrix} R_1 \\ G_1 \\ B_1 \end{pmatrix} = \underline{\underline{M}} \cdot \begin{pmatrix} R_0 \\ G_0 \\ B_0 \end{pmatrix} \quad (3)$$

mit einer 3x3-Matrix $\underline{\underline{M}}$, die die Beleuchtungsverhältnisse charakterisiert. In der Praxis wird die Matrix meist auf die Diagonalform reduziert, was einer getrennten Skalierung der Farbkanäle entspricht. Da es nur auf den Abgleich des Farbtone ankommt und die Bildhelligkeit nach anderen Kriterien eingestellt wird, kann man sich sogar auf zwei unabhängige Skalierungsfaktoren beschränken. Damit vereinfacht sich die Matrix zu

$$\underline{\underline{M}} = \begin{pmatrix} R_{gain} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & B_{gain} \end{pmatrix} \quad (4)$$

mit der „Rotverstärkung“ R_{gain} und der „Blauverstärkung“ B_{gain} .

Man kann die Verstärkungswerte auch allgemeiner für eine beliebige Farbvalenz $(R_0, G_0, B_0)^T$ definieren:

$$R_{gain} = \frac{G_0}{R_0}, \quad B_{gain} = \frac{G_0}{B_0}. \quad (5)$$

Diese Größen lassen sich ähnlich wie die Normfarbwertanteile in einer Farbtabelle darstellen, die wir als Verstärkungswerte-Ebene bezeichnen. Außerdem geben sie Skalierungsfaktoren an, mit denen man den Rot- und Blauwert multiplizieren müsste, um einen

Grauton zu erhalten. Der Farbort der (vermutlichen) Lichtquelle, die bei der Bildaufnahme dominiert hat, liefert also in diesem System direkt die günstigsten Skalierungsfaktoren für den Weißabgleich. Daher wird die Verstärkungswerte-Ebene (siehe Abb. 3) bei der Behandlung des Weißabgleichs gegenüber der Normfarbtafel bevorzugt. Die Verstärkungswerte basieren auf dem RGB-Farbraum eines speziellen Fotosensors und sind daher geräteabhängig.

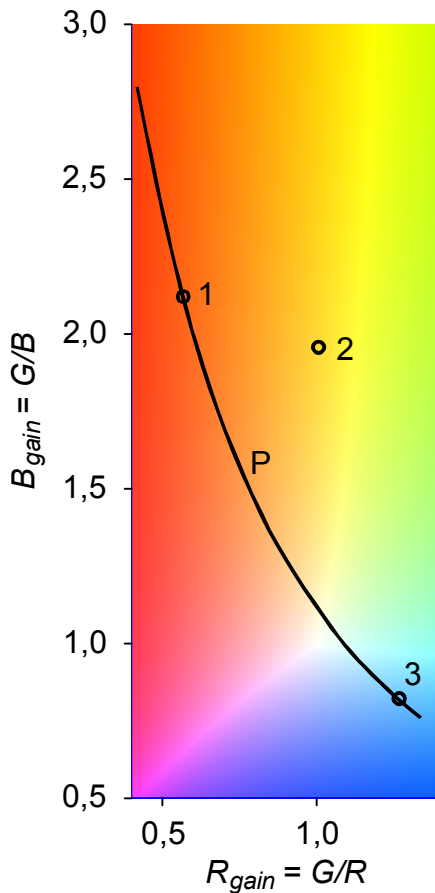


Abb. 3: Verstärkungswerte-Ebene mit den Farborten ausgezeichneten Lichtquellen, basierend auf dem Kodak-Sensor KAF-10010CE.
P – Ortskurve der Schwarzen Strahler im Temperaturbereich 2000...10000 K, 1 – Glühlampe, 2 – Leuchtstofflampe, 3 – Tageslicht bei Sonnenschein.

3 Ausgewählte AWB-Verfahren

Viele Erfindungen zum automatischen Weißabgleich beschreiben komplexe Methoden zur Bestimmung der o. g. Verstärkungswerte. Sie stützen sich entweder allein auf das Rohdatenbild oder verwenden zusätzlich Informationen aus anderen Bestandteilen der Kamera-Hardware.

3.1 Verfahren, die nur auf dem Rohdatenbild beruhen

Beispiel 1. In [5] wird folgender Cluster-Algorithmus beschrieben, der auf der Graue-Welt-Hypothese beruht und Ausreißer unter den Farbvalenzen eliminiert:

1. Das Rohdatenbild wird in Blöcke unterteilt (z. B. 16 x 16).
2. Die mittleren Verstärkungswerte der Blöcke werden als Punkte in der Verstärkungswerte-Ebene interpretiert.
3. Punkte, deren Abstand kleiner als ein Schwellwert ist, werden zu Häufungsmengen zusammengefasst. Häufungsmengen mit wenigen Elementen werden verworfen.

4. In geeigneter Weise werden „Mittelpunkte“ der verbliebenen Häufungsmengen bestimmt.
5. Das gewichtete Mittel der „Mittelpunkte“ definiert die Weißabgleichsfaktoren. Die Wichtung erfolgt nach der Elementezahl der Häufungsmenge.

Beispiel 2. In [7] werden Bildblöcke identifiziert, denen man ideale Remissionseigenschaften zuschreibt. Der sehr differenzierte Algorithmus ist in Abb. 4 vereinfacht dargestellt. Eine Block-Bewertungseinheit (1) teilt die Bildblöcke in drei Klassen ein:

1. die hellsten Blöcke,
2. Blöcke, die das Licht von Leuchtstofflampen ideal remittieren,
3. Blöcke, die das Licht von Temperaturstrahlern, also Tages- oder Glühlampenlicht, ideal remittieren.

Die Zuordnung eines Blocks zur Klasse 2 oder 3 richtet sich nach der Lage der repräsentativen Farbvalenz des Blocks in einer Ebene von Farbwertdifferenzen. Beide Lichtquellenarten sind durch je einen zusammenhängenden Bereich in dieser Ebene definiert. Die Zuordnung zur Klasse 1 wird durch den repräsentativen Luminanzwert des Blocks bestimmt.

Für jede Klasse wird eine mittlere Farbvalenz berechnet (Komponenten 2, 3, 9). Diesen Mittelwerten werden nach verschiedenen empirischen Kriterien, z. B. der Sättigung, Wichtungsfaktoren zugeordnet (Komponenten 4, 5, 10). Dadurch lassen sich Fehldeutungen vermeiden, damit z. B. eine Wiese im Sonnenlicht nicht als weiße Fläche unter einer Leuchtstofflampe interpretiert wird. Aus den Mittelwerten und Wichtungsfaktoren der Klassen wird eine globale mittlere Farbvalenz bestimmt und als Lichtquellenfarbe interpretiert. Der Weißabgleich erfolgt über die Rot- und Blauverstärkung gemäß (5), wobei die RGB-Werte der ermittelten Lichtquellenfarbe einzusetzen sind.

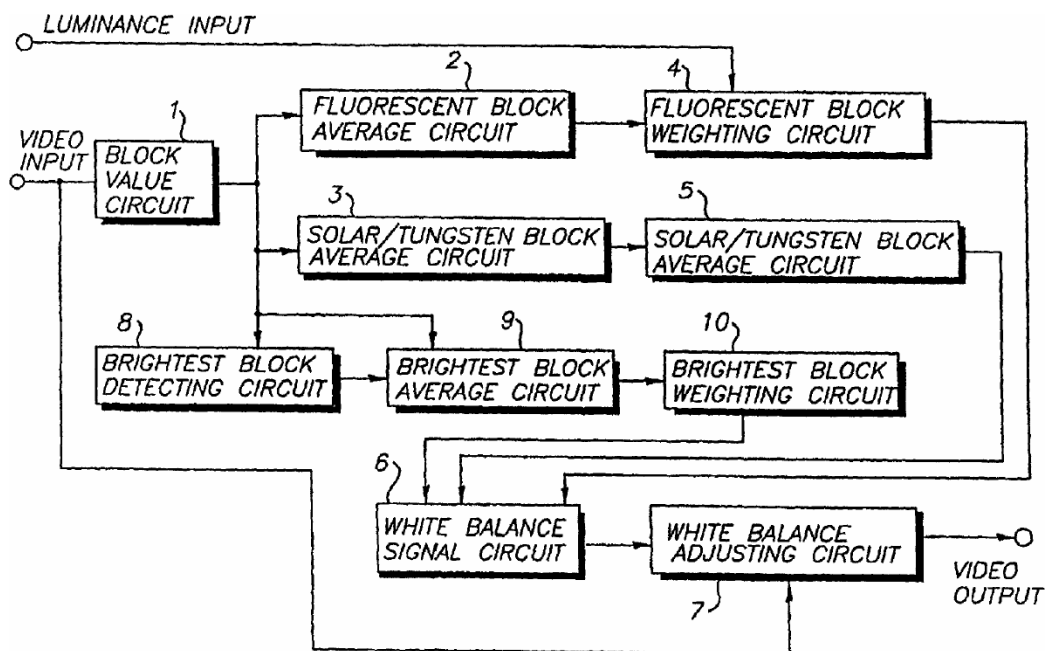


Abb. 4 Zeichnung aus der Patentschrift [7].

3.2 Verfahren, die Zusatzinformationen auswerten

Viele Erfindungen benutzen einen zusätzlichen Farbsensor für das Umgebungslicht. Theoretisch könnte man allein aus dem Signal dieses Sensors die Weißabgleichsfaktoren so berechnen, dass sie die Graue-Welt-Hypothese erfüllen. Im Allgemeinen dient der Sensor aber nur zur Ergänzung der Bildinformationen, eventuell zusammen mit weiteren Informationsquellen (Blitzlicht, Objektstand). Der Farbsensor lohnt sich nur, wenn er einen wesentlich größeren Ausschnitt des Objektraumes erfasst als der Bildempfänger.

Beispiel 3. In [6] ist eine besondere Ausführungsform des Sensors beschrieben, siehe Abb. 5. Über einen Lichtleiter 7 wird das Umgebungslicht auf eine Ecke oder Kante des Bildempfängers 4 gelenkt. Diese Teilfläche ist dann nicht mehr zur Aufnahme von Bildinformationen verwendbar. Sie erscheint jedoch nach dem Auslesen zunächst als Bestandteil des Bildsignals und muss elektronisch vom Bildrahmen getrennt werden. Da lediglich eine integrale Farbvalenz für das Umgebungslicht zu ermitteln ist, kann die erwähnte Teilfläche sehr klein sein.

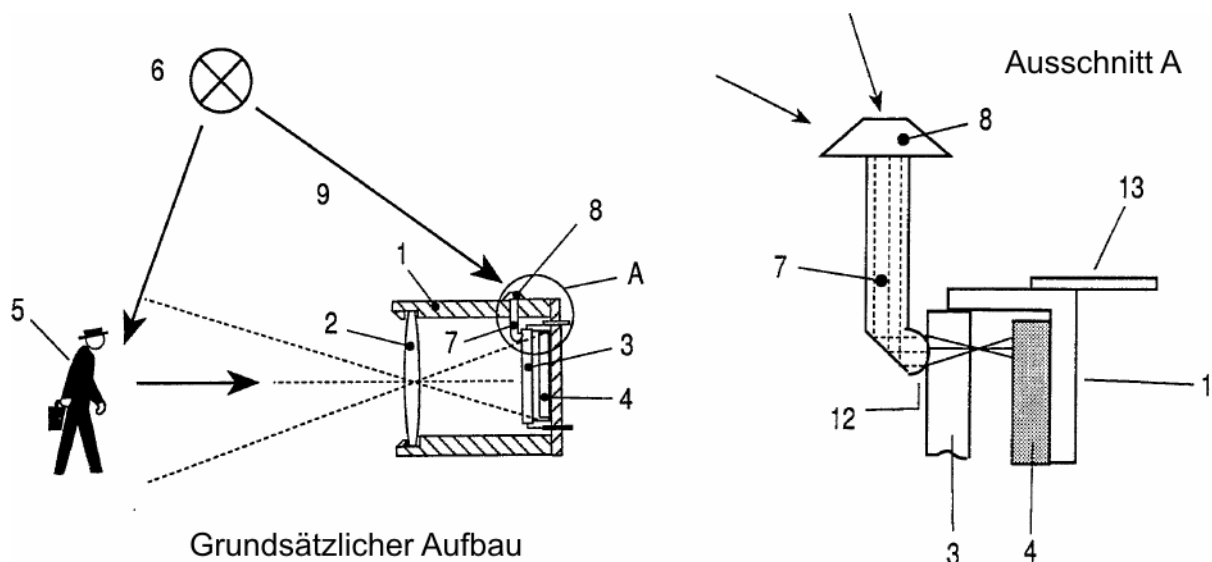


Abb. 5 Modifizierte Zeichnung aus der Patentschrift [6].

Beispiel 4. Weitere Erfindungen beschäftigen sich mit dem Einfluss des Blitzlichtes und nutzen Informationen, die a priori über das Blitzlicht verfügbar sind. Die naheliegende Lösung besteht darin, auf einen gewichteten Mittelwert zwischen der Farbe des Umgebungslichtes und der bekannten Farbe des Blitzlichtes abzugleichen.

Das Patent [3] verwendet dafür ein nichtlineares Interpolationsverfahren, das im Folgenden leicht verallgemeinert beschrieben wird. Zunächst werden aus einer Vorbelichtung ohne Blitz die mittlere Luminanz der Objekte sowie die Verstärkungswerte R_{gain}^a , B_{gain}^a für das Umgebungslicht ermittelt. Letztere lassen sich z. B. mit einem Graue-Welt-Verfahren allein aus dem Rohdatenbild schätzen. Weiter wird aus der mittleren Luminanz und einer Erfahrungs-LUT der Beleuchtungsanteil k des Blitzlichtes bestimmt, der nach Einschalten des Blitzes zu erwarten ist; k liegt zwischen 0 (kein Blitz) und 1 (Blitz dominiert). Die Farbe des Blitzlichtes sei durch die Verstärkungswerte R_{gain}^f , B_{gain}^f gegeben. Bekannt sei auch die Ortskurve der Temperaturstrahler in der Verstärkungswerte-

Ebene (vgl. Abb. 3), die sich als Funktionsgleichung $B_{gain} = S(R_{gain})$ schreiben lässt. Die endgültigen Weißabgleichsfaktoren werden nach folgender Formel bestimmt:

$$\begin{aligned} R_{gain} &= R_{gain}^a + k \cdot (R_{gain}^f - R_{gain}^a) \\ B_{gain} &= B_{gain}^a + S(R_{gain}) - S(R_{gain}^a) + k \cdot (B_{gain}^f - S(R_{gain}^f) - B_{gain}^a + S(R_{gain}^a)) \end{aligned} \quad (6)$$

Abb. 6 zeigt die Wertepaare (6) in Abhängigkeit von k als Ortskurve in der Verstärkungswerte-Ebene (rot gezeichnet). Diese Kurve ist so beschaffen, dass sich ihr vertikal gemessener Abstand zur Ortskurve der Schwarzen Strahler (blau gezeichnet) linear mit R_{gain} ändert. Dadurch wird erreicht, dass der interpolierte Punkt näher an die Ortskurve der Schwarzen Strahler rückt als bei linearer Interpolation. Lichtquellen in der Umgebung Schwarzer Strahler erhalten also ein höheres Gewicht. Exemplarisch erhält man für einen Blitzlichtanteil von $k = 0,45$ den interpolierten Punkt P in Abb. 6.

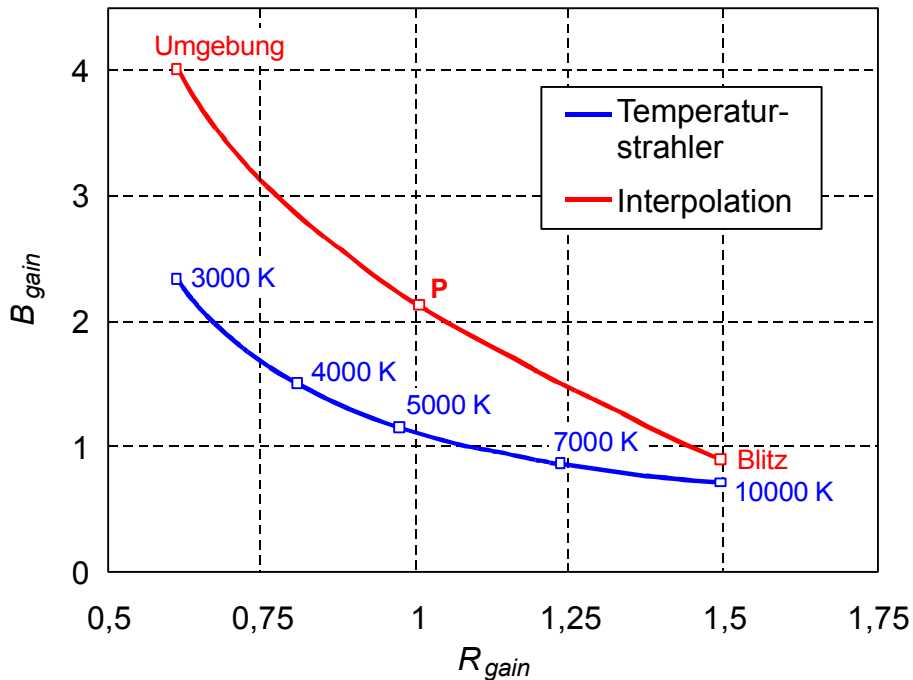


Abb. 6 Zum Interpolationsverfahren nach [3].

4 Ergebnisse

Am ZBS wurde ein kombiniertes AWB-Verfahren für hochwertige Digitalkameras entwickelt. Der Kern-Algorithmus basiert im Wesentlichen auf der Graue-Welt-Hypothese und verwendet außer dem Rohdatenbild keine weiteren bildspezifischen Eingangsdaten. Die Bestimmung der Weißabgleichsfaktoren beruht auf einer Analyse der Bildpunkte in der Verstärkungswerte-Ebene. Eine nähere Beschreibung des Algorithmus ist an dieser Stelle nicht möglich.

Das Verfahren wurde an Rohbildmaterial der Zielkamera getestet. Bei kritischen Objektszenen lieferte es oft bessere Ergebnisse als die AWB-Systeme der Konkurrenz-Produkte. Dies sei anhand von zwei Testbildern in Abb. 7 und Abb. 8 verdeutlicht. Oben links ist jeweils als angestrebtes Optimum das Ergebnis des manuellen Weißabgleichs mittels Graukarte dargestellt, rechts daneben das Resultat unseres AWB-Verfahrens.

Darunter sind Aufnahmen des gleichen Objektes mit kommerziellen Kameras bei aktiviertem AWB zu sehen. Das Motiv „Putzwagen“ (Abb. 7) wurde unter Leuchtstofflampen aufgenommen, das Motiv „Testbox“ (Abb. 8) unter Glühlampen. Da die Eigenfarben dieser Lichtquellen einen deutlichen Buntton haben, ist der Weißabgleich hier unerlässlich. Die Konkurrenzprodukte haben offenbar Schwierigkeiten, solche Farbstiche zu identifizieren, wenn das Bildmotiv nicht die Graue-Welt-Hypothese erfüllt.

Es sei nicht verschwiegen, dass auch das AWB-Verfahren des ZBS in bestimmten Situationen versagt, z. B. bei Mischlicht, unbekannten Beleuchtungsarten oder bunten Bildmotiven ohne farbneutrale Bereiche. Bei derartigen Szenen stößt aber jedes AWB-Verfahren, das keine Zusatzinformationen verwendet, an seine objektiven Grenzen. Es gehört daher zum Stand der Technik, die Messwerte eines Farbsensors für das Umgebungslicht in den Weißabgleich einzubeziehen. Erste Versuche mit simulierten Sensordaten haben hervorragende Ergebnisse bei fast allen Objektszenen geliefert.

5 Literatur

- [1] Gadi, D. u. a.: Tuning Retinex for HDR image visualization, The Second European Conference on Colour Graphics, Imaging and Vision, 2004
- [2] Land, E. u. a.: Lightness and retinex theory, Journal of the Optical Society of America, 61 (1971), S. 1–11
- [3] Nakayama, S.: Image sensing apparatus, white balance adjusting method, and operation processing program therefor, Patent US 2003030730, Anmeldung 2002
- [4] Pomierski, T. u. a.: Neural architecture for chromatic adaption of single images and picture sequences, 2. Nationaler Workshop „Farbbildverarbeitung“, 1996, S. 29–34
- [5] Sato, G.: Automatic white balance adjusting method, Patent US 2004212691, Anmeldung 2003
- [6] Spruck, M.: Videokamera mit Weißabgleich, Patent DE 19514629 A1, Anmeldung 1995
- [7] Toshiki, M.: Method and apparatus for auto white balance, Patent EP 0738085 A2, Anmeldung 1996
- [8] Winter, K. A.: Automatic white balance detection and correction of an image captured in a digital camera, Patent EP 0862336 B1, Anmeldung 1998



AWB - Referenz 1



AWB - Referenz 2



Abb. 7 Weißabgleich am Bildmotiv „Putzwagen“. Oben links: manuell mit Graukarte, oben rechts: AWB mit eigenem Verfahren, unten: AWB mit Vergleichskameras.



AWB - Referenz 1



AWB - Referenz 2



Abb. 8 Weißabgleich am Bildmotiv „Testbox“. Oben links: manuell mit Graukarte, oben rechts: AWB mit eigenem Verfahren, unten: AWB mit Vergleichskameras.