

Abstract

Als Alternative zu herkömmlichen drahtlosen Nahbereich-Funktechnologien ist mit dem technischen Fortschritt der letzten Jahrzehnte die VLC (Visible Light Communication) und speziell die DCC (Display Camera Communication) in Forschung und Industrie populär geworden. Das System DaViD (Data Transmission Using Video Devices), welches am Lehrstuhl für Kommunikationstechnik der technischen Universität Dortmund entwickelt wird, hebt sich von der gewöhnlichen DCC dadurch ab, dass neben dem Datentransfer mit einer Kamera als Empfänger die gleichzeitige Videodarstellung für den Konsumenten ohne Qualitätsminderung möglich sein soll. Die dazu verwendete zeitlich-differentielle Chrominanzmodulation berücksichtigt bisher nicht die negativen Auswirkungen der nichtlinearen Bildschirmwiedergabe (Gradation). In der Arbeit werden diese analysiert, um neue Modulationsansätze vorzustellen, zu prüfen und nach einem Vergleich schließlich die optimale Methode zu ermitteln.

Inhaltsverzeichnis

Abstract	i
1 Einleitung	1
1.1 Das Projekt DaViD	1
1.2 Ziel der Arbeit	2
1.3 Aufbau	2
2 Grundlagen	3
2.1 Farbwahrnehmung und Farbmetrik	3
2.1.1 Das visuelle System des Menschen	3
2.1.2 Farbraum	5
2.2 Farbsysteme	7
2.2.1 RGB-Systeme	8
2.2.2 YUV	10
2.2.3 Empfindungsgemäße Farträume	12
2.2.4 CIE 1960 UCS	13
2.2.5 CIELUV (CIE 1976 L [*] u [*] v [*])	13
2.2.6 CIELAB (CIE 1976 L [*] a [*] b [*])	14
2.2.7 Farbabstände	15
2.2.8 HSV und Alternativen	16
2.3 Das Auflösungsvermögen des Menschen	17
2.3.1 Helligkeitsempfindung	17
2.3.2 Kontrast	18
2.3.3 Modulationsübertragungsfunktion	19
2.3.4 Zeitliche Auflösung	21
2.3.5 Chromatische Auflösung	21
2.4 Bildschirmtechnologie	22
2.4.1 Kathodenstrahlröhren (CRT)	22
2.4.2 Flachbildschirme	22
2.5 Gradation	23
2.5.1 Kennlinie von Kathodenstrahlröhren	23
2.5.2 Wiedergabefehler durch Gammaverzerrung	24
2.5.3 Gammakorrektur	25
2.5.4 Standards und Modelle	27
2.5.5 Constant Luminance	28

2.6 Bildqualität	30
2.6.1 Objektive Qualitätsmaße	30
2.6.2 Subjektive Tests	30
2.7 Stand der VLC-Technik	34
2.7.1 Systematische Einordnung	34
2.7.2 Nutzung der VLC	36
2.7.3 Verwandte Projekte	37
2.7.4 Systembeschreibung von DaViD	39
3 Modulationskonzepte	43
3.1 Motivation und Problemanalyse	43
3.2 Entwicklung und Beschreibung neuer Modulationskonzepte	46
3.2.1 Ziele und Qualitätskriterien	46
3.3 Grünausgleich	47
3.4 Amplituden-Korrektur (AC)	48
3.5 Signalbegrenzung	51
3.5.1 Aussteuerungsreserven	52
3.6 Bildabhängige Modulation	53
3.6.1 Schwellenwert-Amplitudenlimitierung (AL)	53
3.6.2 Übergang	53
3.6.3 Wertabhängige Amplitudenlimitierung (AL)	54
3.7 Lineare Dämpfung LD	54
3.8 Mischung	56
3.8.1 Partielle Amplitudenkorrektur (PAC)	56
3.8.2 Zusätzlicher Grünausgleich	56
3.9 Modulation in CIELAB (L*a*b*-M.)	56
3.10 Übersicht	58
3.11 Auswahl der Modulationsmethoden	60
4 Implementierung	62
4.1 MATLAB	62
4.2 CUDA	62
4.2.1 Code der neuen Modulationsverfahren	62
4.3 Modulierte Einzelbilder	63
5 Messungen	65
5.1 Übertragung und Decodierung	65
5.1.1 Messergebnisse und -auswertung der Übertragung	66
5.2 Luminanzmessung	70
5.2.1 Messergebnisse und -auswertung der Luminanzmessung	71

5.3 Zusammenfassung	72
6 Subjektiver Test	73
6.1 Testmethodik	73
6.2 Testmaterial	74
6.3 Testdurchführung	75
6.4 Testaufbau	76
6.5 Testergebnisse und -auswertung	77
6.6 Zusammenfassung	79
7 Zusammenfassung	80
A Anhang: Bewertungsskalen der ITU	81
B Anhang: Code-Auszug	82
C Anhang: modulierte Testbilder	85
D Anhang: Histogramme der Testsequenzen	86
Abkürzungsverzeichnis	87
Symbolverzeichnis	91
Abbildungsverzeichnis	93
Tabellenverzeichnis	94
Quellcodeverzeichnis	94
Literatur	95

1 Einleitung

Mit der zunehmenden Bedeutung von drahtloser Datenübertragung vor allem im Zuge von Internet of Things [1] und der Verbreitung mobiler Geräte wie Smartphones [2] stellen der wachsende Bedarf an Datenrate und Teilnehmerzahl bei begrenzter Bandbreite und Bedenken bezüglich Interferenzen und Sicherheit große Herausforderungen für die konventionelle Funktechnologie dar. Das Unternehmen Cisco gibt ein Wachstum des globalen mobilen Datenvolumens um 63 % für 2016 an und sagt eine Versiebenfachung zwischen 2016 und 2021 vorher [3].

Haas und die FCC (Federal Communications Commission) warnen vor einer „spectrum crisis“ in den nächsten Jahren [4], [5]. Aktuelle Forschungen an Alternativen wie der VLC (Visible Light Communication) bieten neuartige Lösungs- und Anwendungsmöglichkeiten. Hierzu werden Daten dem Licht im sichtbaren Spektrum beispielsweise von LED-Leuchtmitteln überlagert und am Empfänger mittels Lichtsensoren wiedergewonnen.

1.1 Das Projekt DaViD

Das vom Lehrstuhl für Kommunikationstechnik der TU Dortmund in den letzten Jahren vorangetriebene Forschungsprojekt DaViD (Data Transmission Using Video Devices) [6], [7] repräsentiert eine Sonderform der Nahbereich-DCC (Display Camera Communication). Die Methode zeichnet sich durch die simultane Übertragung von Video und darin versteckten Daten aus, die für den Betrachter möglichst unsichtbar moduliert, aber mithilfe Kameras aufgezeichnet und danach decodiert werden. Dabei werden handelsübliche Geräte wiederverwendet: z. B. Fernseh-, Werbe- oder Computer-Bildschirme zur Wiedergabe und Smartphones-Kameras.

Der bisherige Ansatz basiert auf einer zeitlich-differentiellen Modulation der Farbdifferenzen mit geringer Amplitude. Probleme ergeben sich unter anderem durch die nichtlineare Gradation der eingesetzten Bildschirme, insofern sie die sonst linearen Farbraum-Umrechnungen und die Modulation verzerrn. Dadurch wird unter anderem in helleren und gesättigten Bildsegmenten die Datenüberlagerung als Flimmern deutlich sichtbar, während die Detektion in dunklen Grautönen erschwert wird.

1.2 Ziel der Arbeit

Die in dieser Arbeit vorgestellten neuen Modulationsverfahren sollen die genannten Einflüsse der Gradation berücksichtigen, somit die Sichtbarkeit minimieren und gleichzeitig die Detektierbarkeit durch den Kameraempfänger erhöhen.

1.3 Aufbau

Im Kapitel 2 werden die relevanten theoretischen Grundlagen zur menschlichen Licht- bzw. Farbwahrnehmung, Farbsysteme, die Technologie verbreiteter Bildschirme mit der charakteristischen Gradation und der aktuelle Stand der VLC-Technik sowie des untersuchten Systems DaViD erörtert. Nach einer Evaluation der aktuellen Beeinträchtigungen durch die Gradation werden in Kapitel 3 verschiedene Lösungsvorschläge zur Modulation beschrieben und den Anforderungen entsprechend ausgewählt. Die Realisierung auf Software-Ebene ist in Kapitel 4 erläutert.

Die nächsten Kapitel 5 und 6 stellen die Durchführung der Messungen und des subjektiven Tests mitsamt der Ergebnisse dar. Auf Basis einer komparativen Auswertung wird schließlich das optimale Verfahren ermittelt. Die Zusammenfassung und ein Ausblick auf weitere Forschungsansätze sind im letzten Kapitel 7 zu finden.

2 Grundlagen

Zum Verständnis der vorgestellten Modulationsmethoden werden zunächst fundamentale, multidisziplinäre Theorien und Prinzipien der Farbwahrnehmung und darauf aufbauenden Farbmetrik erörtert sowie besonders das begrenzte Auflösungsvermögen des menschlichen visuellen Systems. Anschließende Abschnitte behandeln geläufige Bildschirm-Technologien und deren nichtlineare Gradation. Zur Evaluation der Bildqualität wird die Methodik subjektiver Tests vorgestellt. Nach einem Überblick über aktuelle VLC-Konzepte folgt zuletzt eine Beschreibung des Projekts DaViD inklusive dessen Modulationsverfahren.

2.1 Farbwahrnehmung und Farbmetrik

Die DIN 5033 zur Farbmessung definiert Farbe im Sinne der Farbempfindung¹ [8]. Allerdings hängt der Eindruck sowohl von den äußeren Umständen als auch vom visuellen Wahrnehmungsapparat des Betrachters ab [9].

In der Farbmetrik als „Lehre der Maßbeziehungen“ [10] wird versucht, Farben quantitativ durch Zahlenwerte zu erfassen [9]. Ein Modell der menschlichen Farbwahrnehmung soll einen Zusammenhang zwischen messbaren Farbreizen und gleichen Farbempfindungen herstellen [10], [11]. Die systematische und quantitative Darstellung aller Farben führt zu Farbsystemen und grafischen Modellen (siehe Abschnitt 2.2).

2.1.1 Das visuelle System des Menschen

Die Funktionsweise des Auges ist zwar seit vielen Jahrhunderten weitgehend erforscht, jedoch liefert die Neurophysiologie in den letzten Jahrzehnten weitere Erkenntnisse [10]. Einfallendes Licht, also elektro-magnetische Strahlung bzw. Photonen, wird im Auge durch eine Linsen- und Blendenoptik ähnlich einer Kamera umgekehrt auf die Netzhaut (Retina) abgebildet [12]. Die Pupillenöffnung reguliert die Adaption an Helligkeiten im Bereich von $10 \cdot 10^{-6} \text{ cd/m}^2$ bis $10 \cdot 10^5 \text{ cd/m}^2$ [10].

Wie Young 1802 erkannte, wandeln dort drei Typen von farbempfindlichen Photorezeptoren, den Zapfen, mit unterschiedlichen, überlappenden spektralen Absorptionskurven² (siehe Abb. 2.1) die Lichtenergie in Nervenimpulse um.

¹ also in Abgrenzung zu Körperfarben von Objekten

² Die Kurven können individuell variieren und sind durch Untersuchungen der Farbpigmente oder Abgleichsexperimente mit Farbenfehlsichtigen messbar [8]

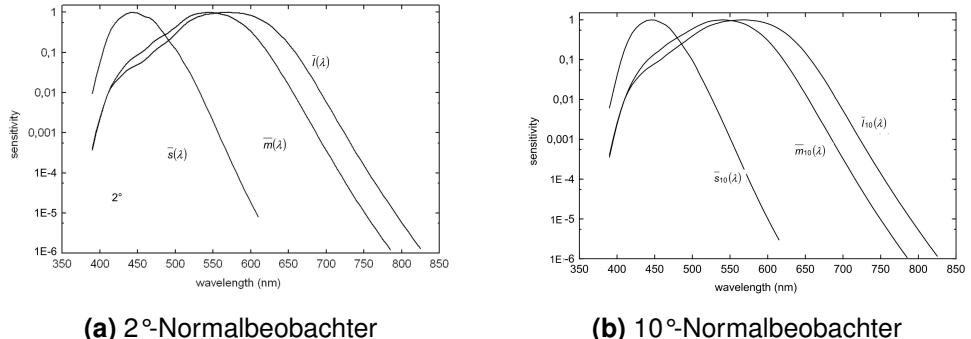


Abbildung 2.1: Empfindlichkeitspektrum der Augen-Rezeptoren $\bar{s}(\lambda)$ (*short*), $\bar{m}(\lambda)$ (*middle*) und $\bar{l}(\lambda)$ (*long*) (normalisiert auf maximal 1) für den 2°- und 10°-Normalbeobachter [13]

Die Wellenlängenbereiche der Empfindlichkeitsmaxima liegen bei ca. 430 nm (Blauviolett), 530 nm (Grün) und 565 nm (Gelb) [10] und geben den Zapfen die gebräuchlichen Namen S, M und L für *short*, *middle* und *long* [8]. Zu identifizieren ist an den Rändern ebenfalls der sichtbare Bereich von 380 nm bis 780 nm nach DIN-Standard [8].

In Abgleichexperimenten sollen Beobachter die Farbe eines vorgegebenen Referenzlichtes durch die Überlagerung von einigen wenigen Farben reproduzieren [14] [15]. Auf Youngs These aufbauend konnte Helmholtz 1867 so die Drei-Farben-Theorie entwickeln, dass zur Erzeugung aller Zwischenfarbtöne drei Primärfarben ausreichen, die wie z. B. Rot, Grün und Blau voneinander unabhängig sind, d.h. nicht gegenseitig herstellbar. Die Farbwahrnehmung entspricht folglich dem Modell der additiven Mischung (s. Abb. 2.2): Zeitlich oder räumlich benachbarte Farbreize überlagern sich im Auge zu einem Farbeindruck mit zunehmender Helligkeit beim Hinzufügen weiterer Farben [8].

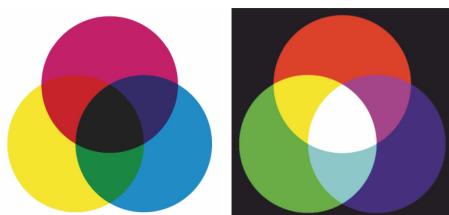


Abbildung 2.2: Vergleich: subtraktive (links) und additive Farbmischung (rechts). Während Rot, Grün und Blau (RGB (Rot-Grün-Blau-Farbraum)) in der additiven Mischung Weiß bilden, erzeugen die in der Druckindustrie üblichen Grundfarben Gelb, Cyan und Magenta der Farbmodelle CMY und CMYK gemischt Schwarz³⁾ [8].

2.1.2 Farbraum

Den Zusammenhang von Farbe zu Farbreiz fasste Grassmann 1853 in drei Gesetzen zusammen [10]:

1. Jeder Farbreiz kann auch ohne genaue Kenntnis des Reizspektrums eindeutig als eine vektorielle Linearkombination von drei Primärvalenzen im Farbraum dargestellt werden. Drei Farbwerte R , G und B bestimmen die Mischanteile und damit die Koordinaten:

$$\vec{F} = R\vec{R} + G\vec{G} + B\vec{B} \quad (2.1)$$

2. Farbreize mit gleicher Farbvalenz erzeugen unabhängig von ihrer spektralen Zusammensetzung identische Wirkungen in der additiven Mischung und werden metamer genannt.
3. Zwischen den Farben gibt es stetige Übergänge (siehe Gleichung (2.2)) [8].

Punkt 3 wird durch die Stetigkeit der Spektralwertfunktionen erreicht. Diese ergeben sich durch Messungen der Farbwerte für monochromatisches Licht⁴ und sie legen die Grundvalenzen fest:

$$R = \int_{\lambda_1}^{\lambda_2} \varphi(\lambda) \bar{r}(\lambda) d\lambda \quad (2.2)$$

$$G = \int_{\lambda_1}^{\lambda_2} \varphi(\lambda) \bar{g}(\lambda) d\lambda \quad (2.3)$$

$$B = \int_{\lambda_1}^{\lambda_2} \varphi(\lambda) \bar{b}(\lambda) d\lambda \quad (2.4)$$

mit dem Farbreiz $\varphi(\lambda) = \Phi(\lambda) / \int_{\lambda_1}^{\lambda_2} \Phi(\lambda) d\lambda$ (spektrale Leistungsdichte der Strahlung) und den Grenzwellenlängen für den sichtbaren Bereich $\lambda_1 = 380 \text{ nm}$ und $\lambda_2 = 780 \text{ nm}$ (siehe Abschnitt 2.1.1) [10].

Die CIE (Commission Internationale de L'Éclairage) etablierte 1931 den farbmetsrischen 2° -Normalbeobachter anhand der Ergebnisse von Wrights Farbabgleichversuch für:

- 17 normalsichtige Versuchspersonen mit helladaptierten Augen
- ein Sichtfeld von 2° entsprechend dem Bereich der größten Zapfendichte und folglich höchsten Farbsensitivität und Auflösung (Fovea)

⁴ sogenannte Spektralfarben

- monochromatisches Licht der Wellenlängen [8]; [9]; [10]:

$$\begin{array}{lll} \lambda_{\text{Rot}} = 700 \text{ nm} & \text{mit} & S(700 \text{ nm}) = 72,096 \\ \lambda_{\text{Grün}} = 546,1 \text{ nm} & \text{mit} & S(546,1 \text{ nm}) = 1,379 \\ \lambda_{\text{Blau}} = 435,8 \text{ nm} & \text{mit} & S(435,8 \text{ nm}) = 1,000 \end{array}$$

Im Jahr 1964 wurde analog zum 2° -Normalbeobachter ein 10° -Normalbeobachter für das Weitwinkelsehen eingeführt, der jedoch weniger Verbreitung findet (siehe [16]).

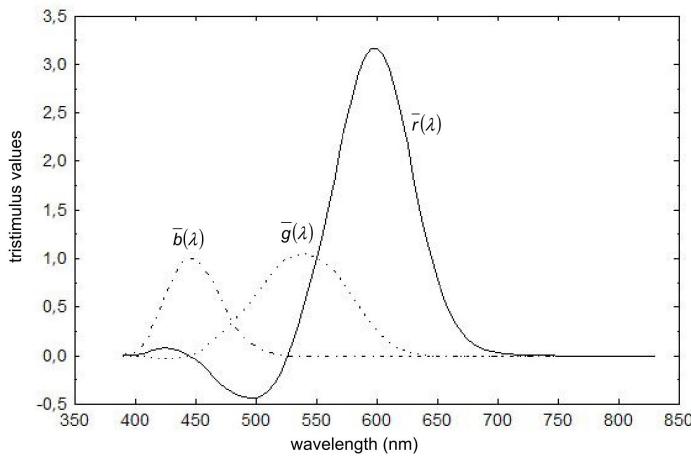


Abbildung 2.3: Spektralwertkurven der spektralen Primärvalenzen für das 10° -Sichtfeld nach Stiles und Burch (1959) [13]

Die ermittelten Spektralwertfunktionen (Abb. 2.3) weisen negative Anteile auf. Die unten angegebene Koordinatentransformation führt zum Normvalenzsystem mit positiven, allerdings nur virtuell existenten Normfarbwerten XYZ (CIE 1931 XYZ). Dabei ist Y als Helligkeit bzw. Luminanz gewählt [9] [17].

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0,49000 & 0,31000 & 0,20000 \\ 0,17697 & 0,81240 & 0,01063 \\ 0,00000 & 0,01000 & 0,99010 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.5)$$

Die Beziehung der Normfarbwerte auf die Summe führt zu den Normfarbwertanteilen xyz:

$$x = \frac{X}{X + Y + Z} \quad y = \frac{Y}{X + Y + Z} \quad z = \frac{Z}{X + Y + Z} \quad x + y + z = 1 \quad (2.6)$$

Die Normfarbtafel in Abb. 2.4 mit den Normfarbwertanteilen x und y als Koordinaten visualisiert den sichtbaren Bereich des Lichts mit dem hufeisenförmigen Spektralfarbenzug als Rand. Die Darstellung vereinfacht die Farbbeschreibung auf zwei Dimensionen [18]: Konstante Farbtöne

liegen auf Strahlen, die vom Weißpunkt ausgehen. Der Abstand zum Weißpunkt gibt die Sättigung an [9]. Bei Bedarf kann im sogenannten xyY-Farbraum die Luminanz ebenfalls dargestellt werden.

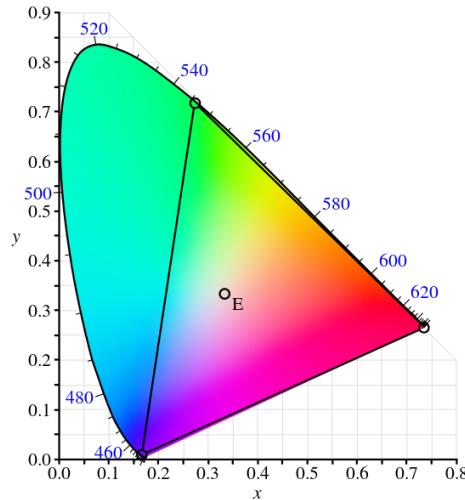


Abbildung 2.4: Normfarbtafel mit Weißpunkt und CIE-RGB [19]

2.2 Farbsysteme

Eine Übersicht über verschiedene Farbsysteme gibt Abb. 2.5. Die üblicherweise eingesetzten Systeme sind von dem Normvalenzsystem abgeleitet, orientieren sich aber am Anwendungsfall: Druckerfarben oder Monitor-Leuchtstoffe zum Beispiel grenzen den erzeugbaren Farbbereich (Gamut) ein.

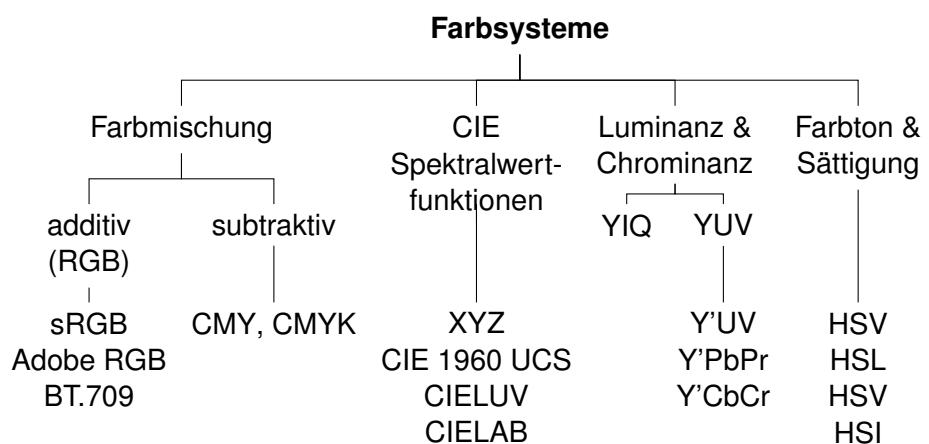


Abbildung 2.5: Klassifikation wichtiger Farbsysteme [8]

2.2.1 RGB-Systeme

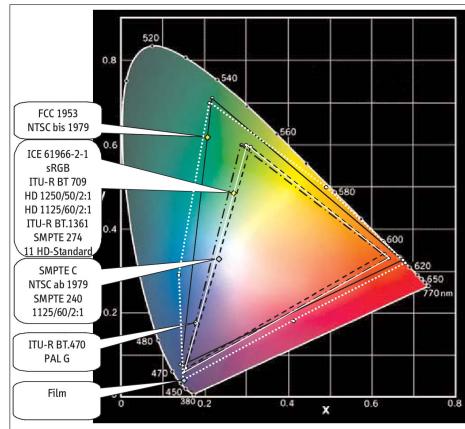


Abbildung 2.6: Gamut bekannter RGB-Systeme in der Normfarbtafel [18]

Der RGB-Raum ähnelt nicht nur dem Modell der menschlichen Farbwahrnehmung, sondern findet Anwendung bei der Bildwiedergabe, da Bildschirmfarben additiv aus drei leuchtstoffspezifischen Grundfarben reproduziert werden können. Das CIE-RGB-System (siehe Abb. 2.3) besitzt nur Bedeutung als Bezugspunkt [20]. In der Praxis wurden leuchtstoffspezifischen Primärvalenzen standardisiert. Zur vollständigen Definition genügt die Angabe von x_r, y_r, x_g, y_g, x_b und y_b der neuen Primärvalenzen und des zur Normalbeleuchtung passenden Weißpunktes W für die Luminanz Y . Dargestellt sind in Tabelle 2.1 xy -Farbwerte⁵ für RGB und der Weißpunkt wichtiger RGB-Farbräume sowie in Abb. 2.6 die sich daraus ergebenden Farbräume in der CIE-Normfarbtafel. Tabelle 2.2 listet Spezifikationen der gebräuchlichsten Normbelichtungen auf, die festlegen, welche Art Weiß W wiedergegeben wird.

Die Transformations-Matrix M von RGB nach XYZ

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = M \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \text{mit: } M = \begin{bmatrix} S_r X_r & S_g X_g & S_b X_b \\ S_r Y_r & S_g Y_g & S_b Y_b \\ S_r Z_r & S_g Z_g & S_b Z_b \end{bmatrix} \quad (2.7)$$

berechnet sich bei gegebenen xy -Koordinaten mit Gleichung (2.6) für $Y = 1$ [17], [21], [22]:

$$X = x \cdot \frac{Y}{y} \quad Y = 1 \quad Z = (1 - x - y) \cdot \frac{Y}{y} \quad (2.8)$$

⁵ Bei keiner anderen Angabe wird im Folgenden der Farbraum CIE 1931 XYZ mit 2°-Normalbeobachter zugrunde gelegt

Tabelle 2.1: Genormte xy-Farbwerde und Weißpunkte einiger RGB-Räume [20]

Normen	Weißpunkt	Primärvalenzen					
		R		G		B	
		x_r	y_r	x_g	y_g	x_b	y_b
sRGB, HDTV (ITU-R BT.709)	D65	0,6400	0,3300	0,3000	0,6000	0,1500	0,0600
	D65	0,6400	0,3300	0,2100	0,7100	0,1500	0,0600
	D65	0,6400	0,3300	0,2900	0,6000	0,1500	0,0600
	D65	0,6300	0,3400	0,3100	0,5950	0,1550	0,0700
	C	0,6700	0,3300	0,2100	0,7100	0,1400	0,0800
	E	0,7347	0,2653	0,2738	0,7174	0,1666	0,0089

Danach ist das lineare Gleichungssystem zu lösen:

$$\begin{bmatrix} S_r \\ S_g \\ S_b \end{bmatrix} = \begin{bmatrix} X_r & X_g & X_b \\ Y_r & Y_g & Y_b \\ Z_r & Z_g & Z_b \end{bmatrix}^{-1} \begin{bmatrix} X_n \\ Y_n \\ Z_n \end{bmatrix} \quad (2.9)$$

Am gebräuchlichsten ist der Farbraum gemäß BT.709 – sowohl für 1280×720 und 1920×1080 HDTV (High Definition Television) als auch im sRGB-Standard für die meisten Bildschirmmonitoren [20].

Tabelle 2.2: xy-Koordinaten und Verwendung gebräuchlicher Normbelichtungen der CIE [9]

Normbeleuchtung	x_n	y_n	Verwendung
CIE-Normlicht C	0,3101	0,3162	NTSC
CIE-Normlicht E	1/3	1/3	Energiegleiches Weiß: $X = Y = Z$
D50 (5000 K ⁶)	0,3457	0,3585	
D65 (6500 K)	0,312713	0,329016	Tageslicht, sRGB und PAL/SECAM-TV

⁶ Farbtemperatur: hohe Werte wirken bläulich („kalt“), kleine rötlich („warm“)

2.2.2 YUV

Wie in Abschnitt 2.3.5 erläutert ist die menschliche Wahrnehmung für Farbunterschiede deutlicher weniger empfindlich als für Helligkeit [20]. Um die Bandbreite bei der Farbsignalübertragung zu reduzieren, wurden die Farbdifferenzen $R - Y$ und $B - Y$ als reine Chrominanzwerte eingeführt, die von der weiterhin schwarz-weiß-kompatiblen Luminanz Y dekorreliert kodiert und übertragen werden können. Da der Grünkanal G am meisten zur Luminanz beiträgt, wird die zu erwartende geringe $G - Y$ -Differenz nicht verwendet [9] – erreicht wird gleichzeitig ein maximaler Signalrauschabstand (SNR, signal-to-noise ratio) [20]. Die Farbdifferenz- oder Chrominanzsignale $B' - Y'$ und $R' - Y'$ oder U und V bilden mit Luma Y' die Grundlage für davon abgeleitete Komponentensignale.

In der Praxis unterscheidet man:

- Y'UV für NTSC (National Television Systems Committee) oder PAL (Phase Alternating Line)
- Y'PbPr für analoges Fernsehen
- Y'CbCr für digitales Fernsehen, M-JPEG und MPEG und
- YIQ vom früheren NTSC (eine heutzutage nicht mehr gebräuchliche, 33°-gedrehte Variante von Y'UV) [20].

Analog zur Luminanz Y wird Luma Y' als gewichtete Summe mithilfe der Koeffizienten (siehe Tabelle 2.3) gebildet. Zu berücksichtigen sind die in Abschnitt 2.5 erläuterten gammakorrigierten, nichtlinearen Signale $R'G'B'$ (gammakorrigierter RGB-Farbraum)⁷. Das „Constant Luminance“-Prinzip, die Übertragung der reinen Luminanz Y unabhängig von den Chrominanzen, als Voraussetzung für die anfangs erwähnte Bandbreiteneinsparung wird so verletzt (siehe Abschnitt 2.5.2) [20]. Zur Berechnung werden die Signale auf die Wertebereiche⁸ in Tabelle 2.4 skaliert.

$$Y' = K_R \cdot R' + K_G \cdot G' + K_B \cdot B' \quad (2.10)$$

$$U = K_U(B' - Y') \quad (2.11)$$

$$V = K_V(R' - Y') \quad (2.12)$$

Beispielhaft anhand der Darstellung des kleineren RGB-Farbraums im Y'CbCr-Würfel in Abb. 2.7 ist zu erkennen, dass für gewisse Werte Y'CbCr keine RGB-Werte existieren [18].

⁷ nach Empfehlung von Poynton gekennzeichnet mit ', z. B. Y' , R' etc.

⁸ Die Reserven (*footroom* und *headroom*) sollen bei Analogtechnik ein Überschwingen verhindern und sind für digitale Signale optional

Tabelle 2.3: Koeffizienten zur Berechnung von Luma Y' für verschiedene Standards [20]

Standard	K_R	$K_G = 1 - K_R - K_B$	K_B
ITU-R BT.601	0,299	0,587	0,114
BT.709	0,2126	0,7152	0,0722
SMPTE 240M	0,212	0,701	0,087

Tabelle 2.4: Wertebereiche für YUV-Systeme

System	Wertebereich			
	$R' G' B'$	Y'	U	V
Y'UV	[0 ... 1]	[0 ... 1]	[-0,436 ... 0,436]	[-0,615 ... 0,615]
Y'CbCr → (full range)	[0 ... 255]	[16 ... 235]	[16 ... 240]	[... 255]
Y'PbPr	[0 ... 255]	[16 ... 235]	[-127,5 ... 127,5]	[... 127,5]

Die Umwandlung von RGB nach Y'UV für BT.709 erfolgt als Matrixmultiplikation:

$$\begin{bmatrix} Y' \\ U \\ V \end{bmatrix} = \underbrace{\begin{bmatrix} 0,2126 & 0,7152 & 0,0722 \\ -0,1146 & -0,3854 & 0,5000 \\ 0,5000 & -0,4542 & -0,0458 \end{bmatrix}}_{T_{RGB \rightarrow Y'UV, BT.709}} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (2.13)$$

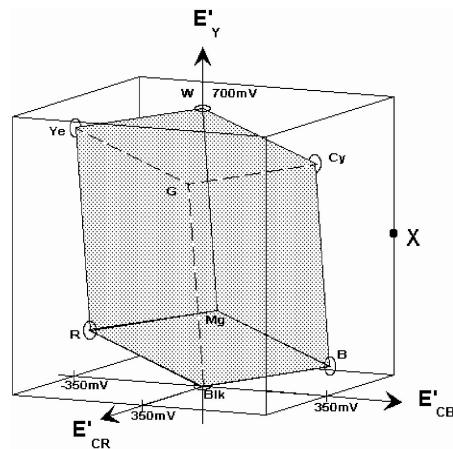


Abbildung 2.7: RGB im Y'CbCr-Farbraum [18]

Eine Inversion ermöglicht die Rücktransformation:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 0 & 1,575 \\ 1 & -0,187 & -0,468 \\ 1 & 1,856 & 0,000 \end{bmatrix}}_{T_{Y'UV \rightarrow RGB, BT.709}} \begin{bmatrix} Y' \\ U \\ V \end{bmatrix} \quad (2.14)$$

2.2.3 Empfindungsgemäße Farbräume

In Abschnitt 2.2.3 wird die wahrnehmungsbezogene Ungleichmäßigkeit des XYZ-Farbraums und davon abgeleiteten Systemen deutlich. Es wurden seitdem viele Versuche unternommen, einen empfindungsgemäßen (metrischen) Farbraum zu definieren, in dem Abstände den wahrgenommenen Farbdifferenzen entsprechen. Am meisten Beachtung finden:

- CIE 1960 UCS (Abschnitt 2.2.4),
- der CIELUV- (Abschnitt 2.2.5) und
- der CIELAB-Farbraum (Abschnitt 2.2.6)

Eingesetzt werden sie besonders bei der geräteunabhängigen Farbspezifikation oder Auswertung von Farbabständen [18], [23].

Farbabstände nach MacAdam

Ein Nachteil des NTSC-XYZ-Systems besteht in den ungleichmäßigen Darstellung empfundener Farbabstände. MacAdams Experimente (1942) ergaben die so genannten MacAdam-Ellipsen: Gebiete, in denen Farbunterschiede zur Farbe im Mittelpunkt vom menschlichen Auge nicht mehr wahrgenommen werden [8]. Abb. 2.8 zeigt diese 10-fach vergrößert in der Normfarbtafel. Insbesondere Grüntöne korrespondieren mit einer geringen Farbauflösungsfähigkeit (große Ellipsenflächen). Dies gilt auch für Farben entlang der Gelb-Blau-Achse im Verhältnis zur Rot-Grün-Achse [9].

Gegenfarbentheorie

Hering präsentierte 1878 die Gegenfarbentheorie (Vierfarbentheorie)⁹, derzufolge Menschen Rot-Grün und Gelb-Blau als Gegensatzpaare in verschiedenen Kanälen wahrnehmen. So

⁹ Während die Dreifarbentheorie auf den drei Zapfentypen in der Rezeptorschicht fundiert, konnte die nachgeschaltete Signalverarbeitung nach der Gegenfarbentheorie später als antagonistischer neuronaler Effekt in den Gangliazellen bestätigt werden [8], [24]

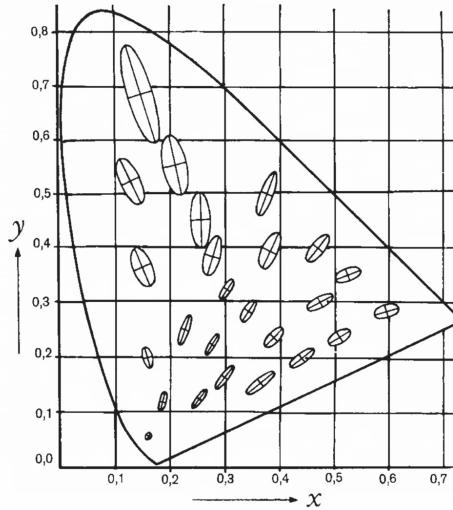


Abbildung 2.8: MacAdam-Ellipsen in der Normfarbtafel [8]

existiert ein „rötliches Gelb“, aber nicht ein „rötliches Grün“, d. h. Gegenfarben schließen sich in einer einzigen Farbempfindung aus. Diese Erkenntnis legt eine Farbdarstellung durch zwei getrennte Chrominanzinformationen und der Helligkeit nahe, wie z. B. im CIELAB-System.

2.2.4 CIE 1960 UCS

Die CIE führte 1960 erstmals mit dem UCS-Diagramm eine Transformation der xy -Ebene so durch, dass die Schwankung der Ellipsenflächen von 1 : 20 in der Normfarbtafel auf 1 : 2 reduziert werden konnte. Die neuen Koordinaten u und v ergeben sich folgendermaßen [10]

$$u = \frac{4x}{12y - 2x + 3} \quad v = \frac{6y}{12y - 2x + 3} \quad (2.15)$$

2.2.5 CIELUV (CIE 1976 L^{*}u^{*}v^{*})

Die CIE korrigierte 1979 den v -Wert [23]. Unter Einbezug des annähernd wahrnehmungsgetreuen Helligkeitswerts L^* als dritte Dimension wurde CIELUV (CIE 1976 L^{*}u^{*}v^{*}) entwickelt.

Mit der abkürzenden Funktion $f(t)$ gilt:

$$f(t) = \begin{cases} \left(\frac{29}{3}\right)^3 t, & t \leq \left(\frac{6}{29}\right)^3 \\ 116 \cdot t^{1/3} - 16, & t > \left(\frac{6}{29}\right)^3 \end{cases} \quad (2.16)$$

$$L^* = f\left(\frac{Y}{Y_n}\right) \quad (2.17)$$

$$u^* = 13 \cdot L^* \cdot (u' - u'_n) \quad (2.18)$$

$$v^* = 13 \cdot L^* \cdot (v' - v'_n) \quad (2.19)$$

$$u' = \frac{4x}{12y - 2x + 3} = \frac{4X}{X + 15Y + 3Z} \quad (2.20)$$

$$v' = \frac{9y}{12y - 2x + 3} = \frac{9Y}{X + 15Y + 3Z} \quad (2.21)$$

Zur Berechnung ist eine Referenz-Helligkeit durch Y_n sowie der Weißpunkt durch u'_n und v'_n anzugeben.

2.2.6 CIELAB (CIE 1976 $L^*a^*b^*$)

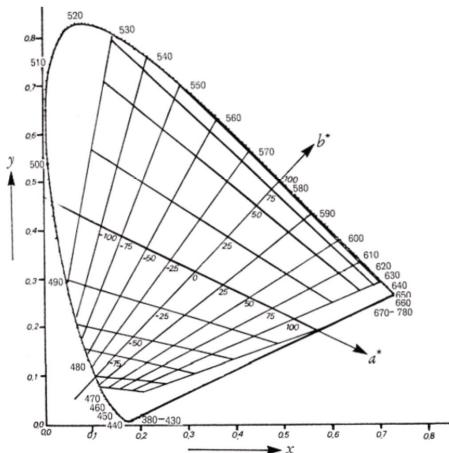


Abbildung 2.9: Transformation von xy nach a^*b^* veranschaulicht für die Helligkeit $Y = 20$ [8]

Der CIELAB (CIE 1976 $L^*a^*b^*$)-Farbraum greift mit den Farbwerten a^* und b^* auf der Rot-Grün- bzw. Blau-Gelb-Achse die Gegenfarbentheorie in Abschnitt 2.2.3 auf (vgl. Achsen in Abb. 2.9).

Die Helligkeit L^* ist identisch zur Definition in CIELUV.

$$a^* = 500 \left(f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right) \right) \quad (2.22)$$

$$b^* = 200 \left(f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right) \right) \quad (2.23)$$

$$\text{mit: } X_n = \frac{X}{X_n} \quad Y_n = \frac{Y}{Y_n} \quad Z_n = \frac{Z}{Z_n} \quad (2.24)$$

Das CIELAB-System eignet sich vor allem zur Bewertung von Körperfarben, während das CIELUV-System besonders bei der additiven Farbmischung an Monitoren von Bedeutung ist [8].

Die Rücktransformation erfolgt mit:

$$x_n = f^{-1} \left(\frac{L^* + 16}{116} + \frac{a^*}{500} \right) \quad (2.25)$$

$$y_n = f^{-1} \left(\frac{L^* + 16}{116} \right) \quad (2.26)$$

$$z_n = f^{-1} \left(\frac{L^* + 16}{116} - \frac{b^*}{200} \right) \quad (2.27)$$

$$f^{-1}(t) = \begin{cases} t^3 & t > \frac{6}{29} \\ 3 \left(\frac{6}{29} \right)^2 \left(t - \frac{4}{29} \right) & t \leq \frac{6}{29} \end{cases} \quad (2.28)$$

$$\text{mit: } X = x_n \cdot X_n \quad Y = y_n \cdot Y_n \quad Z = z_n \cdot Z_n \quad (2.29)$$

2.2.7 Farbabstände

Zur Messung der Wahrnehmung von Farbabständen ΔE können je nach Farbsystem unterschiedliche Formeln (meist die Euklidische Distanz) herangezogen werden:

- im RGB-Raum:

$$\Delta E_{RGB} = \sqrt{(R_2 - R_1)^2 + (G_2 - G_1)^2 + (B_2 - B_1)^2} \quad (2.30)$$

- im CIELUV-Raum:

$$\Delta E_{uv}^* = \sqrt{(L_2^* - L_1^*)^2 + (u_2^* - u_1^*)^2 + (v_2^* - v_1^*)^2} \quad (2.31)$$

- im CIELAB-Raum:

$$\Delta E_{ab}^* = \sqrt{(L_2^* - L_1^*)^2 + (a_2^* - a_1^*)^2 + (b_2^* - b_1^*)^2} \quad (2.32)$$

Bessere Gleichabständigkeit z. B. durch eine andere Gewichtung geht meist mit einem höheren Rechenaufwand einher. Vor allem im CIELAB-Raum optimierte die CIE 1994 und 2000 die Formeln. Ein Wert von 1 entspricht im CIELUV- oder CIELAB-System näherungsweise einer JND (Just Noticeable Difference) [20].

2.2.8 HSV und Alternativen

Menschen beschreiben Farben häufig intuitiv anhand der Helligkeit (engl. *luminance*) und Farbart. Letztere setzt sich zusammen aus dem Farbton¹⁰ (engl. *hue*) und aus der Farbsättigung (engl. *saturation*) [9], [25]. Es existieren verschiedene, wenn auch nicht einheitlich festgelegte HS-Farbmodelle, die diese Terminologie mit drei Valenzen reflektieren: *hue*, *saturation* und einem in nachstehender Tabelle zugeordneten Hellwert (siehe Tabelle 2.5)

Tabelle 2.5: unterschiedliche HS-Farbmodelle

Farbmodell	Helligkeitswert		
	Buchstabe	engl.	dt.
HSV (hue saturation value)	V	value	Hellwert
HSL (hue saturation lightness)	L	lightness	relative Helligkeit
HSV (hue saturation brightness)	B	brightness	absolute Helligkeit
HSI (hue saturation intensity)	I	intensity	Lichtintensität

Eine HSV-Variante berechnet [26] mit den normierten Werten $r, g, b, h, s, v \in [0,1]$:

$$M = \max(r, g, b) \quad m = \min(r, g, b) \quad (2.33)$$

$$h' = \begin{cases} 0, & M = m \Leftrightarrow r = g = b \\ \frac{g - b}{M - m}, & r = M \\ \frac{b - r}{M - m} + 2, & g = M \\ \frac{r - g}{M - m} + 4, & b = M \end{cases} \quad M = m \Leftrightarrow r = g = b = 0 \quad (2.34)$$

$$s = \begin{cases} 0, & M = 0 \Leftrightarrow r = g = b = 0 \\ \frac{M - m}{M}, & \text{sonst} \end{cases} \quad v = M \quad (2.35)$$

¹⁰ nach der DIN (Deutsches Institut für Normung) eigentlich „Bunton“ zu nennen [8]

2.3 Das Auflösungsvermögen des Menschen

Folgende Abschnitte erörtern die Grenzen des menschlichen Sehvermögens, kleine visuelle Reize zu wahrzunehmen.

2.3.1 Helligkeitsempfindung

In der Netzhaut sind Zapfen und Stäbchen für die Helligkeitswahrnehmung verantwortlich – letztere allerdings nur für das farblose, photopische Sehen im Dunklen. Die beiden spektralen Hellempfindlichkeitskurven $V(\lambda)$ nach Messungen der CIE sind in Abb. 2.10 zu sehen [10]. Erkennbar ist der große Beitrag von Grüntönen zum Helligkeitseindruck [9]. Die Helligkeitsempfindung hängt nicht nur von der Leuchtdichte und Wellenlänge, sondern zusätzlich von der Adaption des Auges an die Belichtungssituation und Hintergrundbeleuchtung sowie der Größe des Helligkeitsfeldes ab [15].

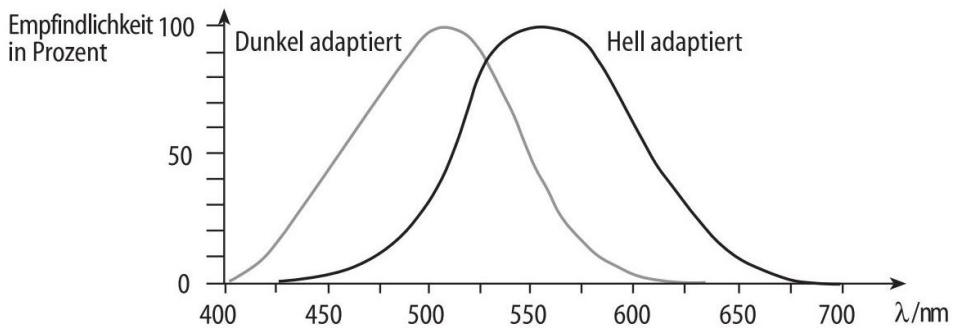


Abbildung 2.10: Spektrale Hellempfindlichkeit $V(\lambda)$ des menschlichen Auges [9]

Verschiedene Relationen zwischen der subjektiven Helligkeit H und der Leuchtdichte L wurden vorgeschlagen. Fechner folgerte aus Webers Arbeiten (siehe Abschnitt 2.3.2) einen logarithmischen Zusammenhang zwischen der Hellempfindung H und der Leuchtdichte L [8]:

$$H = c \log(kL) \quad (2.36)$$

Spätere Erkenntnisse beschränken die Gültigkeit des Weber-Fechner-Gesetzes auf mittelstarke Reize (vgl. Weber-Kontrast, Abb. 2.11). Die CIE favorisiert mit der Potenzfunktion im CIELUV-System (vgl. Abschnitt 2.2.5) beispielsweise eine Variante der Stevensschen Potenzfunktion, die für Sinnesreize allgemein gilt (Intensität S und Empfindung E) [8]:

$$E = kS^\alpha \quad (2.37)$$

2.3.2 Kontrast

Die menschliche Helligkeitswahrnehmung hängt weniger von absoluten Luminanzen als von relativen Unterschieden ab (s. Abschnitt 2.3.2, Weber-Kontrast), für die unterschiedliche Kontrast-Maße C definiert wurden [14].

Kontrast-Verhältnis

Für Monitore wird meistens das Kontrast-Verhältnis CR (contrast ratio)

$$C_R = \frac{L_{max}}{L_{min}} \quad (2.38)$$

zwischen dem hellsten und dunkelsten Bildelement angegeben, die zeitlich oder örtlich getrennt sind. ITU-R BT.815-1 [27] empfiehlt die Messung an einem Testbild mit einem weißen und vier rundherum angeordneten schwarzen Quadranten der maximalen bzw. minimalen Signalamplitude auf einem grauen Hintergrund mit 50% des Weißwertes. Die einzelnen gemessenen Leuchtdichten in den Feldern (L_{b1} bis L_{b4} für Schwarz, L_w für Weiß) ergeben:

$$C_{BT.815} = \frac{L_{b1} + L_{b2} + L_{b3} + L_{b4}}{4L_w} \quad (2.39)$$

Weber-Kontrast

Der Weber-Kontrast

$$C_W = \frac{\Delta L}{L_0} \quad (2.40)$$

gilt für eine kleine Helligkeitsabweichung ΔL_0 auf einem einheitlich hellen Hintergrund (L_0). Wird nach dieser Definition der JND (Just Noticeable Difference) über die absolute Helligkeit aufgetragen (Abb. 2.11), zeigt sich in einem großen Bereich die 1834 von Weber gefundene Gesetzmäßigkeit: Der empfundene Reizunterschied ist proportional zum Ausgangsreiz oder anders ausgedrückt:

$$\frac{\Delta L}{L_0} = konst. \quad (2.41)$$

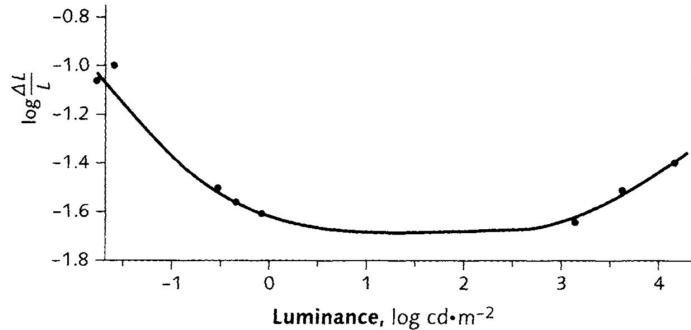


Abbildung 2.11: visuelle Wahrnehmungsschwelle für verschiedene Helligkeiten [20]

Michelson-Kontrast und Ortsfrequenz

Bei periodischen Mustern eignet sich als Maß am besten der Michelson-Kontrast¹¹ [14]:

$$C_M = \frac{L_{max} - L_{min}}{L_{max} + L_{min}} \quad (2.42)$$

2.3.3 Modulationsübertragungsfunktion

Die sogenannte Modulationsübertragungsfunktion MTF (modulation transfer function) ist ein wichtiges Qualitätsmaß für das örtliche Auflösungsvermögen optischer Systeme. Sie beschreibt, wie optische Systeme Kontrastdetails abzubilden vermögen. Zur Messung können verschiedene Testmuster eingesetzt werden, wie ein Linienspektrum oder Sinus-Raumgitter (Abb. 2.12), die bestimmt werden durch den Kontrast C_M (gemessen nach Michelson) und die Ortsfrequenz $f = 1/d$ als Kehrwert der räumlichen Periode d (gebräuchliche Einheiten: cpd (cycles per degree) oder lpmm (lines per mm)). Im räumlichen Frequenzbereich korrespondieren feine Details mit hohen Frequenzen.

Wird das abgebildete Sinusgitter mit der Leuchtdichte $L = L_0(1 + m \cdot \cos(2\pi x/d))$ und Modulationsgrad m abgebildet auf die Verteilung $L' = L_0(1 + m_B \cos((2\pi x/d) + \psi))$ im Bildbereich, so kann das Verhältnis m_B/m in Abhängigkeit von der Ortsfrequenz als MTF gedeutet werden [9], [10]. Mathematisch entspricht die MTF der Fouriertransformierten eines Linienspektrums bzw. eines Punktes. Die Fourieranalyse ermöglicht eine Auswertung allgemeiner Strukturen.

¹¹ Er deckt sich bis auf einen Faktor 2 für die mittlere Helligkeit $L_0 = (L_{max} + L_{min})/2$ und dem Unterschied $\Delta L = L_{max} - L_{min}$ mit Webers Definition.

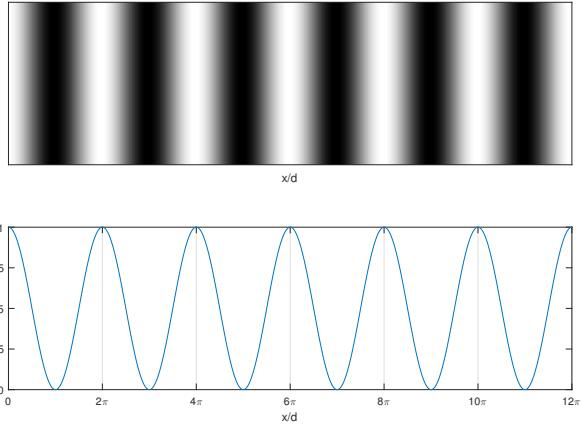


Abbildung 2.12: Sinusgitter mit der Leuchtdichte $L = L_0(1 + m \cdot \cos(2\pi x/d))$ mit Modulationsgrad m

MTF und CSF des Menschen

Die experimentelle Bestimmung der minimal erkennbaren Kontrastschwelle (CT, contrast threshold; vergleichbar: JND) anhand oben genannter Muster liefert eine analoge psychophysische Beschreibung der Kontrastwahrnehmung des Menschen wie die MTF: Die Inverse der CT (contrast threshold) für verschiedene Ortsfrequenzen ergibt die Kontrastsensitivitätsfunktion (contrast sensitivity function CSF). Die MTF des menschlichen visuellen Systems (Beispiel in Abb. 2.13) wird physiologisch durch das optische Augensystem (z. B. durch Akkommodation und Abbildungseigenschaften) sowie die Rezeptordichte für hohe Frequenzen limitiert.

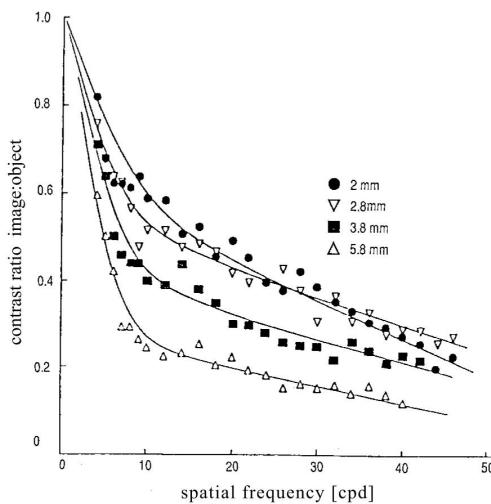


Abbildung 2.13: MTF des Menschen gemessen für verschiedene Pupillen-Durchmesser [28]

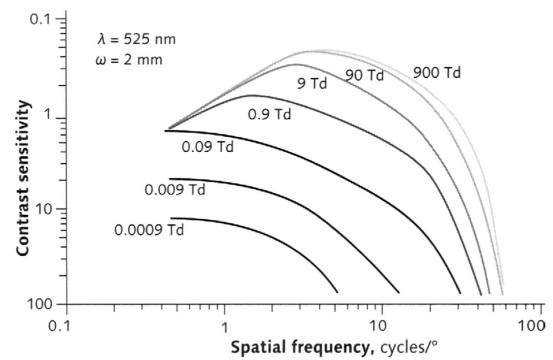


Abbildung 2.14: CSF in Abhängigkeit vom Adaptionszustand; gemessen für Linienspektren verschiedener Helligkeiten [20]

Die CSF berücksichtigt zusätzlich neuronale Effekte wie die laterale Inhibition für kleine Ortsfrequenzen¹² [9], [15]. Abbildung 2.14 verdeutlicht die Abhängigkeit der CSF vom Adaptionszustand. Die CSF weist insgesamt eine Bandpasscharakteristik mit einem Hochpunkt bei ca. 3 cpd auf [20], [28], [29].

In den letzten Jahrzehnten wurden komplexere CSF-Modelle vorgeschlagen, die auch andere Parameter zu berücksichtigen versuchen (vgl. [30]–[32]).

2.3.4 Zeitliche Auflösung

Untersuchungen des menschlichen Auges zeigen ein degressives zeitliches Auflösungsvermögen bei höheren Frequenzen mit einem Maximum für ruhende Bilder. Ab einer Wiederholrate von ca. 20 Hz nehmen Menschen Einzelbilder als gleichmäßige Bewegung wahr und ab ca. 50 Hz kein störendes Großflächenflimmern, wobei diese Flimmer-Verschmelzungsfrequenz für größere Helligkeiten und im äußeren Blickfeld höher anzunehmen ist [9]. Die Kurve Abb. 2.15 lässt erkennen, dass die zeitliche Kontrastsensitivität wie die räumliche zwar auch als Bandpass funktioniert, aber nur für hohe Frequenzen von dieser getrennt modelliert werden kann [28].

2.3.5 Chromatische Auflösung

Beispielhaft sind chromatische und achromatische CSF zum Vergleich in Abb. 2.15 dargestellt. Allgemein festzustellen ist eine höhere Kontrastempfindlichkeit für reine Helligkeitsunterschiede als für Farben, vor allem bei hohen räumlich-zeitlichen Frequenzen. Ferner fällt die Reaktion auf Blau-Gelb-Unterschiede unempfindlicher als für Rot-Grün-Stimuli aus und geht bei höheren zeitlichen Frequenzen früher in eine reine Schwarz-Weiß-Wahrnehmung über [14].

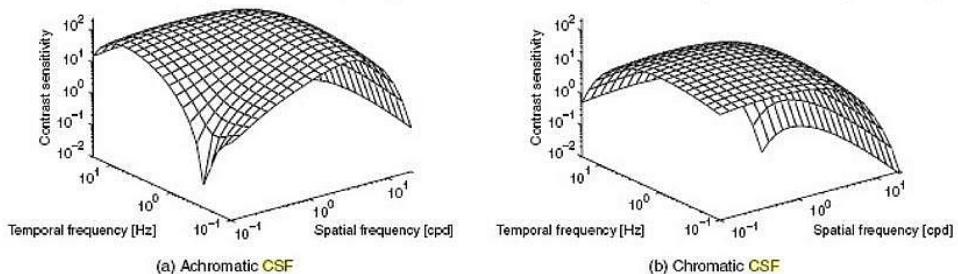


Abbildung 2.15: chromatische und achromatische zeitlich-räumliche CSF [14]

¹² Abrupte Helligkeitsunterschiede werden somit kontrastverschärfend wahrgenommen (bekannter Effekt: Machsche Bänder)

2.4 Bildschirmtechnologie

Bildwiedergabegeräte wandeln das Videosignal in sichtbares Licht um, welches dem Betrachter als Bild erscheinen soll. Die Trägheit der Wiedergabe und des menschlichen Sehsinns ermöglichen den Eindruck eines kontinuierlichen Videos trotz sequentiell ausgegebener Einzelbilder (*frames*).

Das begrenzte menschliche Auflösungsvermögen zusammen mit der additiven Mischung erlaubt die Rekonstruktion eines kohärenten Farbbildes durch drei Grundfarben RGB, die als Subpixel die Bildpunkte bilden und durch ein Tripel aus elektrischen Farbwertsignalen gesteuert werden [9]. Diese werden zur effizienten Transmission codiert und im digitalen Video durch Bits diskret repräsentiert. Üblich ist die 8-Bit-Codierung für Stufen von 0 bis 255. Das Primärsystem legt den reproduzierbaren Farbumfang (Gamut) fest¹³ [12].

Bei aktiven Display-Typen strahlen die Pixel selbst, während passive Displays die Transmission oder Reflexion der Hintergrundbeleuchtung steuern¹⁴. Die technische Ausführung der Monitore bestimmt maßgeblich die Bildqualität – besonders hinsichtlich Auflösung, Größe, Helligkeit, Kontrast, Blickwinkelabhängigkeit, der farbentreuen Wiedergabe, Zeitverhalten und der in dieser Arbeit behandelten Gradation [9].

2.4.1 Kathodenstrahlröhren (CRT)

Kathodenstrahlröhren (Cathode Ray Tube, CRT) wurden im 20. Jahrhundert vorherrschend als Monitore eingesetzt und beeinflussten mehrere noch heutzutage gängige Standards in der Videotechnologie. In Farbbildröhren erzeugt die mithilfe eines Glühdrahtes erhitzte Kathode einen gelenkten Elektronenstrahl, der Phosphore zum Leuchten in einer der drei Farben Rot, Grün oder Blau anregt [9] (siehe Abb. 2.16).

2.4.2 Flachbildschirme

Neuere Techniken erlauben flache Bauweisen. Zu den aktiven FPDs (Flat Panel Displays) zählen OLEDs (Organic Light Emitting Diodes) und Plasmadisplays, die organische Leuchtdioden bestimmter Wellenlängen als Subpixel bzw. UV-Plasmastrahlung zur Anregung von Leuchttstoffen nutzen. Das vollständige Ausschalten einzelner OLEDs erlaubt tiefe Schwarzwerte.

¹³ Den Beispielen der Normfarbtafel entnehmbar deckt der Gamut zwar hoch gesättigte (Spektral-)Farben nicht komplett ab, diese kommen allerdings selten als Körperfarben in der Natur vor.

¹⁴ Auf Projektions-Systeme, die im Gegensatz zu Bildschirmen für die direkte Wiedergabe auf subtraktiver Mischung basieren, wird in dieser Arbeit nicht weiter eingegangen

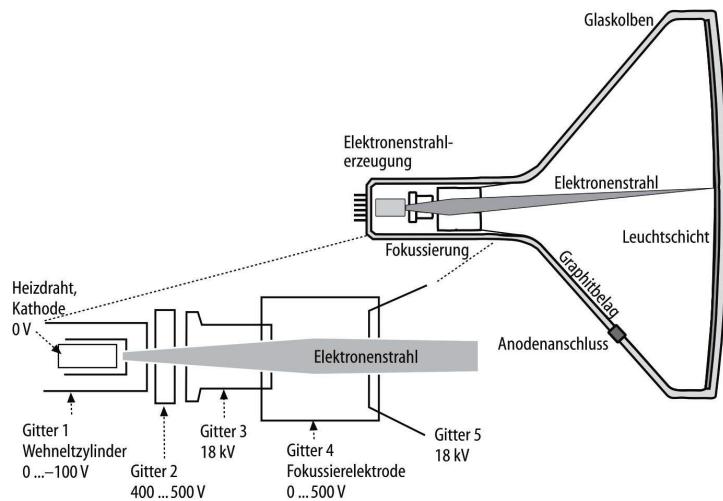


Abbildung 2.16: Schema einer Kathodenstrahlröhre [9]

Plasmamonitor erreichen hohe Leuchtdichten, allerdings häufig mit einer bildhelligkeitabhängigen Regelung. Flüssigkristallbildschirme (Liquid Crystal Display, LCD) steuern als passive Typen die Durchlässigkeit der Pixel für die Hintergrundbeleuchtung, indem sie die Polarisation des Lichts verwenden [9], [33].

2.5 Gradation

Das in dieser Arbeit primär analysierte Problem, welches in der bisherigen Modulationsmethode unberücksichtigt ist, besteht in der Gradation von Bildschirmen. Diese ist historisch und technisch auf die lange Zeit vorherrschenden Kathodenstrahlröhren (siehe Abschnitt 2.4.1) zurückzuführen [34].

2.5.1 Kennlinie von Kathodenstrahlröhren

Während der Anoden-/Elektronenstrahlstrom I_a eine proportionale Leuchtdichte $L = cI_a$ im Leuchtstoff generiert, hängen I_a und somit L nichtlinear von der Steuerspannung U_{st} ab. Die elektro-optische Wandlungskennlinie (EOTF, electro-optical transfer function) kann angenähert werden durch eine Potenzfunktion [35]:

$$L = \begin{cases} kU_{st}^\gamma & U_{st} \geq 0 \\ 0 & \text{sonst} \end{cases} \quad (2.43)$$

Der Wert γ nimmt Werte im Bereich 2,7 bis 2,9 an und die hervorgerufene Nichtlinearität wird Gradation bzw. Gammaverzerrung genannt. Die doppelt-logarithmische Darstellung der EOTF in Abb. 2.17 ermöglicht ein Ablesen von γ anhand der Steigung.

$$\gamma = \frac{\Delta \log I_a}{\Delta \log U_{st}} \quad (2.44)$$

Für R , G und B können sich diese bedingt durch verschiedene Leuchtstoffeigenschaften oder abhängig vom Blickwinkel und der Hintergrundbeleuchtung unterscheiden [36]. Besonders die EOTF von Grün weist für hohe Signalspannungen häufig eine Sättigung in Form eines geringeren Anstiegs der Helligkeit auf (vgl. Abb. 2.18 für die auf den Weißpunkt normierten, emittierten Strahldichten der RGB-Leuchtstoffe; die Abweichungen von der angenommenen Proportionalität gemäß Abschnitt 2.5.1 können in der Praxis häufig vernachlässigt werden) [35].

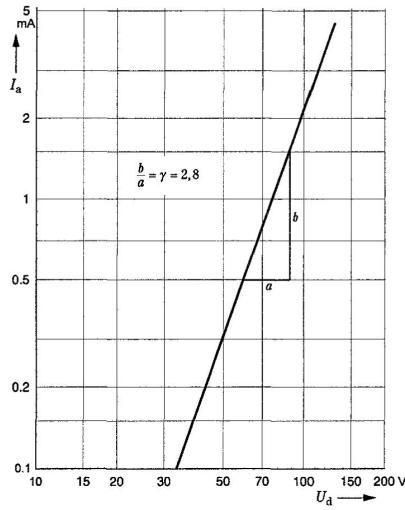


Abbildung 2.17: Strahlstromkennlinie $I_a(U_{st})$ einer Farbbildröhre [35]

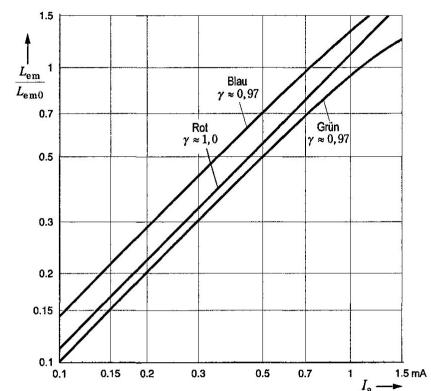


Abbildung 2.18: normierte Strahldichten L/L_0 der R -, G - und B -Leuchtstoffemission als Funktion des Elektronenstrahlstroms I_a [35]

2.5.2 Wiedergabefehler durch Gammaverzerrung

Die Gammaverzerrung bewirkt eine dunklere Wiedergabe des Originalbildes (siehe Abb. 2.19c) und eine Erhöhung der Sättigung s (siehe Abb. 2.19b) gemäß

$$s = 1 - (1 - s_1)^\gamma \quad (2.45)$$

mit der Originalsättigung s_1 ohne Gradation (d. h. für $\gamma = 1$). Außerdem verschieben sich Farbtöne in Richtung der Primärvalenzen, wie Abb. 2.19a und Abb. 2.20 veranschaulichen.

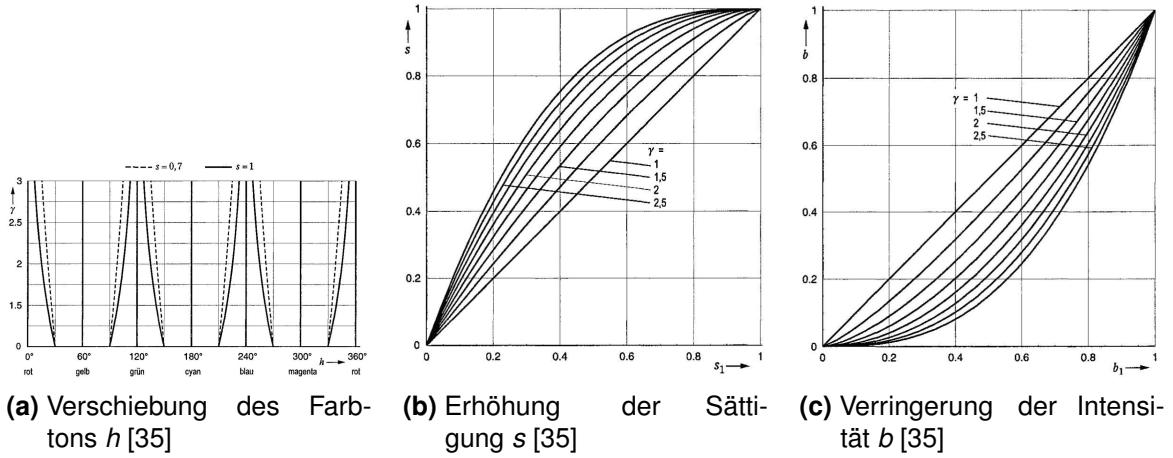


Abbildung 2.19: Wiedergabefehler durch die Gammaverzerrung im hsb-System in Abhängigkeit vom nichtverzerrten Originalwert h_1, s_1, b_1

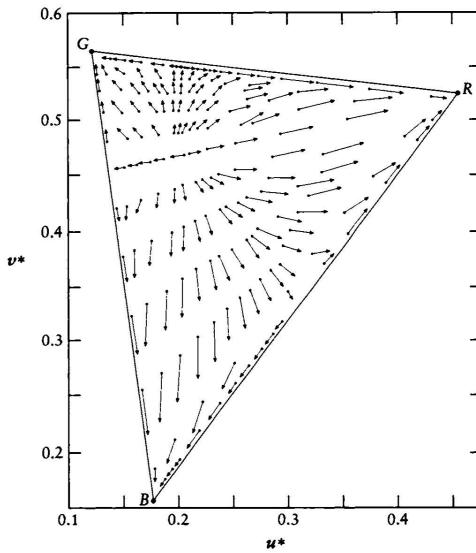


Abbildung 2.20: Farbtonverschiebung in $u^* v^*$ -Koordinaten durch Gammakorrektur [37]

2.5.3 Gammakorrektur

Zum Ausgleich der Gammaverzerrung und zur richtigen Bildreproduktion werden die Farbwertsignale R , G und B mit einer inversen Übertragungsfunktion OETF (opto-electronic transfer function) korrigiert. Es wurde vereinbart, diese Gammavorverzerrung bzw. -korrektur mit dem Wert $\gamma_C = \gamma$ einmal aufnahmeseitig durchzuführen, um den technischen Aufwand für die Emp-



Abbildung 2.21: Gammaverzerrung am Bildbeispiel mit $\gamma = 2.8$

fangsgeräte einzusparen:

$$R' = R^{1/\gamma_C} \quad \text{analog für } G \text{ und } B \quad (2.46)$$

Bei Unterkompensation mit einem kleineren Wert γ_C - z. B. $\gamma_C = 2,2$ bzw. $1/\gamma_C = 0,45$ und $\gamma_C = 2,8$ - ergibt sich für die gesamte Übertragungsstrecke ein sogenanntes Überalles-Gamma von $\gamma_{ges} \approx 1,27$. Die resultierende Kontrastanhebung wird, z. B. bei dunklen Umgebungsbedingungen, subjektiv meist präferiert [35], [37].

Neben der technisch-inhärenten Erklärung bietet die Gammakorrektur den Vorteil, in etwa der logarithmischen bzw. der durch eine Potenzfunktion beschriebenen menschlichen Helligkeitswahrnehmung zu entsprechen (siehe Abschnitt 2.3.1 und Gleichung (2.17)). Eine gleichmäßige Signalstufung wird so ebenfalls gleichmäßig hell z. B. als lineare Grautonskala empfunden.

Dies gilt auch für eventuelle Rauschamplituden. Ohne die Vorverzerrung würde Rauschen in dunklen Bildsegmenten störender wahrgenommen werden. Die Gammakorrektur erhöht also den SNR [35] und reduziert gleichzeitig die nötige Bitanzahl zur Grauwert-Repräsentation. Bei einer angenommenen minimal wahrgenommenen Helligkeitsabweichung von 1 % reicht eine 8-Bit-Codierung bei nichtlinearer Skalierung aus, sonst wären 12 Bit notwendig [15], [20]

Modernere Bildschirmtechnologien (FPDs, s. Abschnitt 2.4.2) weisen zwar auch nichtlineare Kennlinien auf (z. B. s-förmig bei LCDs [38]), können diese aber technisch ausgleichen. Plasmadisplays regeln die Pixelhelligkeiten beispielsweise durch Pulsweitenmodulation [20], [36]. Aus Kompatibilitätsgründen wird die Gammavorverzerrung lokal durch Signal-Prozessoren bzw. mithilfe einer LUT (Look-Up-Tabelle) berechnet – meist entsprechend verbreiteter Standards wie BT.709 oder BT.1886 (siehe Gleichung (2.49)) oder neuerer Vorschläge wie der EOTF für HDR-Video in SMPTE ST2084 PQ zur effizienten Bit-Ausnutzung bezüglich der menschlichen

Helligkeitswahrnehmung [39], [40] [9]:

$$V' = \left(\frac{c_1 + c_2 L^{m_3}}{1 + c_3 L^{m_1}} \right)^{m_2} \quad (2.47)$$

$$L = \left(\frac{V'^{(1/m_1)} - c_1}{c_2 - c_3 V'^{(1/m_2)}} \right)^{(1/m_2)} \quad (2.48)$$

$$m_1 \approx 0,159, \quad m_2 \approx 78,844, \quad c_1 \approx 0,836, \quad c_2 \approx 18,852, \quad c_3 \approx 18,688$$

Zu Beachten ist der Einfluss weit verbreiterter Einstellregler wie *brightness* und *contrast*. Konventionell legt *brightness* den wiedergegebenen Schwarzwert L_B ohne Eingangssignal ($V' = 0$) fest. Der Faktor a (*contrast*) skaliert danach die gesamte Kennlinie entsprechend und bestimmt die Bildschirmhelligkeit L_W für Weiß ($V' = 1$) [34], [41]:

$$L = a (\max [V' + b, 0])^\gamma \quad (2.49)$$

$$a = (L_W^{1/\gamma} - L_B^{1/\gamma})^\gamma \quad (2.50)$$

$$b = \frac{L_B^{1/\gamma}}{L_W^{1/\gamma} - L_B^{1/\gamma}} \quad (2.51)$$

2.5.4 Standards und Modelle

Für die Gammakorrektur und Gradation wurden in der Vergangenheit diverse Standards eingeführt. Die CIE definiert im HD-Standard BT.709, der auch rückwirkend für SDTV (Standard Definition Television) eingesetzt wird, für die vorverzerrten Videosignale $V' = R', G'$ bzw. B' und die linearen Eingangssignale $L = R, G$ bzw. B [20] [42]

$$V' = \begin{cases} 4,5L & 0 \leq L < 0,018 \\ 1,099L^{0,45} - 0,099 & 0,018 \leq L \leq 1 \end{cases} \quad (2.52)$$

Die exakte Inversion der reinen Potenzfunktion in Gleichung (2.43) würde zu einer unendlichen Steigung im Ursprung und starkem Rauschen in dunklen Pixeln führen. Deshalb ist anfangs eine stückweise Geradengleichung festgelegt [20]. Für die EOTF wird $L = (V')^{2,5}$ angenommen, sodass sich effektiv ein Überalles-Gamma von $\gamma_{ges} \approx 1,25$ einstellt [20].

BT.709 löst heutzutage auch größtenteils den Standard von 1953 NTCS (mit $\gamma_C = 2,2$) sowie die europäische Spezifikation für 576i oder frühe PAL- und SECAM-Systeme (mit $\gamma_C = 2,8$) ab. Der Standard SMPTE (Society of Motion Picture and Television Engineers) 240M für SDTV

galt vor BT.709 allerdings zeitweise mit folgender OETF [20], [43]:

$$V' = \begin{cases} 4L & 0 \leq L < 0,0228 \\ 1,1115L^{0,45} - 0,1115 & 0,0228 \leq L \leq 1 \end{cases} \quad (2.53)$$

und einer Gradation gemäß:

$$L = \begin{cases} \frac{V}{4} & 0 \leq V < 0,0913 \\ \frac{V^{1/0,45} + 0,1115}{1,1115} & 0,0913 \leq L \leq 1 \end{cases} \quad (2.54)$$

Der Vorschlag BT.1361 erweitert den Farb-Gamut von BT.709 durch die Nutzung von negativen Signalwerten und setzt eine Funktion nach Gleichung (2.55) fest [44], er wurde aber nicht in die Praxis umgesetzt [45].

$$V' = \begin{cases} -\frac{1,099(-4L)^{0,45} - 0,099}{4} & -0,25 \leq L < -0,0045 \\ 4,5L & -0,0045 \leq L < 0,018 \\ 1,099L^{0,45} - 0,099 & 0,018 \leq L \leq 1,33 \end{cases} \quad (2.55)$$

Der Standard sRGB (Standard-RGB-Farbraum) findet bei PC-Monitoren, Druckern und im Internet Anwendung und gibt eine ähnliche Transfercharakteristik wie BT.709 an [20]:

$$V' = \begin{cases} 12,92L & 0 \leq L < 0,0031308 \\ 1,055 \cdot L \left(\frac{1}{2,4} \right) - 0,055 & 0,0031308 \leq L \leq 1 \end{cases} \quad (2.56)$$

$$L = \begin{cases} \frac{V'}{12,92} & 0 \leq V' < 0,04045 \\ \left(\frac{V' + 0,055}{1,055} \right)^{2,4} - & 0,04045 \leq V' \leq 1 \end{cases} \quad (2.57)$$

Abb. 2.22 stellt die Transfercharakteristiken im Vergleich dar und verdeutlicht die Ähnlichkeit zur perzeptuellen Helligkeit L^* aus dem CIELUV-System (siehe Abschnitt 2.2.5). Dennoch weisen die Kurven untereinander merkliche Abweichungen auf.

2.5.5 Constant Luminance

Werden die nichtlinearen Komponenten R' , G' und B' nach der Gammakorrektur durch Matrixmultiplikation in den YUV-Farbraum transformiert, resultiert dies in einer Abweichung von

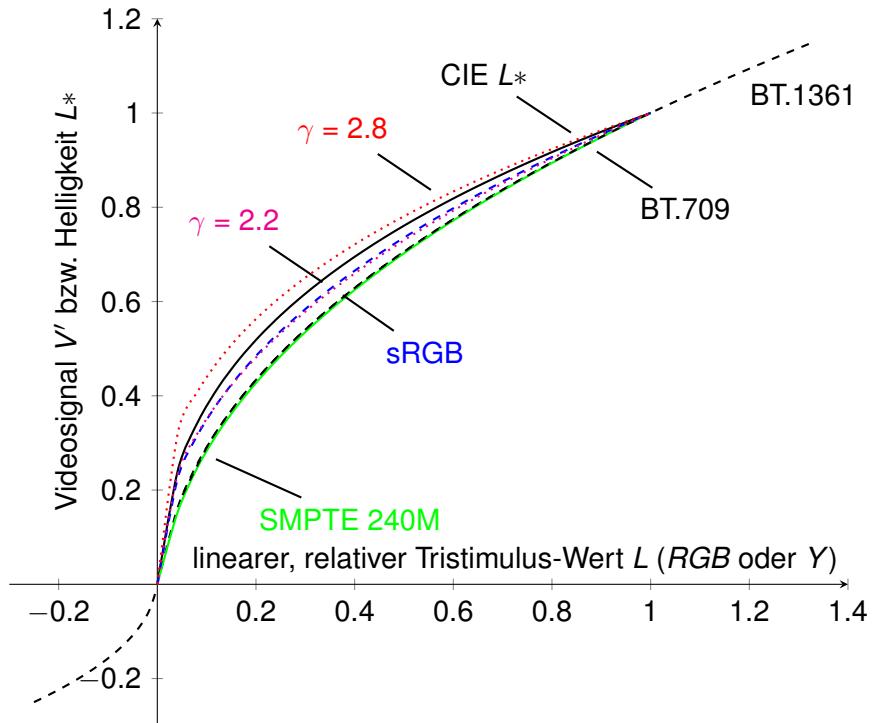


Abbildung 2.22: Kennlinien im Vergleich: EOTFs und L^*

den eigentlichen Werten YUV gebildet aus dem linearen RGB-Tristimulus. Luma¹⁵ Y' spiegelt im Gegensatz zur Luminanz Y die menschliche Helligkeitsempfindung nicht exakt wieder (siehe Weber-Fechner-Gesetz in Abschnitt 2.3.2). Der Fehler erscheint als Übersprechen (engl. *crosstalk* oder *leakage*):

- von den Chrominanzen U' und V' in die Luminanz Y bzw. relative Helligkeit L (*cross luminance*) oder umgekehrt
- von Y' in die Chrominanzinformationen U und V (*cross colour*).

Dadurch wird die Unabhängigkeit der Kanäle und das Constant-Luminance-Prinzip (vgl. Abschnitt 2.2.2) verletzt. Der Effekt tritt besonders bei gesättigten Farben und feinen Strukturen¹⁶ auf [9].

¹⁵ Poynting rät zur abgrenzenden Notation der neuen Größe mit Y' , U' und V' , um Missverständnisse zu vermeiden [20]

¹⁶ d. h. hohen Ortsfrequenzen (vgl. Abschnitt 2.3.3); Nadelstreifenanzüge bspw. erscheinen so bunt und scharfe Farb-Kanten fein gemustert [43]

2.6 Bildqualität

Das Hauptziel der Videotechnik besteht meistens in der Verbesserung der Bildqualität. Keimel [46] wählt dafür die in [47] genannte Definition als QoE (Quality of experience):

the degree of delight or annoyance of the user of an application or service [47].

Die Beurteilung ist stark vom Betrachter und den Umständen abhängig. Individuelle Maßstäbe und Erwartungen sowie Betrachtungsdistanz oder Umgebungslicht können diese beeinflussen. Es bieten sich zwei Methoden zur Messung an: subjektive Tests oder objektive Maße [14].

2.6.1 Objektive Qualitätsmaße

Da subjektive Tests zeitintensiv und aufwändig sind, werden sie häufig durch objektive Qualitätsmaße ersetzt, die diese auf Basis objektiv messbarer Parameter vorhersagen sollen. Das am weitesten verbreitete Maß, der PSNR (peak signal-to-noise ratio), stellt eine schnelle und mathematisch einfach beschreibbare Rechengröße dar. Er modelliert nach Ansicht Winklers [14] die subjektive Qualität jedoch nicht ausreichend.

2.6.2 Subjektive Tests

In subjektiven Experimenten müssen Betrachter unter definierten Testbedingungen zu dargebotenen Bildsequenzen Entscheidungen treffen. Winkler [14] kategorisiert diese im Sinne:

1. einer Anpassung, d. h. dass der Beobachter den Stimulus entsprechend vorgegebener Klassifikationen wählt – z. B. eine Amplituden-Schwelle oder einen Stimulus anpasst (vgl. experimentelle Farbabgleiche in Abschnitt 2.1.2) – oder umgekehrt
2. einer Beurteilung: Der Reiz ist zu klassifizieren, z. B. anhand zweier Entscheidungsalternativen (z. B. „Ja“ oder „Nein“) oder einer Bewertungsskala.

Erstere sind laut Winkler leichter umzusetzen, aber nicht ausreichend für eine visuelle Qualitätsbestimmung. Zur zweiten Kategorie zählen die empfohlenen, psychophysikalischen subjektiven Tests in ITU-R REC. BT.500-11¹⁷ [48] oder ITU-R REC. P.910¹⁸ [49].

Eine Gruppe von mindestens 15 repräsentativen, normalsichtigen¹⁹ Betrachtern [48] bewertet unter definierten Umgebungsbedingungen Videosequenzen. In der Trainingsphase vorab ist

¹⁷ anzuwenden bei Fernsehbildern

¹⁸ konzipiert für Multimedia-Anwendungen wie Videotelefonie oder Video-Speicherungs- und -Wiedergabe-systeme [49]

¹⁹ zu prüfen mit einer Snellen- oder Landolt-Tafel und Farbtafel, bspw. von Ishara

eine Testsequenz mit begleitender Erklärung zur Aufgabenstellung und Methodik zu zeigen. Die nachfolgenden, einzelnen Tests (*sessions*) sollten aus einer zufälligen Reihenfolge aller möglichen Kombinationen von Sequenzen und Modifikationsparametern²⁰ bestehen und maximal 30 min dauern. Die ITU-R [48] gibt zur Bewertung der Qualität (*quality assessment*) oder Beeinträchtigung (*impairment assessment*) verschiedene Skalen an (siehe Tabelle A.1), die meist semantische Bezeichnungen auf Zahlenwerte abbilden.

Die folgende Übersicht stellt mögliche Methoden sowie ihre Anwendungsfälle gegenüber [14].

1. Bewertung mit Referenz:
 - a) DSCQS (Double Stimulus Continuous Quality Scale): Beobachter sehen in zufälliger Reihenfolge paarweise eine Referenz und die modifizierte Sequenz für ca. 10 s, bei Wunsch auch wiederholt, und bewerten beide auf einer kontinuierlichen Qualitäts-skala mit fünf definierten Stufen von „schlecht“ bis „exzellent“ (siehe Tabelle A.1). Der Test ist sensitiv für kleine Veränderungen und eignet sich für ähnliche Bilder.
 - b) DSIS (Double Stimulus Impairment Scale)²¹: Wie beim DSCQS wird ein Paar gezeigt, jedoch die Referenz immer zuerst. Danach erfolgt die Bewertung z. B. der Qualitätsminderung auf einer fünfstufigen „impairment scale“ (siehe Tabelle A.1) von „sehr störend“ bis „nicht wahrnehmbar“. Eingesetzt wird die Methode vor allem bei deutlich erkennbaren Störungen.
2. Kontinuierliche Bewertung:
 - a) SSCQE (Single Stimulus Continuous Quality Evaluation): Längere Sequenzen von mindestens fünf Minuten Dauer werden kontinuierlich mithilfe eines Reglers beurteilt.
 - b) SDSCE (Simultaneous Double Stimulus for Continuous Evaluation): Wenn die Wiedergabetreue zu evaluieren ist, wird zusätzlich zur Methode der SSCQE simultan ein Referenzbild entweder seitlich nebeneinander auf einem oder zwei angeglichenen Monitoren gezeigt [48].
3. Alternative Methoden – SS (Single Stimulus) und SC (Stimulus Comparison): Sequenzen werden einzeln (SS) oder relativ zueinander (SC) entweder:
 - a) Kategorien zugeordnet. Diese können
 - i. semantische Bezeichnungen tragen (ACJ (Adjectival Categorical Judgement)) und auf einer fünfstufigen Skala (Tabelle A.1) oder Vergleichsskala Tabelle A.2 stehen oder
 - ii. numerisch durch Werte bspw. auf einer NCS (Numerical Category Scale) definiert sein.

²⁰ Die Parameter sollten die gesamte Qualitätsbeeinflussung in gleichen Abständen abdecken [48]

²¹ auch EBU (European Broadcasting Union)-Methode genannt; entspricht dem DCR (Degradation Category Rating) aus P.910

- b) Oder sie werden nichtkategorisch bewertet (NCJ (Non-Categorical Judgement)), d. h. umgekehrt wird
 - i. eine sprachliche Bewertung wie beim ACJ oder auf einer analogen Skala beim *continous scaling* vergeben oder
 - ii. ein Zahlenwert zugeteilt.
- c) PC (Pair Comparison) gemäß ITU-R REC. P.910 [49]: Aus zwei, durch unterschiedliche Systeme veränderten Sequenzen wählt der Proband einen Favoriten bzw. bewertet sie zueinander. Dadurch können geringe Unterschiede der Systeme festgestellt werden [49].
- d) ACR (Absolute Category Rating) entspricht der SS-Methode: Einzelne Sequenzen ohne Referenz werden auf einer fünfstufigen Skala bewertet. Das Verfahren benötigt am wenigsten Zeit und eignet sich für einfache Qualifizierungstests [49].

Kritisiert wird an Single-Rating-Verfahren²², dass Störungen zeitlich nicht unbedingt gleichmäßig verteilt sind und psychologisch am Ende stärker gewichtet werden, weswegen die Dauer zu limitieren ist. Demgegenüber wird häufig die Beeinträchtigung normaler Sehgewohnheiten durch die Wiedergabe eines Referenzbilds bemängelt [14].

Alternativ bietet sich die SSCQE an – besonders, wenn die Bildqualität zeitlich oder szenenabhängig variiert [48]. Fehler können aber aufgrund einer verzögerten Bewertung oder Abhängigkeiten vom Szeneninhalt auftreten [14].

Betrachtungsbedingungen

Tabelle 2.6 nennt verschiedene Vorgaben an die Testumgebung hinsichtlich Beleuchtung, Betrachtungsabstand und Bildschirmanforderungen. Die Heiumgebung (*home environment*) simuliert typische Fernseh-Konsum-Bedingungen, während die Laborumgebung (*laboratory environment*) kritischere Systemanforderungen stellt [50], [51]. Die ITU-R-Standards ITU-R BT.710-4 [51] für HDTV und am aktuellsten ITU-R BT.2022 [50] für Flachbildschirme (HDTV und SDTV) lösen die für CRTs konzipierten Betrachtungsbedingungen aus ITU-R BT.500 [48] ab. Für Multimedia-Anwendungen sind zum Testen ferner ITU-R BT.910 [49] und ITU-R BT.1788 [52] anzuwenden.

Auswahl des Testmaterials

Das untersuchte Bewertungsproblem bestimmt maßgeblich die Auswahl des Testmaterials. Das Sortiment sollte eine hohe Qualität und eine große Bandbreite an Merkmalen aufweisen,

²² d. h., die Szene wird nach der Wiedergabe einmalig bewertet

wie beispielsweise:

- sowohl Videos als auch statische Bilder
- Bilder mit betonten Farben
- grafische (Text-)Elemente und Animationen
- örtliche Details
- verschiedene künstlerische Effekte und Kameratechniken [53]
- Videos mit Schnitten (*scene cuts*) [54]
- Bewegung in verschiedenen Geschwindigkeiten und Richtungen

Geeignete Empfehlungen nennen [55] und [56]. Der Einbezug kritischer Sequenzen, die eine angemessene Herausforderung für die Anwendung darstellen, ist notwendig für eine Interpretation der Ergebnisse [48]. Allerdings sollten vor allem anstößige, polarisierende, ablenkende oder kontroverse Inhalte, die die Bewertung manipulieren könnten, gemieden werden. Zudem sind Verbreitungsrechte einzuhalten [55] [53].

Auswertung subjektiver Tests

Zur statistischen Auswertung wird der MOS (Mean Opinion Score), ein Mittelwert der Bewertungen, für jede Testbedingung und -sequenz ermittelt:

$$\bar{u}_{jkr} = \frac{1}{N} \sum_{i=1}^N u_{ijkr} \quad (2.58)$$

mit der Bewertung u_{ijkr} des Betrachters i (von insgesamt N) für die Testbedingung j , die Sequenz k und die Wiederholung r .

Ferner ist jeweils ein Konfidenzintervall anzugeben. Empfohlen wird ein 95 %-Niveau [48]:

$$[\bar{u}_{jkr} - \delta_{jkr}, \bar{u}_{jkr} + \delta_{jkr}] \quad (2.59)$$

$$\text{mit: } \delta_{jkr} = 1,96 \frac{s_{jkr}}{\sqrt{N}} \quad (2.60)$$

und der Standardabweichung:

$$s_{jkr} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (\bar{u}_{jkr} - u_{ijkr})^2} \quad (2.61)$$

Tabelle 2.6: vorgegebene Betrachtungsbedingungen von ITU-R-Standards für subjektive Tests (Laborverhältnisse) und im Experiment umgesetzte Werte

Betrachtungsbedingung	Vorgaben der Standards			durchgeföhrter Test	
	BT.500	BT.710-4	BT.2022	absolut	relativ
maximale Bildschirmleuchtdichte (cd/m^223	–	150 bis 250	70 bis 250	210	–
dazu im Verhältnis die Leuchtdichte:					
• des schwarzen Bildschirms ²⁴	$\leq 0,02$			ca. $0,09 \text{ cd}/\text{m}^2$	0,0004
• eines Schwarzbildes im total abgedunkelten Raum	ca. 0,01		–	$\leq 0,001 \text{ cd}/\text{m}^2$	$\leq 5 \cdot 10^{-6}$
• des Hintergrunds ²⁵	ca. 0,15			ca. $30 \text{ cd}/\text{m}^2$	ca. 0,143
Farbwert des Hintergrunds ²⁵	D65			ca. D65 bzw. D50	
Bildschirmhelligkeits- und Konstrasteinstellung	mithilfe PLUGE ²⁶	–	mithilfe PLUGE	konform	
sonstige Raumbeleuchtung	gering				
Betrachtungsdistanz	PVD (Tabelle)	$3H^{27}$	PVD (Graph) oder DVD (Tabelle)	$3H$	
maximaler horizontaler Betrachtungswinkel	± 30		–	konform	
Prüfung der Bildschirmauflösung	empfohlen	–	empfohlen	erfolgreich	
Bildschirmdiagonale	–	1,4 m (55 in)	–	55 in	–

2.7 Stand der VLC-Technik

Auch wenn VLC auf eine lange Geschichte zurückblicken kann, ist sie seit ca. 15 Jahren besonders durch den Fortschritt der LED-Technik und durch die Suche nach alternativen Übertragungsmethoden im Fokus theoretischer und kommerzieller Überlegungen [57]. Das Projekt DaViD soll zuerst kontextuell und danach im Detail erörtert werden.

2.7.1 Systematische Einordnung

Im Gesamtsystem nach Abb. 2.23 ordnet sich VLC neben den Anwendungen unsichtbarer Lichtspektren – IR (Infrarot) und UV (Ultraviolett) – der OWC (Optical Wireless Communication)

²³ gemessen mit einem Photometer in einem kleinen weißen Fenster mit dem maximalem Signalpegel von 255

²⁴ für ITU-R BT.2022[50] Messung der CR nach ITU-R BT.815 – siehe Gleichung (2.39)

²⁵ nach ITU-R BT.710-4 einzuhalten für einen Winkelbereich von 53° Höhe und 83° Weite

²⁶ entsprechend ITU-R BT.814 und ITU-R BT.815

²⁷ Bildschirmhöhe H

bzw. OFS (Optical Free-space Communication)²⁸ als Freiraum-Variante der allgemeineren Datenübertragung per Licht – der OC (optical communication) – unter. Parallel können leitergebundene Technologien wie z. B. die Glasfaserkabel-Übertragung und HF-Nutzungen mit den bekanntesten drahtlosen Vertretern WLAN (Wireless Local Area Network), Bluetooth und ZigBee klassifiziert werden.

VLC kann auf historische Methoden wie Signalfeuer oder das Photophon²⁹ zurückgeführt werden. Neuere Anwendungen nutzen Laser oder Dioden sowie vorhandene Licht-Infrastrukturen. In Anlehnung an IEEE 802.11 für WLAN („Wi-Fi“) wurde für ein drahtloses Netzwerk auf Basis von Licht von Haas die Bezeichnung „Li-Fi (light fidelity)“ [60] sowie von der IEEE (Institute of Electrical and Electronics Engineers) der Standard 802.15.7-2011 eingeführt [61], [62]. Die Vorschläge zur Erhöhung der Reichweite und Datenrate von üblichen Methoden mit SISO (Single Input Single Output) und in der Frequenz beschränkter LEDs (Light Emitting Diodes) inkludieren Multiplexing über parallele Kanäle und MIMO (Multiple Input Multiple Output) durch mehrere Sender und Empfänger z. B. mithilfe LED-Arrays [63] oder der DCC³⁰. Auch DaViD stellt eine DCC her, die allerdings die simultane Videopräsentation nicht stören soll.

Mit den genannten Techniken beschäftigt sich hauptsächlich die Nachrichten- bzw. Kommunikationstechnik, aber auch die Photonik bzw. Optoelektronik, Bild- und Signalverarbeitung oder Hochfrequenztechnik steuern theoretische Grundlagen bei.

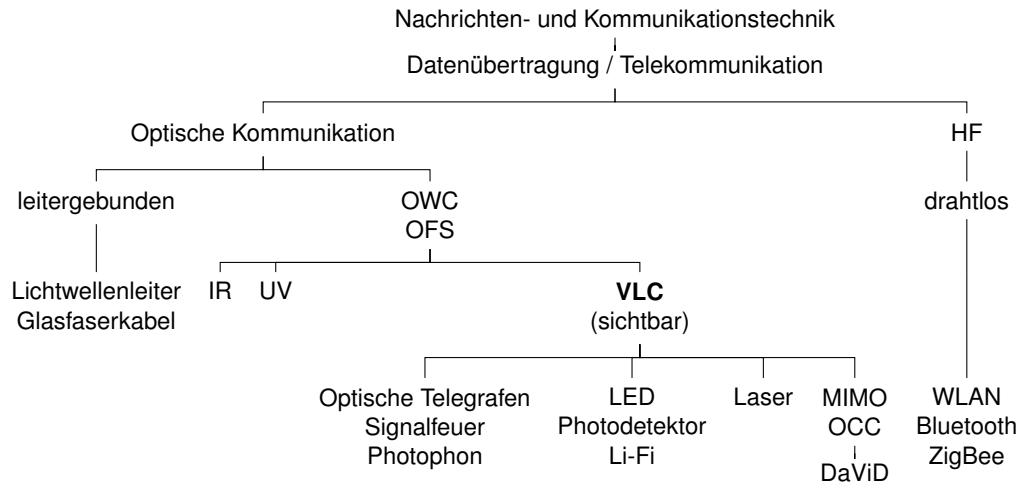


Abbildung 2.23: Systematik der VLC

²⁸ auch: FSO (Free-space Optical Communication); manche Autoren bezeichnen mit OFS abgrenzend zur OWC nur Punkt-zu-Punkt-Verbindungen im Außenbereich [58]

²⁹ Eine Erfindung von Bell, Tonsignale über modulierte Lichtstrahlen zu senden [59]

³⁰ auch: SCC (Screen Camera Communication) oder allgemein bei Einsatz von Kamerasensoren OCC (Optical Camera Communication)

2.7.2 Nutzung der VLC

VLC weist gegenüber vergleichbaren Funktechnologien folgende Vor- und Nachteile vor:

Vorteile

- Nutzbarkeit der frei verfügbaren Bandbreite des Licht-Spektrums [64]
- keine Interferenz z. B. [65]
 - mit sensibler Elektronik in Krankenhäusern oder Flugzeugen
 - mit anderen Teilnehmern durch lokale Abschirmung; kleine Zellen ermöglichen hohe Datendichten
- Möglichkeit des Voll duplex [65]
- Datenraten bis 100 Mbit/s [64]
- Verwendung vorhandener Infrastrukturen und Hardware wie öffentlicher oder häuslicher Beleuchtung, Verkehrslichter, Mobilgeräte, Fahrzeug-Lichter und TV-/Computer-Bildschirme weitestgehend ohne Einschränkung der primären Funktionalität sowie dadurch:
 - ökonomisch und ökologisch effiziente Wiederverwendung ohne zusätzlichen Installations- oder Platzbedarf
 - häufig räumliche oder inhaltliche Verknüpfung mit der Kommunikations-Anwendung [66]
 - Profitieren von der Verbreitung und Entwicklung moderner Technologien wie LEDs oder Smartphones [67] [66]
 - einfacher und kostengünstiger Systemaufbau mit weniger notwendigen Komponenten als für die HF-Technik [67]
- keine Gefahr für die Gesundheit
- Datensicherheit durch Abschirmbarkeit der Übertragung und direkter Verbindung
- simple Verbindmöglichkeit ohne Pairing (im Gegensatz zu Bluetooth) [68]

Nachteile

- begrenzter Anwendungsbereich
- häufige notwendige Sichtverbindung
- geringere Datenrate als WLAN [64]
- Störanfälligkeit gegenüber Umwelteinflüssen wie veränderlichen Beleuchtungssituativen im Außenbereich mit folglich relativ hohen Fehlerraten [66]
- zwingend aktive Beleuchtung für die leistungsfähige Nutzung
- teils Nichtlinearitäten optischer Systeme [69]

- wenig Standardisierung³¹

Anwendungsszenarien

Für folgende Anwendungen eignet sich VLC demgemäß unter anderem:

- als Alternative oder heterogene Ergänzung zu WLAN [67], [71] und 5G [2], [72]
- für die Fahrzeug- oder Nahbereich-Kommunikation
- zur präzisen Indoor-Positionsbestimmung [73]
- in der Medizintechnik (z. B. für Body Sensor Networks [74])
- für die öffentliche oder private Datenverbreitung wie Audio- oder Bild-Medien, Informationen oder Werbung und Dateien auch ohne Anmeldung z. B. für den einmaligen Gebrauch in Geschäften oder Museen; als Kiosk System
- in der gemeinsamen Nutzung von Smartphone-Kameras und -Apps
- in der Industrie: Kommunikation im Internet of Things oder zwischen Maschinen; Bereitstellung von Prozess-Informationen oder Instruktionen [64], [75]

2.7.3 Verwandte Projekte

Exemplarische Projekte mit vergleichbaren Technologien wie in DaViD werden im Folgenden beleuchtet:

- *COBRA*:
Die Autoren erörtern die Datenübertragung zwischen Smartphones mit Farb-Barcodes, die für eine schnelle Verarbeitung und Robustheit gegen Unschärfe adaptiv optimiert werden. Typischerweise möglich sind so 90 kbit/s bis 175 kbit/s [65], [68], [76].
- *SoftLight*:
Die Idee von *COBRA* wird durch die Modulation von *U* und *V* und durch einen Fountain-Code (*rateless erasure code*) zur Adaption der Datenrate an den Übertragungskanal ergänzt. *SoftLight* leistet bis zu etwa 300 kbit/s [77], [78].
- *LightSync*:
Das System synchronisiert die Kommunikation zwischen Kamera und Display, solange der Empfänger mindestens die halbe Framerate nutzt. Der Rolling-Shutter-Effekt wird mithilfe Farbbalken herausgefiltert und ein FEC (forward error correction)³² eingesetzt, um unvollkommene Frames wiederzugewinnen [68].

³¹ Die überarbeitete Fassung IEEE 802.15.7r1 soll beispielsweise auch OCC einbeziehen [70]

³² Löschkodierung (*erasure code*, EC)

- *HiLight*:

Die Daten werden durch Frequenzmodulation des Alpha-Kanals³³ im Video versteckt übertragen und mit einer Smartphone-Kamera empfangen. Erzielt werden bis zu 180 kbit/s Durchsatz [79]–[81].

- *DisCo*:

DisCo stellt eine DCC bei simultaner Videoanzeige dar. Durch zeitliche Helligkeitsmodulation wird ein Datenmuster im Bild eingebettet und mittels einer Rolling-Shutter-Kamera mittels verschiedener Belichtungen decodiert [82], [83].

- *VRCodes*:

Auch *VRCodes* nutzt den Rolling-Shutter-Effekt zur Decodierung von im Video versteckter Barcodes und die Flimmer-Verschmelzungsfrequenz zur Mischung der komplementär gefärbten Pixel aus [68], [81], [84].

- [85] und [86] schlagen eine bildabhängige Auswahl der Pixelblöcke in Kanten oder texturierten Bereichen vor, um dort die Daten bei zeitlich differentieller Luminanzmodulation³⁴ möglichst flimmerfrei zu platzieren.

- „A Pixelated MIMO Wireless Optical Communication System“:

Hranilovic und Kschischang beschreiben eine MIMO-Kommunikation zwischen verschiedenen mehrdimensionalen optischen Geräten zum Senden³⁵ und Empfangen³⁶ von sequentiellen 2D-Codebildern [87]. OFDM (orthogonal frequency-division multiplexing) gemäß WLAN-Standard wird verwendet. Die Autoren berücksichtigen zwar die Gradiation kommerzieller Displays, setzen allerdings eine Kompensation durch Vorverzerrung voraus [87] – so wie auch eine korrekte Ausrichtung und Synchronisierung [88].

- *PixNet*:

Die SCC von *PixNet* nutzt eine Luminanzmodulation und wie Hranilovic und Kschischang OFDM zur Codierung der Daten, um robuster gegen perspektivische Verzerrungen, Unschärfe und Umgebungslichteinflüssen zu werden [73], [88].

- *InFrame*:

Wie bei DaViD wird die hochfrequente Luminanzmodulation von Blöcken ausgenutzt, um Daten im eigentlichen Video versteckt zu integrieren und so eine duale Kommunikation sowohl für Anwender als auch Gerät umzusetzen. Das Konzept beinhaltet transiente Frame-Übergänge, eine hierarchische Daten-Frame-Struktur und Codemultiplexverfahren und erzielt bis zu 300 kbit/s³⁷ [81].

- *Temporal Psychovisual Modulation (TPVM)*:

Wu und Shu [66], [89] stellen vergleichbar mit DaViD eine SCC mit zeitlich differentieller Modulation vor, die im Originalbild verborgen wird, sodass das Ansehen des Videos

³³ entspricht der Transparenz des Bildes

³⁴ vgl. Abschnitt 2.7.4

³⁵ z. B. LCDs, LED-Arrays, Laserstrukturen, OLEDs

³⁶ Photodioden-Arrays, CCD- oder CMOS-Sensoren

³⁷ bei 120 Hz Framerate

unbeeinträchtigt bleibt. Es werden quadratische Pixel-Blöcke für eine vereinfachte Decodierung moduliert, weiche Signalübergänge und eine ausreichende Framerate eingeplant, die Signalmarkierungen zur Synchronisierung und Datenredundanz zum Fehlerschutz verwendet. Das System erreicht bis zu 1 Mbit/s bei akzeptablen Fehlerraten [64].

- Steganographie:

Auch die Steganographie versucht Informationen in Signalen zu verbergen – z. B. mittels Wasserzeichen zur Authentizitäts- oder Integritätsprüfung. Allerdings steht häufig die Robustheit gegen Modifizierungen im Vordergrund und eine Echtzeitverarbeitung ist nicht notwendig [81], [90]. Einige Methoden schlagen die Modulation der Chrominanz vor [91]–[94] und berücksichtigen teilweise Gammaverzerrungen [95].

2.7.4 Systembeschreibung von DaViD

DaViD [64], [96]–[98] beruht auf der unauffälligen Datenübertragung von einem Bildschirm als Sender zu einer Videokamera als Empfänger, wobei die gleichzeitige Videowiedergabe für den Konsumenten unbeeinträchtigt bleiben soll. Das Projekt repräsentiert eine Form der DCC, erweitert das MIMO-Konzept jedoch durch Multiplexing um einen weiteren Teilnehmer, dem Betrachter, zu einer dualen *Display Camera Communication* und *Display User Communication*. Die Anwendung ist für kurze Sichtverbindungen (*line-of-sight*) gedacht.

Von den umgesetzten Methoden soll an dieser Stelle nur die zeitlich-differentielle Modulation³⁸ erläutert werden. In der Form der Chrominanzmodulation ermöglicht sie die geringste Sichtbarkeit bei robuster Decodierbarkeit und den Einsatz von Kameras geringerer Auflösung.

Übertragungsstrecke

Abbildung 2.24 stellt die Übertragungsstrecke von DaViD schematisch dar. Die Eingangsdaten werden zuerst den Pixeln an den horizontalen und vertikalen Positionen i und j des $N_x \times N_y$ großen Bildschirms zugeordnet: Nach Teilung des Datenstroms in Pakete $d(l) \in \{-1, 1\}$ der maximalen Länge L pro Frame m , wird jedes Bit auf einen Block von $B_x \times B_y$ Pixeln

³⁸ vgl. [64] für die örtlich-differentielle Modulation

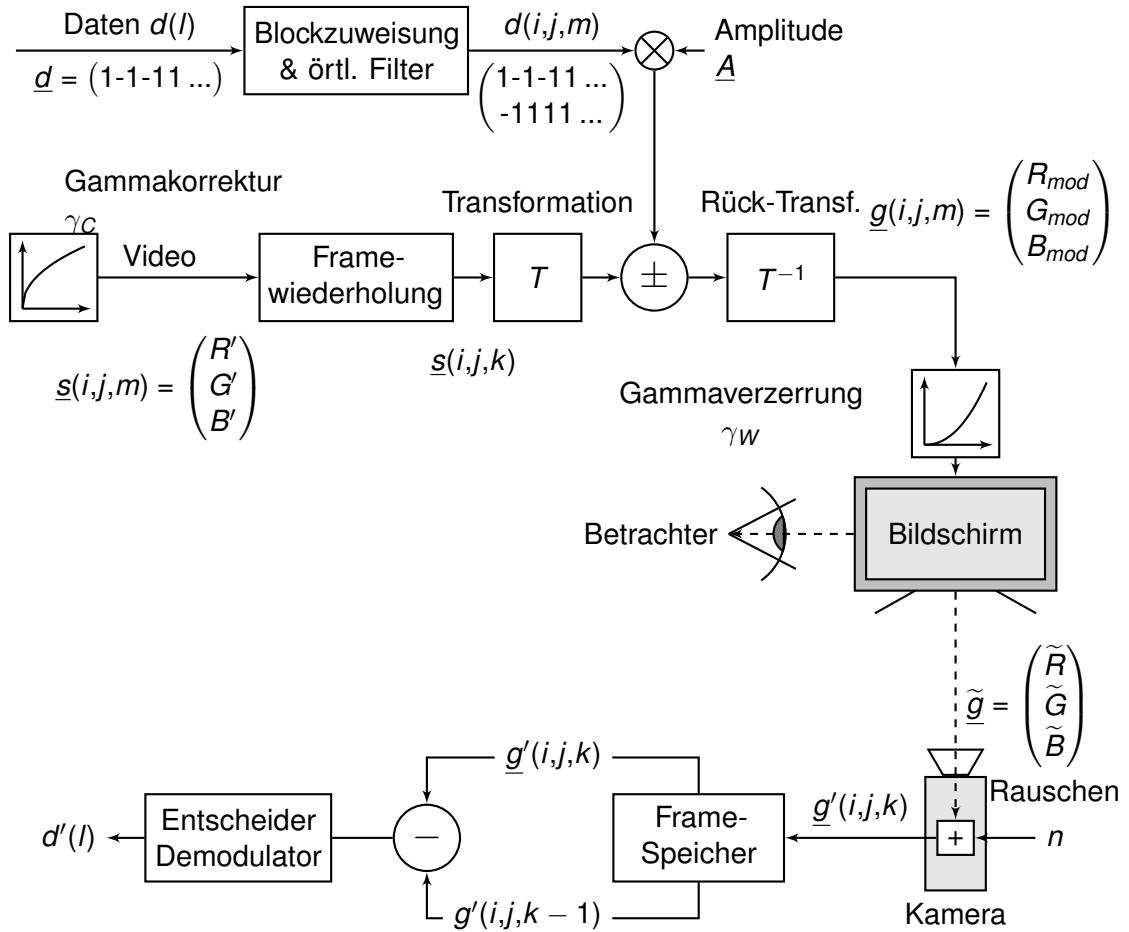


Abbildung 2.24: Blockdiagramm der Übertragungsstrecke

abgebildet:

$$d(l) \rightarrow d(i,j,m) \quad 0 \leq l < L \quad (2.62)$$

$$L = \left\lfloor \frac{N_x}{B_x} \right\rfloor \cdot \left\lfloor \frac{N_y}{B_y} \right\rfloor \quad (2.63)$$

$$i = (l \cdot B_x) \bmod (N_x) + r_x, \quad r_x = 0 \dots (B_x - 1) \quad (2.64)$$

$$j = \left\lfloor l / \left\lfloor \frac{N_x}{B_x} \right\rfloor \right\rfloor \cdot B_y + r_y, \quad r_y = 0 \dots (B_y - 1) \quad (2.65)$$

Die resultierende Unterabtastung kommt der geringeren notwendigen Kamera-Auflösungen und einer robusteren Übertragung entgegen. Nachfolgend wird von der experimentell bewährten Blockgröße $B = B_x = B_y = 4$ sowie von keinen zusätzlichen örtlichen Filtern ausgegangen.

Das Video-Signal \underline{s} liegt bereits als R'G'B'-Tripel³⁹ nach Gammakorrektur z. B. durch die Aufnahme-Kamera vor und kann nicht beeinflusst werden. Für die zeitlich-differentielle Modulation müssen die Frames wiederholt werden:

$$\underline{s}(i,j,k+1) = \underline{s}(i,j,k), \quad k = 2m, m \in \mathbb{Z} \quad (2.66)$$

Modulationsverfahren

Bei der zeitlich-differentiellen Modulation werden die Datenmuster dem Videosignal überlagert, ein Frame mit der positiven und eines mit der negativen Amplitude \underline{A} . Wird \underline{A} für alle Komponenten RGB gleich gewählt, entspricht dies der Modulation der Luminanz bzw. Helligkeit.

Verschiedene Kanal-Amplituden ergeben die Farbmodulation mit dem Vorteil, einer geringeren Wahrnehmbarkeit für Menschen aufgrund der reduzierten Empfindlichkeit für chromatische Kontraste (siehe Abschnitt 2.3.5). Vor allem der Blau-Kanal qualifiziert sich an dieser Stelle, da er am wenigsten zur Gesamthelligkeit beiträgt, aber auch die Nutzung der Farbdifferenzsignale U und V empfiehlt sich, insofern die Luminanz Y konstant bleibt. Zudem steigert Multiplexing die Datenrate, indem die Nachrichten separaten Farb-Kanälen zugewiesen werden. Die Modulation kann mit der Matrix T der eventuell notwendigen Farbraum-Transformation⁴⁰ allgemein folgendermaßen beschrieben werden:

$$\underline{g}_{RGB}(i,j,k) = [R_{mod} \quad G_{mod} \quad B_{mod}] \quad (2.67)$$

$$= \begin{cases} T^{-1} \cdot [T \cdot \underline{s}_{RGB}(i,j,m) + d(i,j,m) \cdot \underline{A}] & k = 2m \\ T^{-1} \cdot [T \cdot \underline{s}_{RGB}(i,j,m) - d(i,j,m) \cdot \underline{A}] & k = 2m + 1 \end{cases} \quad (2.68)$$

$$= \underline{s}_{RGB}(i,j,m) \pm d(i,j,m) \cdot \underbrace{T^{-1} \cdot \underline{A}}_{\underline{A}_{RGB}} \quad (2.69)$$

Für die Modulation in RGB gilt $T = T^{-1} = 1$ und für die Luminanzmodulation speziell $A_R = A_G = A_B$. Chrominanz-Amplituden A_U und A_V lassen sich in äquivalente RGB-Werte überführen – so z. B. für $A_U = A_V = 9$ und sRGB mit Gleichung (2.14):

$$\begin{aligned} A_{RGB} &= T_{Y'UV \rightarrow RGB, BT.709} \cdot A_{YUV} \\ \begin{bmatrix} A_R \\ A_G \\ A_B \end{bmatrix} &= \begin{bmatrix} 1 & 0 & 1,575 \\ 1 & -0,187 & -0,468 \\ 1 & 1,856 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 9 \\ 9 \end{bmatrix} = \begin{bmatrix} 14,172 \\ -5,899 \\ 16,701 \end{bmatrix} \end{aligned} \quad (2.70)$$

³⁹ im Folgenden wird auf die Kennzeichnung der nichtlinearen gammavorverzerrten Größen wie R' außer zur Hervorhebung verzichtet

⁴⁰ die Transformation ist abhängig vom Farbraum des Monitors; für die eingesetzten Modelle kann der sRGB-Farbraum angewendet werden[33], [98]

$$\underline{g}_{RGB}(i,j,k) = \begin{bmatrix} R_{mod} \\ G_{mod} \\ B_{mod} \end{bmatrix} = T^{-1} \cdot \left[T \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \pm \begin{bmatrix} 0 \\ d_U A_U \\ d_V A_V \end{bmatrix} \right] \quad (2.71)$$

$$= \begin{bmatrix} R \\ G \\ B \end{bmatrix} \underbrace{\pm T^{-1} \cdot \begin{bmatrix} 0 \\ d_U A_U \\ d_V A_V \end{bmatrix}}_{dA(i,j,m)} \quad (2.72)$$

Demodulation

Ein Bildschirm mit ausreichender Auflösung gibt das modulierte, jedoch gammaverzerrte Signal \tilde{g} wieder (siehe Abschnitt 2.5 und zur späteren genaueren Analyse Abschnitt 3.1). Der menschliche Betrachter nimmt das Video idealerweise unverfälscht wahr, sofern die Wiedergabefrequenz hoch sowie die Blockgröße und Modulationsamplitude gering genug sind, während die Kamera⁴¹ die Datenmuster miterfasst. Diese mit Rauschen n behafteten Bilder \underline{g}' werden perspektivisch entzerrt, gefiltert und abgetastet, wobei eine zeitliche Synchronisierung herzustellen ist. Je zwei Frames eines Paares werden voneinander subtrahiert, um daraus nach einer Schwellenwertentscheidung für R und B die Daten d' wiederzugewinnen:

$$\Delta = \underline{g}'(i,j,k) - \underline{g}'(i,j,k-1) \quad (2.73)$$

$$\begin{aligned} &= \underline{s}'(i,j,m) + d'(i,j,m) \cdot \underline{A}'_{RGB} \\ &\quad - (\underline{s}'(i,j,m) - d'(i,j,m) \cdot \underline{A}'_{RGB}) \\ &= 2 \cdot d'(i,j,m) \cdot \underline{A}'_{RGB} \end{aligned} \quad (2.74)$$

⁴¹ beispielsweise eines Smartphones