



Grundlagen der Multimediaetechnik

Menschliche Wahrnehmung

29.10.2021, Prof. Dr. Enkelejda Kasneci



Termine und Themen

22.10.2021	Einführung
29.10.2021	Menschliche Wahrnehmung – visuell, akustisch, haptisch, ...
05.11.2021	Informationstheorie, Textcodierung und -komprimierung
12.11.2021	Bildverbesserung
19.11.2021	Bildanalyse
26.11.2021	Grundlagen der Signalverarbeitung
03.12.2021	Bildkomprimierung
10.12.2021	Videokomprimierung
17.12.2022	Audiokomprimierung
14.01.2022	Videoanalyse
21.01.2022	Dynamic Time Warping
28.01.2022	Gestenanalyse
04.02.2022	Tiefendatengenerierung
11.02.2022	FAQ mit den Tutoren
15.02.2022	Klausur (noch nicht bestätigt)



Menschliche Wahrnehmung

- Visuell
- Akustisch

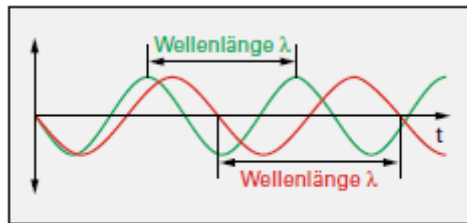
Visuelle Wahrnehmung



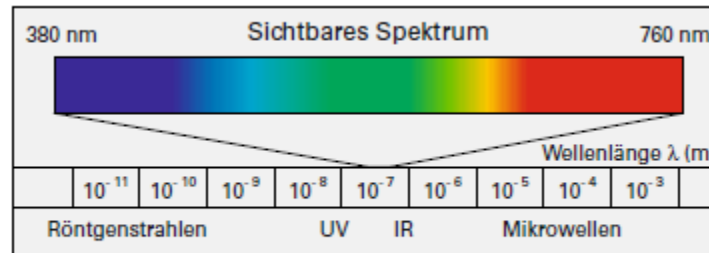


Was ist „Farbe“

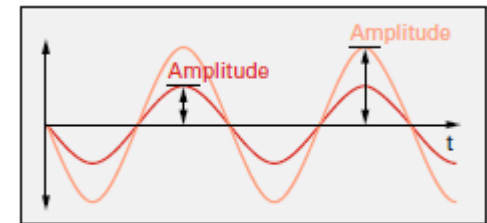
- **Licht** ist eine **Welle** als auch **Teilchen** (Welle-Teilchen-Dualismus)
- **Farbe** wird durch die **Wellenlänge (Frequenz)** erzeugt
- **Helligkeit** durch die **Amplitude** der Welle



Wellenlänge



Elektromagnetisches
Spektrum



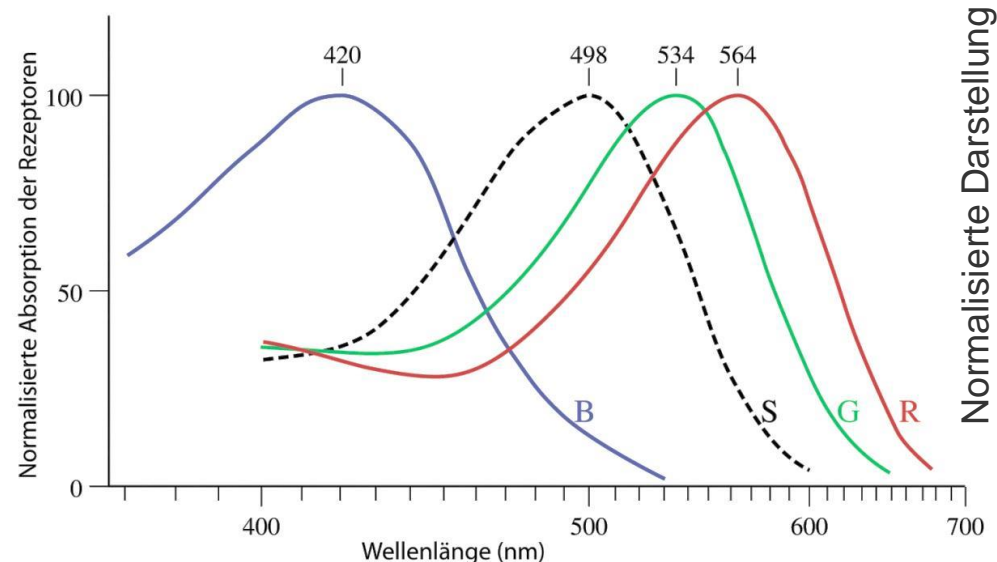
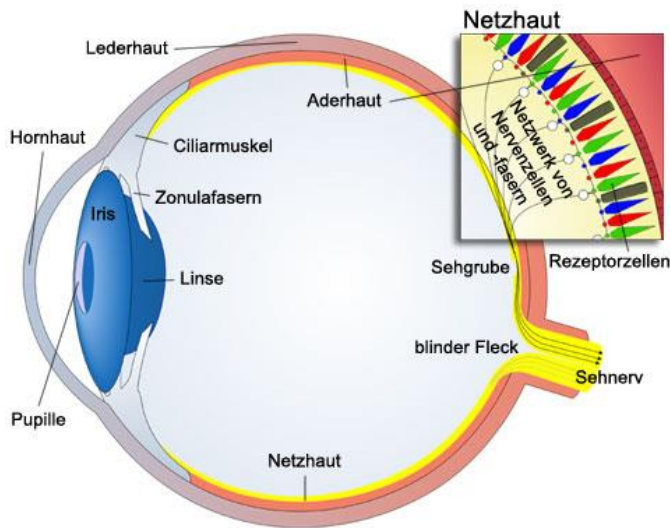
Amplitude

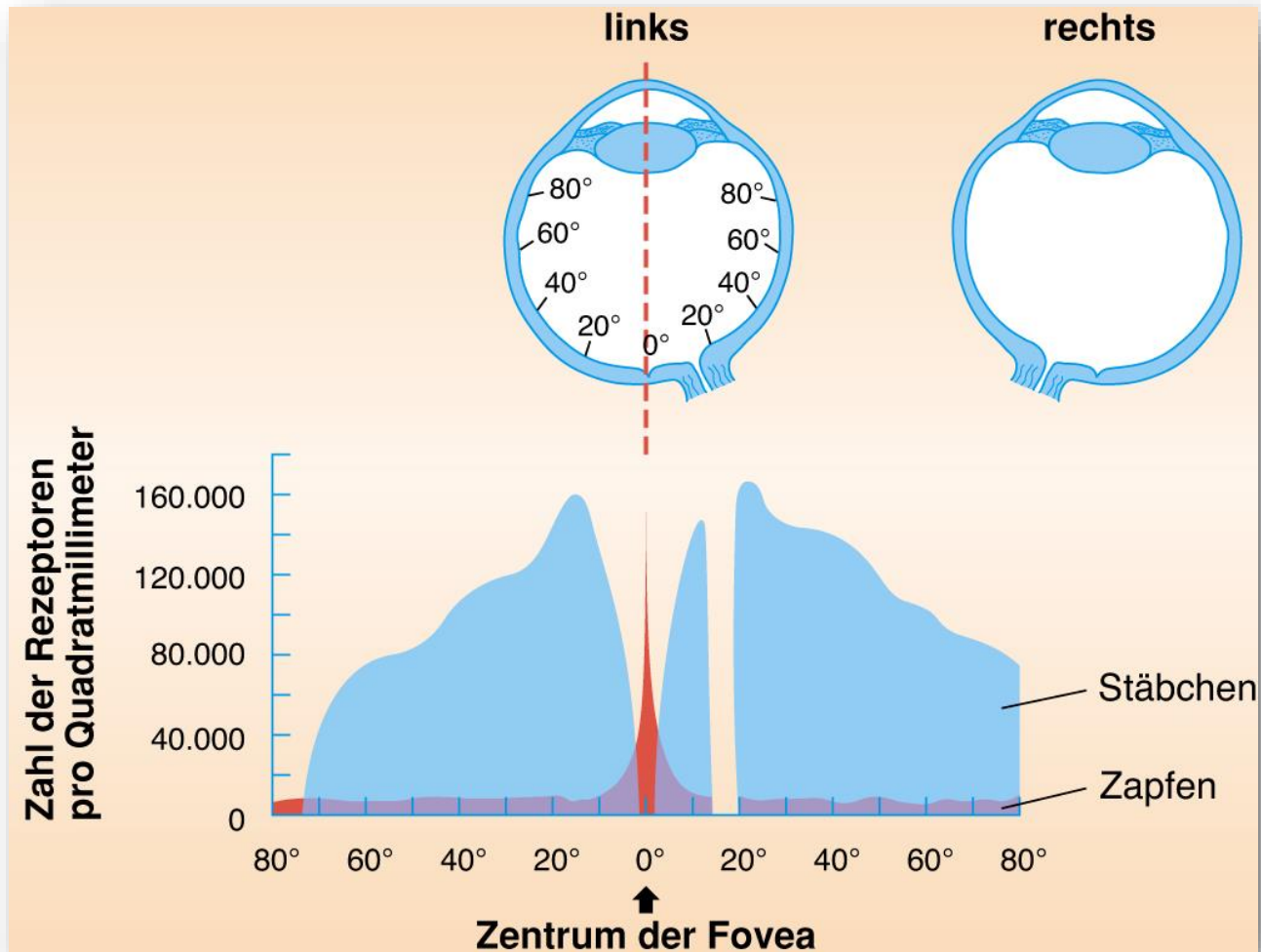
- Farben können durch Mischung (Interferenzen) anderer Farben erzeugt werden
- Das menschliche Auge nimmt Rot, Grün und Blau wahr



Farbsehen

- Menschliche visuelle Wahrnehmung erfolgt über drei unterschiedlich für das wahrgenommene Spektrum empfindliche Zelltypen.
 - drei **Zapfentypen** zur Erkennung von **Farben** (Wellenlänge, höhere Reizschwelle)
 - **Stäbchen** zur **Helligkeitsdetektion** (Sättigung bereits bei Dämmerung)
- Der Wahrnehmung fast aller sichtbaren Wellenlängen kann durch Kombination von drei Wellenlängen angenähert werden.





[Speckmann, E. J., Hescheler, J., & Köhling, R. (Eds.). (2012).
Repetitorium Physiologie. Elsevier, Urban&Fischer Verlag.]

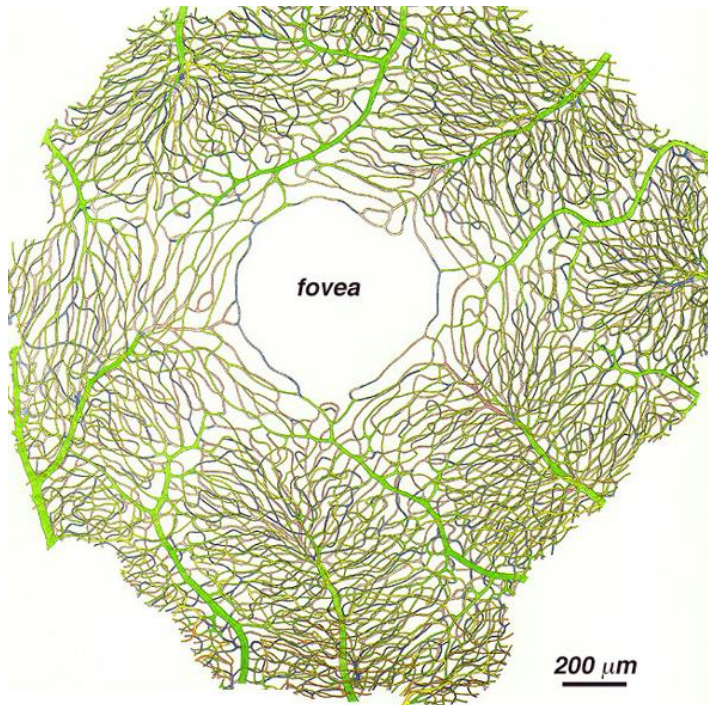
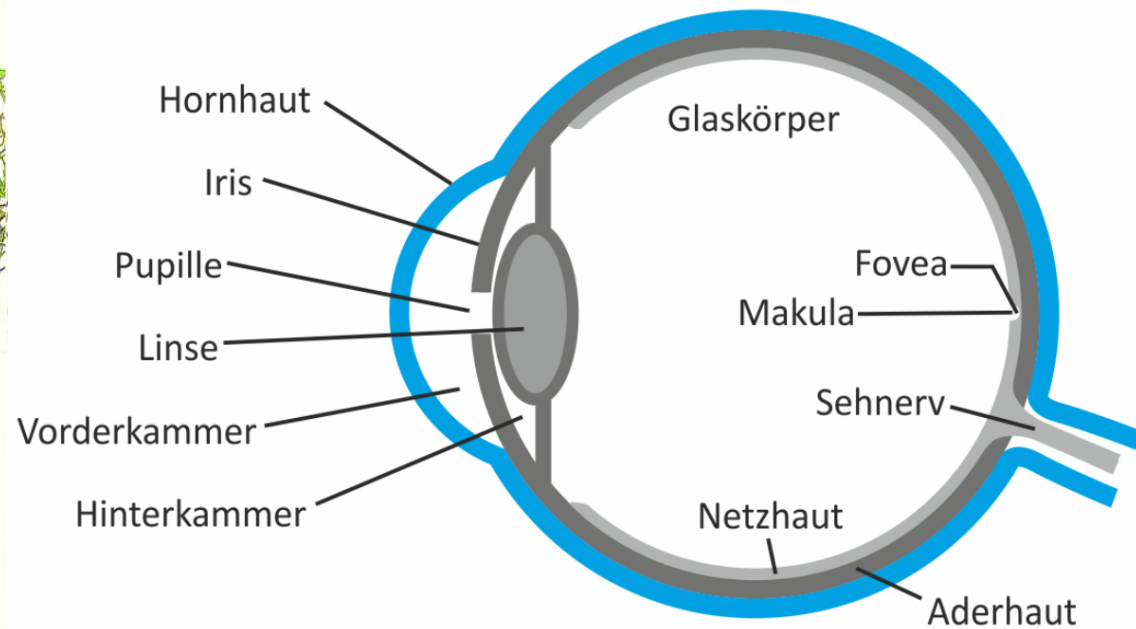
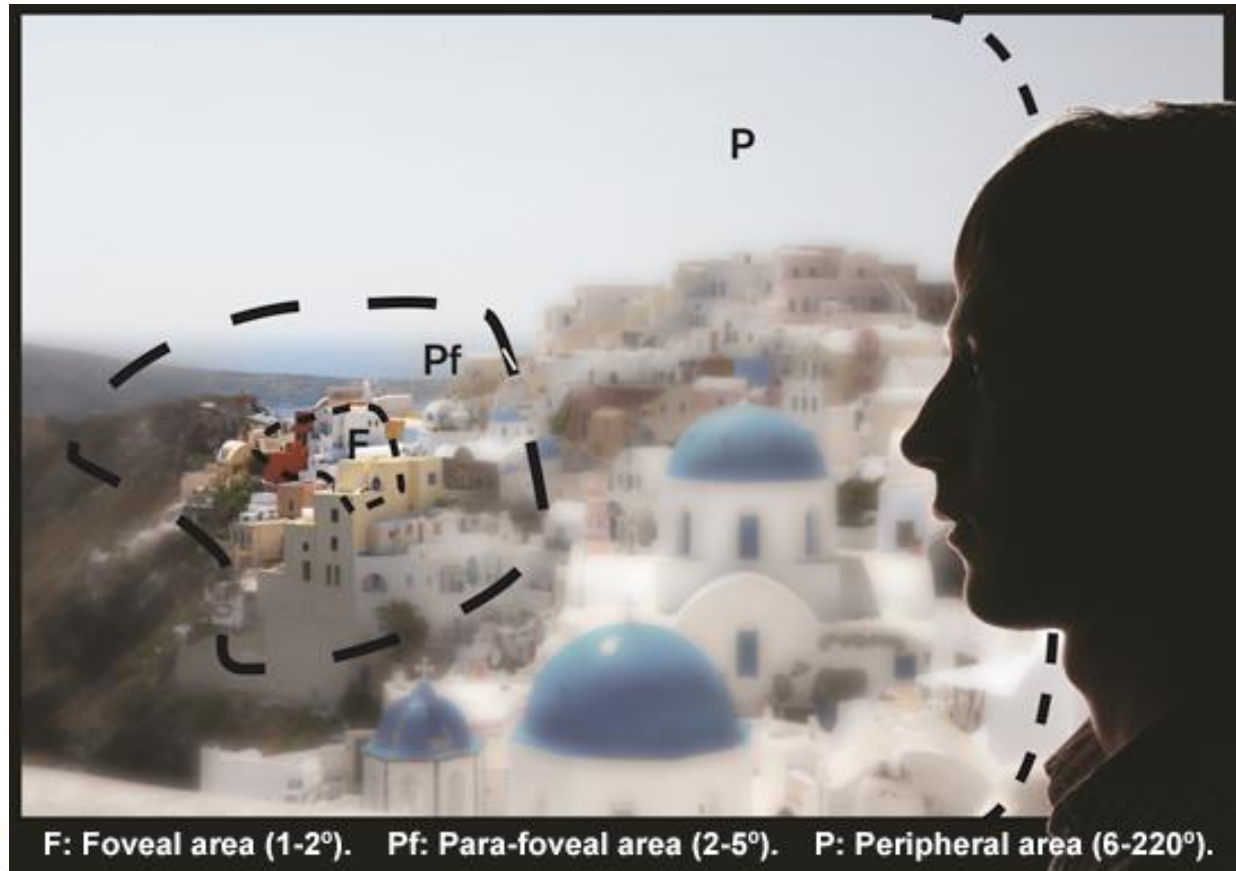
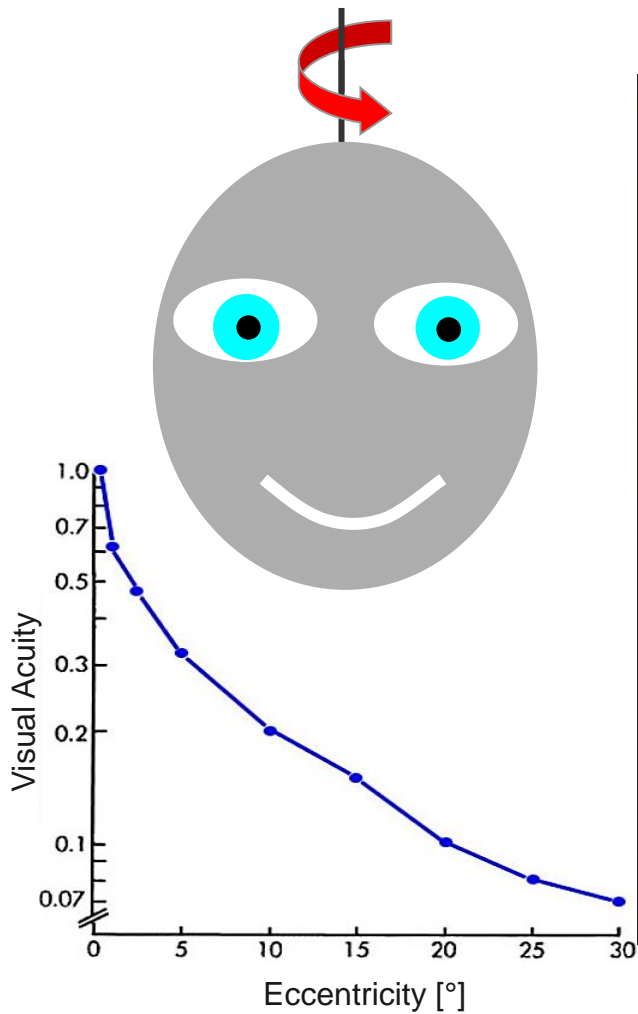


Fig. 20. Diagram of the retinal vasculature around the fovea in the rhesus monkey derived from more than 80 microscope fields.



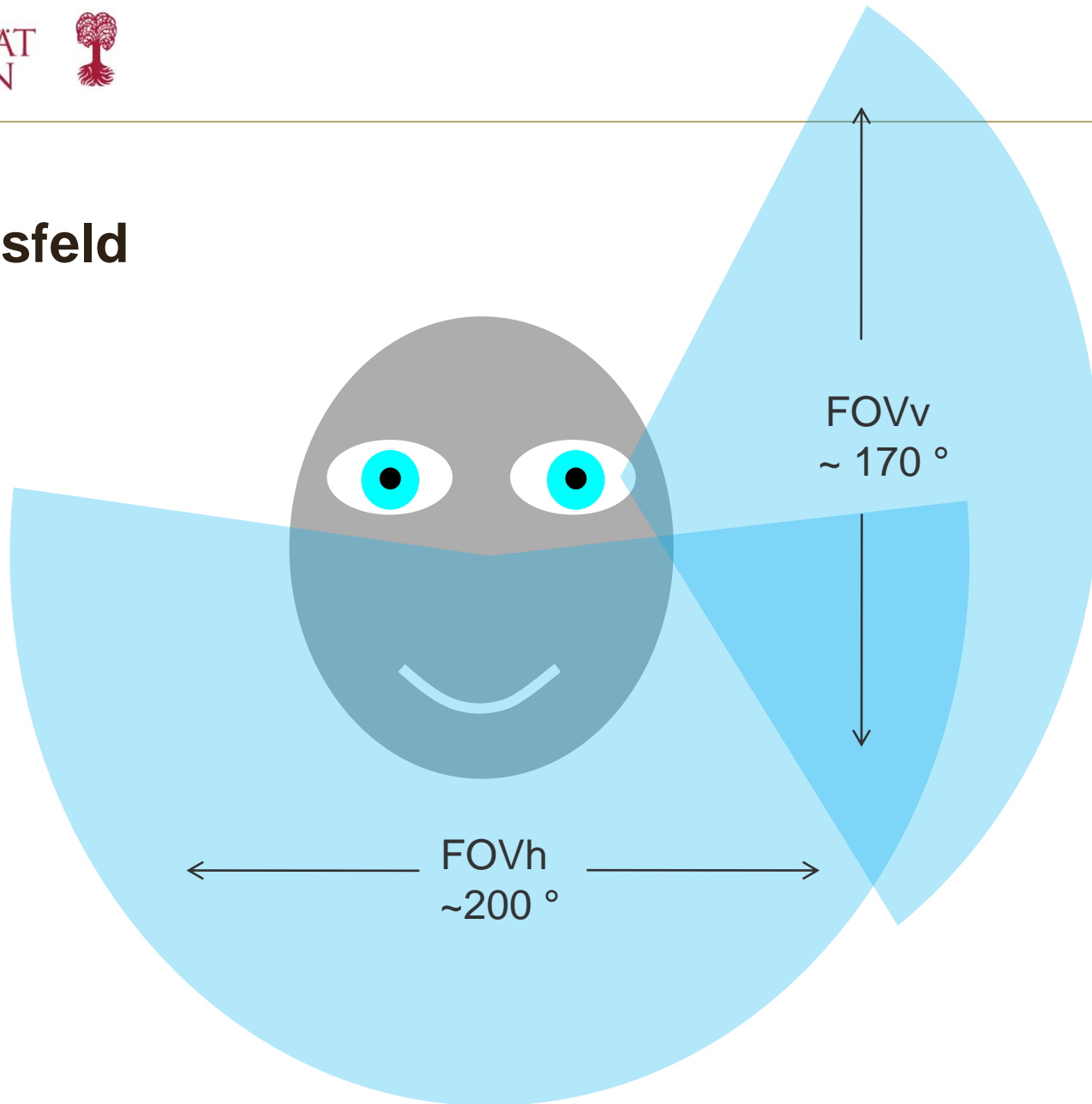
[<http://webvision.med.utah.edu>]



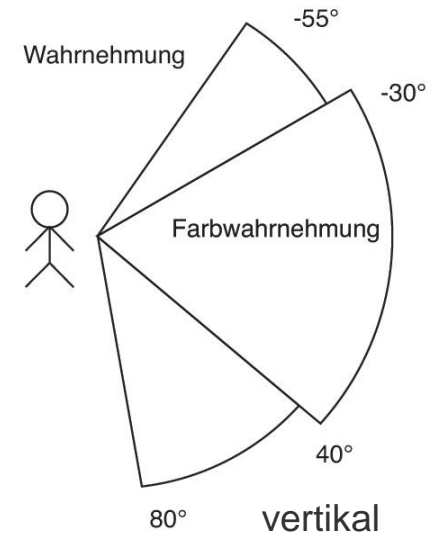
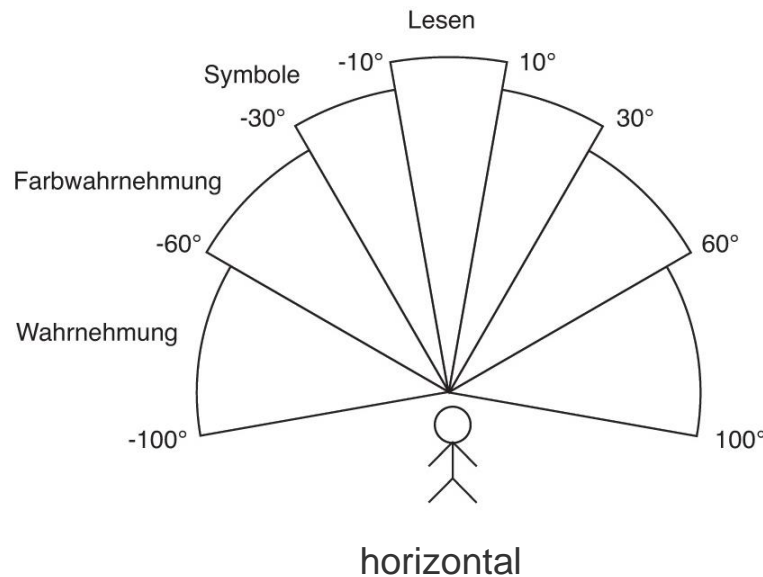
[Bergstrom & Schall. *Eye tracking in user experience*. Elsevier, 2014]



Gesichtsfeld



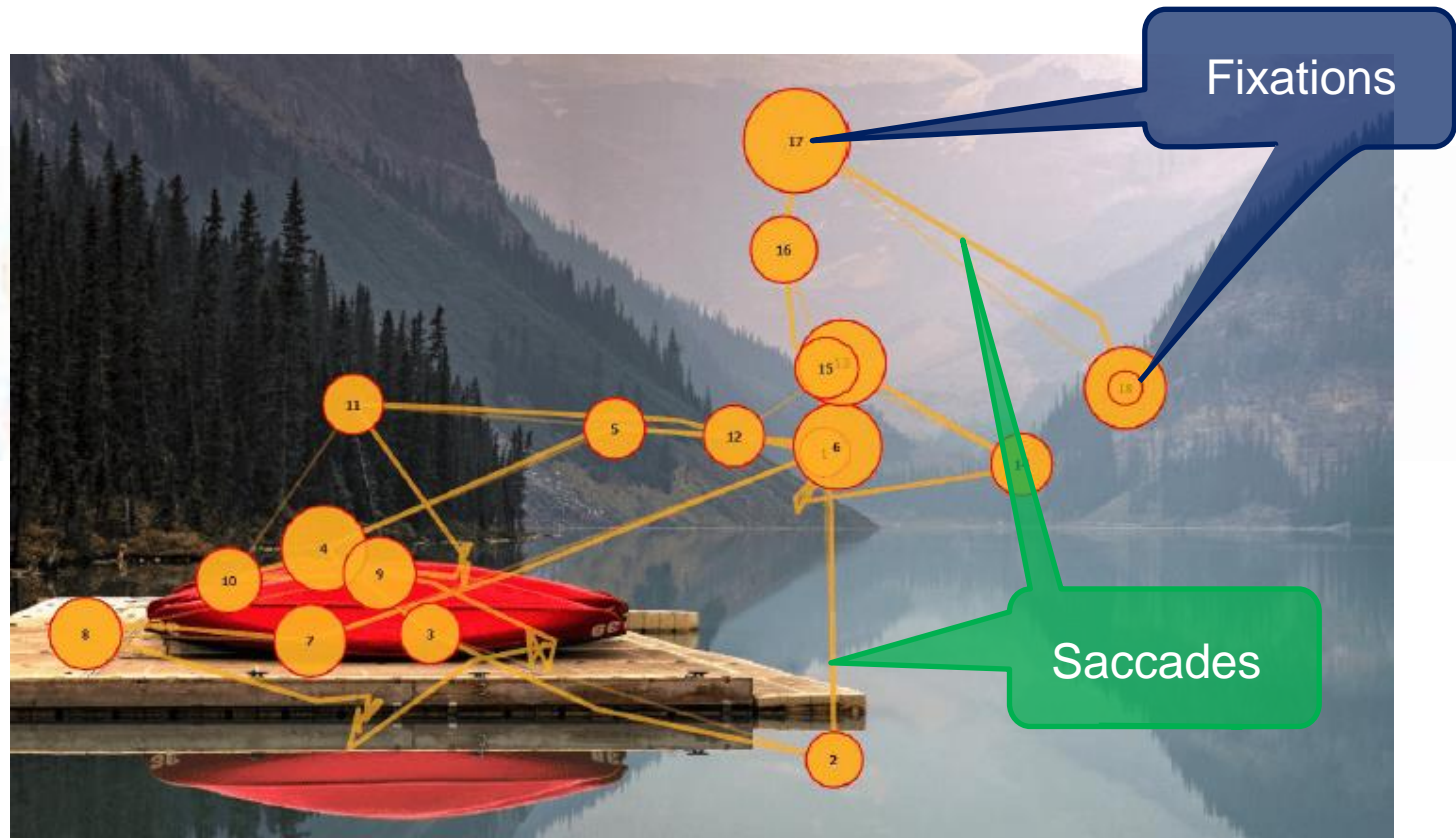
Sehfeld des Menschen



- **Höchste Auflösung: Mitte des Sehfelds** (Sehzentrum)
(3 mm Objekt kann in 10 m Entfernung erkannt werden)
- **Sehzentrum: viele Zapfen, wenig Stäbchen**
→ Nachts keine hohe Auflösung



Eye movements enable visual perception



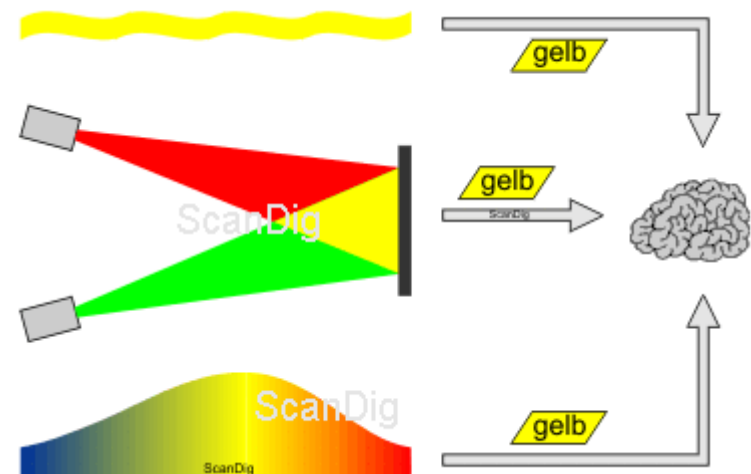
[<https://imotions.com/blog/7-terms-metrics-eye-tracking/>]



Farbwahrnehmung

- **Metamerie** bezeichnet den Effekt, dass unterschiedliche Farbspektren denselben Farbeindruck beim Menschen hervorrufen können
 - **Monochromatisches Licht** (Licht einer Wellenlänge)
 - **Mischung** mehrerer **monochromatischer Lichtquellen**
 - **Polychromatisches Licht** (Licht vieler Wellenlängen)
- Der Mensch kann Licht nicht in seine Spektralfarben zerlegen
(*im Gegensatz zu Tönen*)

Der Wahrnehmung fast aller sichtbaren Wellenlängen kann durch Kombination von drei Wellenlängen angenähert werden!





Farbmischung

- Unterscheidung zwischen
 - **Additive Farbmischung**
 - Aktive lichterzeugende Medien (Monitore)
 - Rot, Grün und Blau → RGB
 - **Spektrale Intensitäten** der einzelnen Lichtkomponenten **werden addiert**
 - **Subtraktive Farbmischung**
 - Ausgabe auf Druckern (reflektierende Medien)
 - Cyan, Magenta und Gelb (Yellow) → CMY
 - **Spektrale Intensitäten** werden entsprechend dem Farbwert aus dem weißen Licht **entfernt**
→ additive Mischung von Absorptionskomponenten





Farbsysteme

- **Digitale Medien: RGB**

(bei 8 bit pro Farbkanal)

- 255: maximaler Anteil
- 0: kein Farbanteil

- **Printmedien: CMY(K)**

- 100: maximaler Farbanteil
- 0: keine Farbe → Papierfarbe
- K oder Key (Schwarz) zur Verstärkung des Kontrastes oder um klares Schwarz darzustellen
- Ebenso kommen im Printbereich Sonderfarben zum Einsatz



CC-License <http://www.flickr.com/photos/blueridgekitties/>

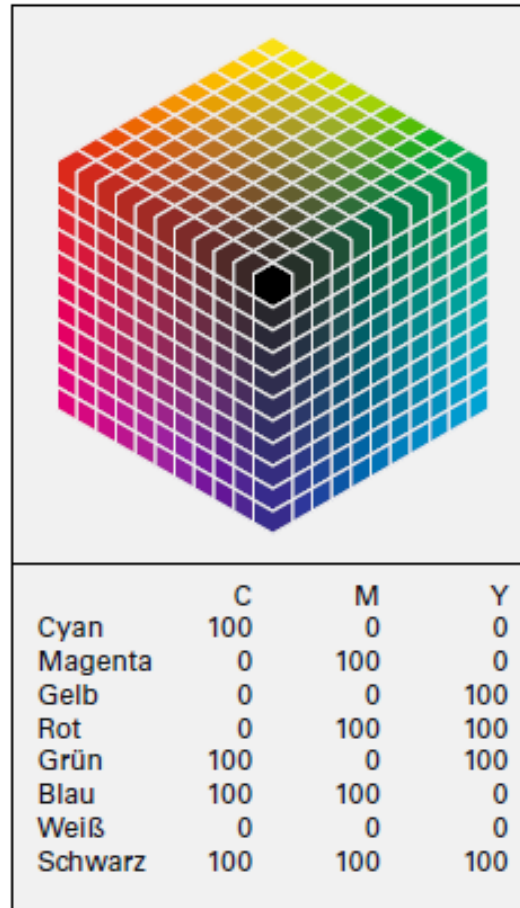


Abbildung: Heidelberger Druckmaschinen AG



Farbsysteme – CMY

- 100: maximaler Farbanteil, 0: keine Farbe → Papierfarbe

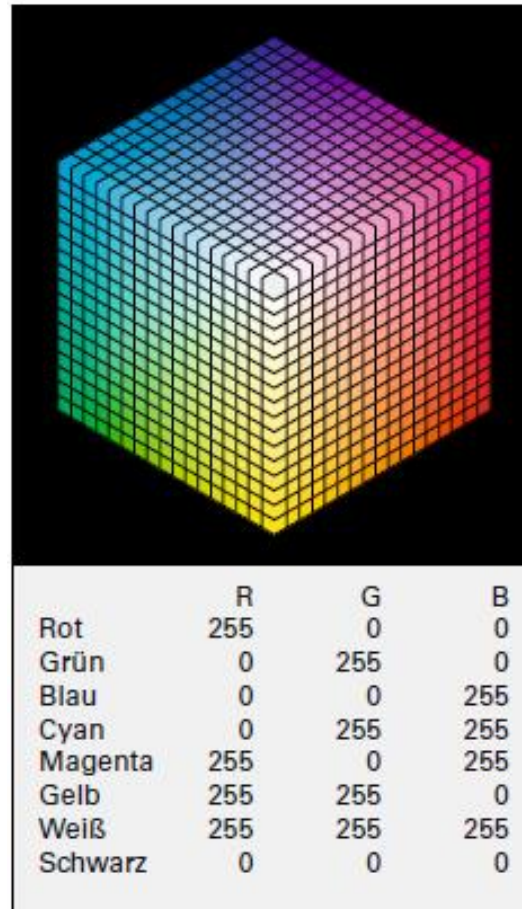


Bildquelle: Wikipedia



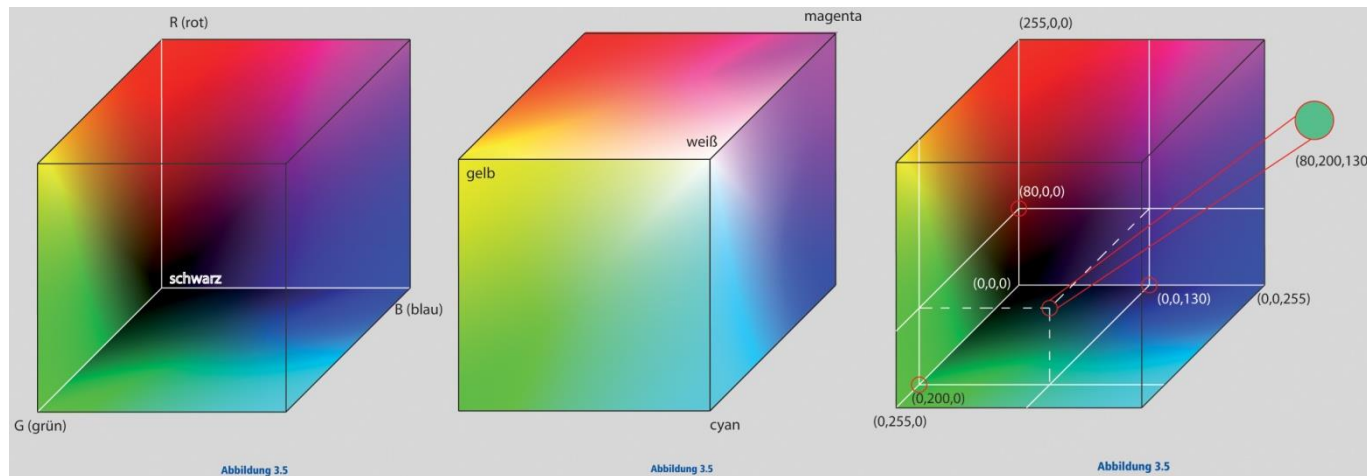
Farbsysteme – RGB

- 255: maximaler Anteil, 0: kein Farbanteil (bei 8 bit pro Farbkanal)

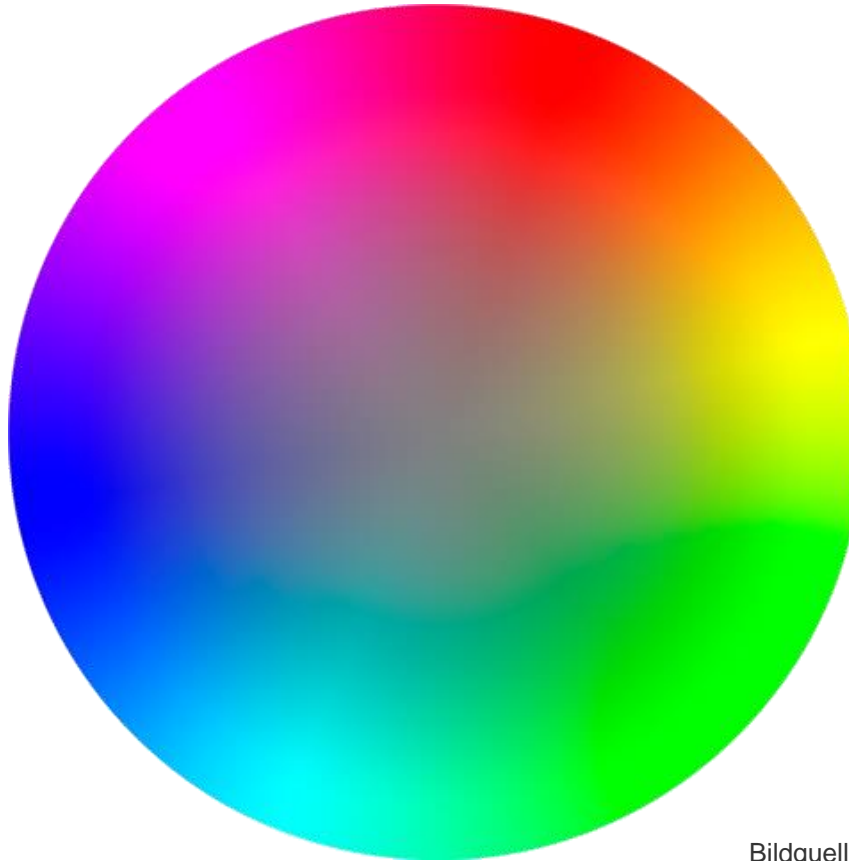


Bildquelle: Wikipedia

- **Additive Komposition jeder beliebigen Farbe** durch **Mischung** der **Primärfarben Rot, Gelb** und **Grün**
 - Mischungsverhältnis bestimmt Farbton
 - Weiß $\rightarrow R:G:B=1:1:1$
 - Schwarz $\rightarrow R:G:B=0:0:0$
 - Grau $\rightarrow R:G:B=x:x:x \forall 0 < x < 1$
 - Sattes Gelb: $\rightarrow R:G:B=1:1:0$
 - Darstellung als Farbwürfel



Bildquelle: Wikipedia



Bildquelle: Wikipedia

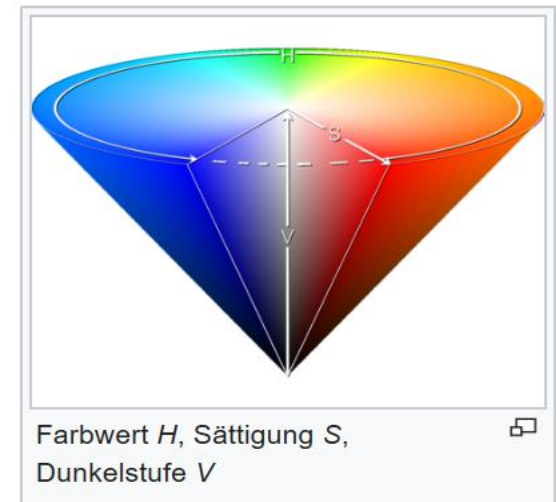
- Die **Farben** werden entsprechend ihrer **Abfolge im Spektrum in einem Farbkreis angeordnet** (entspricht der visuell wahrgenommenen „Sortierung“)
- **Komplementärfarben stehen sich gegenüber**
- **Sättigung nimmt nach außen zu**



Farbdarstellung

HSL/HSB- und HSV-Farbdarstellung

- **HSV** (*Hue/Saturation/Value*)-Model : Entwickelt durch A. R. Smith 1978. Es basiert auf drei Merkmalen von Farben, die für uns besonders intuitiv sind:
 - Farbton (rot, gelb, grün ...)
 - Sättigung (brillant oder matt?)
 - Helligkeit (hell oder dunkel?).
- **HSB**:
 - Hue, Saturation, Brightness (absolute Helligkeit)
- **HSL (auch HLS)**:
 - Hue (Farbton), Saturation (Sättigung), Lightness (relative Helligkeit)

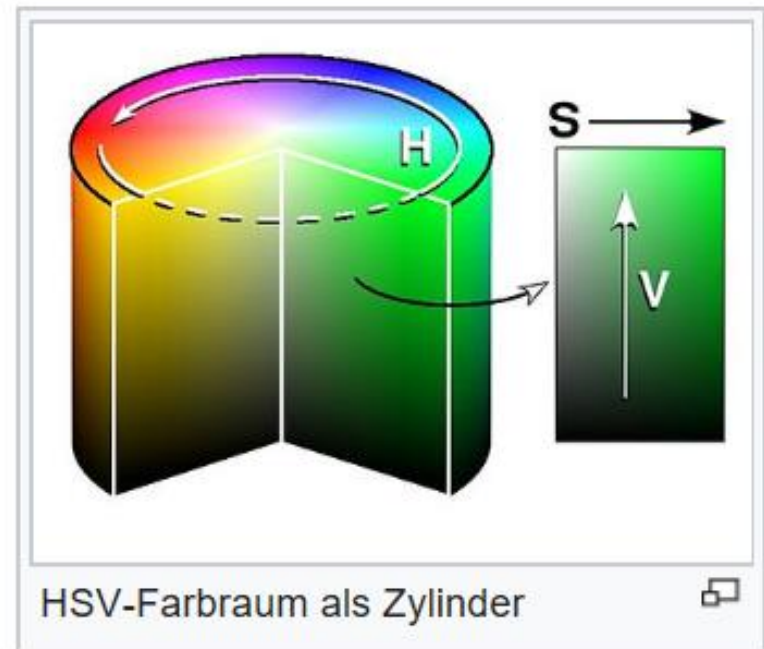
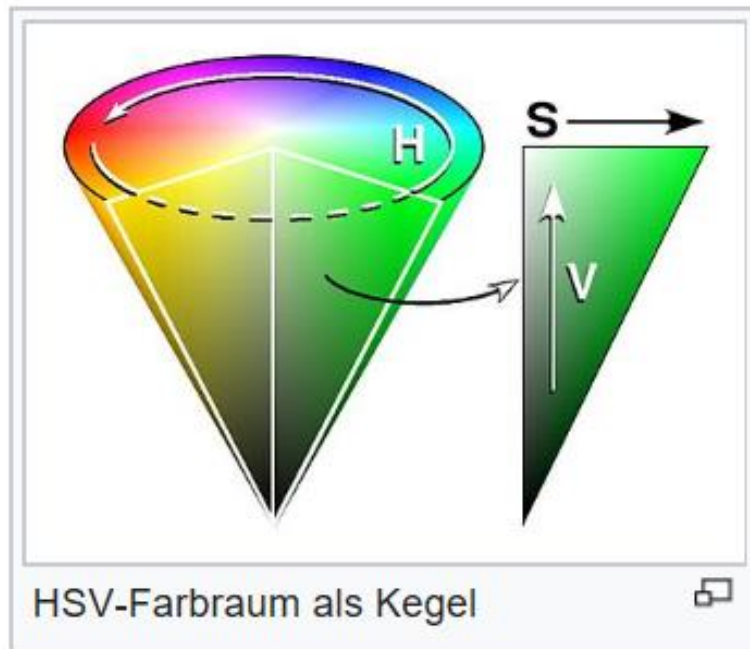


Bildquelle: Wikipedia



Anschauliche Darstellung des HSV Farbraums

- Kegel: Helligkeit nimmt von oben (Farbkreis) nach unten (Spitze = Schwarz) ab.
- Zylinder: Helligkeit nimmt zu einer Fläche abnimmt (anstatt zu einem Punkt).



Bildquelle: Wikipedia



Umrechnung RGB nach HSV/HSL

Vorbedingung: $R, G, B \in [0, 1]$

$$MAX := \max(R, G, B), MIN := \min(R, G, B)$$

$$H := \begin{cases} 0, & \text{falls } MAX = MIN \Leftrightarrow R = G = B \\ 60^\circ \cdot \left(0 + \frac{G-B}{MAX-MIN}\right), & \text{falls } MAX = R \\ 60^\circ \cdot \left(2 + \frac{B-R}{MAX-MIN}\right), & \text{falls } MAX = G \\ 60^\circ \cdot \left(4 + \frac{R-G}{MAX-MIN}\right), & \text{falls } MAX = B \end{cases}$$

falls $H < 0^\circ$ dann $H := H + 360^\circ$

$$S_{HSV} := \begin{cases} 0, & \text{falls } MAX = 0 \Leftrightarrow R = G = B = 0 \\ \frac{MAX-MIN}{MAX}, & \text{sonst} \end{cases}$$

$$S_{HSL} := \begin{cases} 0, & \text{falls } MAX = 0 \Leftrightarrow R = G = B = 0 \\ 0, & \text{falls } MIN = 1 \Leftrightarrow R = G = B = 1 \\ \frac{MAX-MIN}{1-|MAX+MIN-1|}, & \text{sonst} \end{cases}$$

$$V := MAX$$

$$L := \frac{MAX + MIN}{2}$$

Nachbedingung: $H \in [0^\circ, 360^\circ]$, $S, V, L \in [0, 1]$

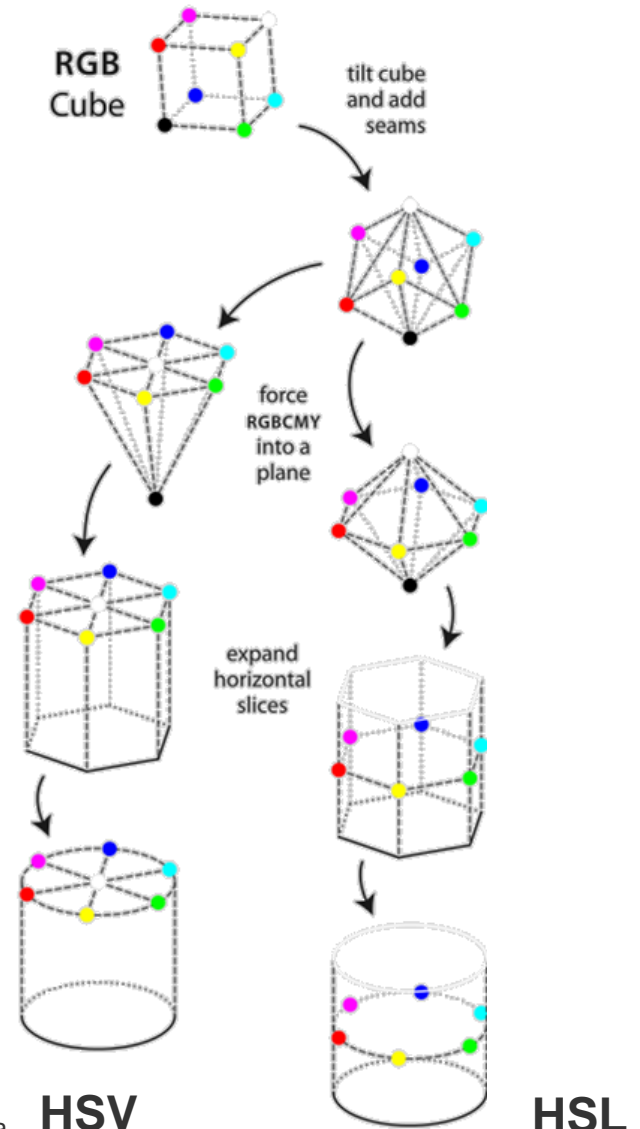
Demo: <https://web.cs.uni-paderborn.de/cgvb/colormaster/web/color-systems/hsl.html>

Bildquelle: Wikipedia



Ableitung von RGB zu H*-Modellen

- Repräsentation des Farbraums mit Polarkoordinaten



Bildquelle: Wikipedia

HSB

HSL

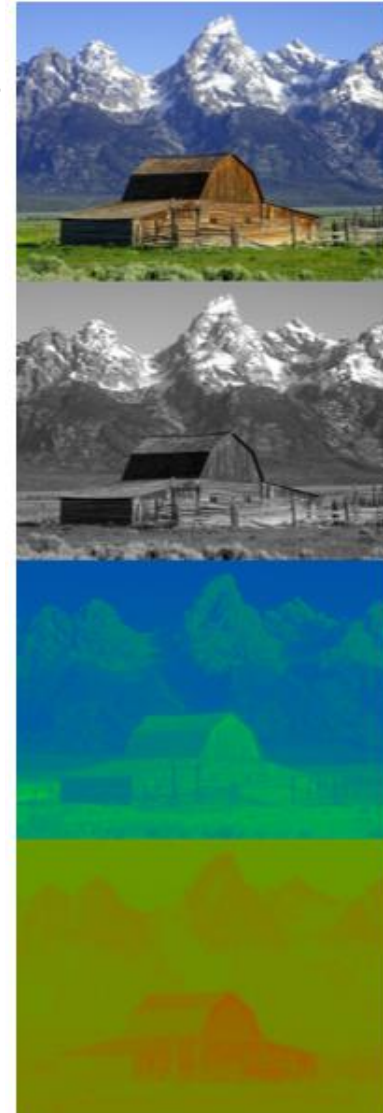


Farbdarstellung

- **Videosignal besteht aus Luminanz- und Chrominanzanteil**
 - **Luminanz:** Pixelhelligkeit
 - **Chrominanz:** Pixelfarbe (Farbton und Sättigung)
- Umkodierung von RGB in Komponentensignal
 - YUV-Kodierung (PAL) → YCbCr-Kodierung (HDTV, DVD)
 - YIQ-Kodierung (NTSC)



- Erzeugen von Helligkeits- und Farbinformation aus RGB
 - $Y = WR \cdot R + WG \cdot G + WB \cdot B$ (Luminanz)
 - $U = U_{max} \cdot (B - Y)/(1 - WB)$ (Chrominanz 1)
 - $V = V_{max} \cdot (R - Y)/(1 - WR)$ (Chrominanz 2)
 - U und V sind Farbdifferenzsignale
 - YUV ist verlustfrei zurück nach RGB transformierbar
- Gewichtungsfaktoren
 - YUV (analog): $WR = 0,299$, $WB = 0,144$
 - YCbCr (digital): $WR = 0,2126$, $WB = 0,0722$
 - $WG = 1 - WR - WB = 0,577$
 - Komponentenskalierung:
 $U_{max} = 0,436$, $V_{max} = 0,615$
- Kompatibilität zu S/W-Geräten (nur Y)
- Einsatz: PAL (YUV), HDTV (YCbCr)



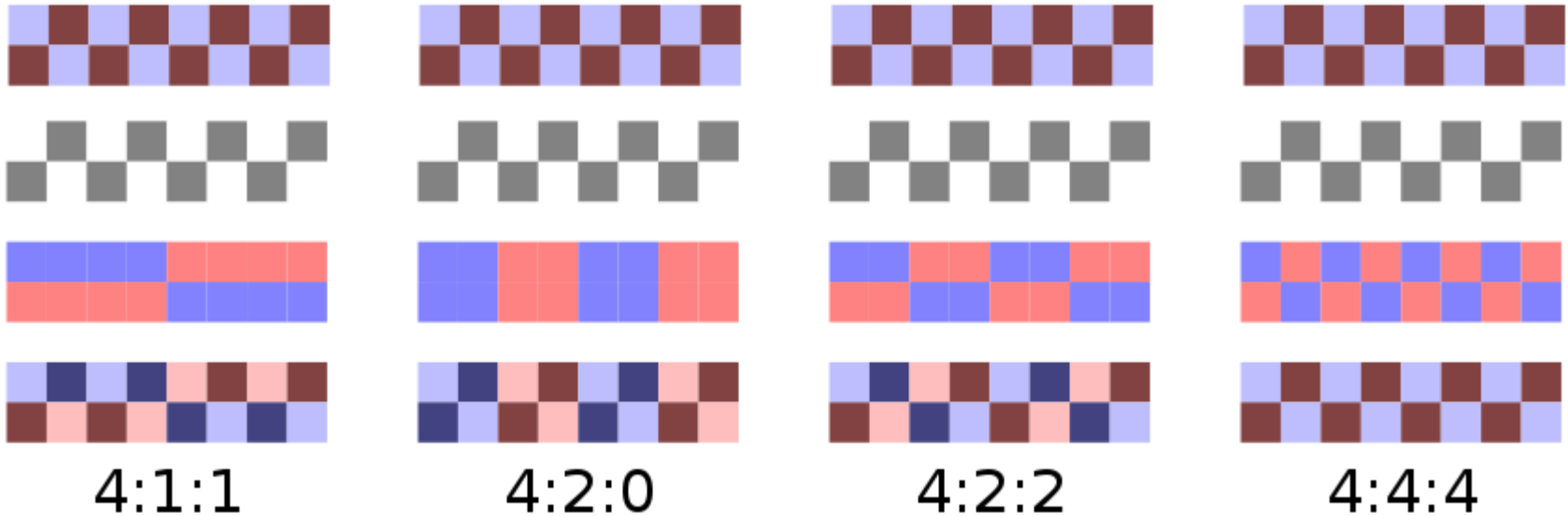
Bildquelle: Wikipedia



- **Nutzung verschiedener Auflösungen für Luminanz und Chrominanz**
 - Farbwahrnehmung ggü. Helligkeitswahrnehmung eingeschränkt
 - Hohe Auflösung für Luminanz Y
 - Niedrigere Auflösung für U/V
- **Allgemeine Angabe** des sog. **Subsampling** als **R:a:b**
 - **R**: horizontale Abtastreferenz des Luminanzkanals
 - **a**: U/V-Abtastrate aktuelle Zeile in Relation zu R
(Chrominanzpixel/Chrominanzauflösung in aktueller Zeile, horizontal)
 - **b**: U/V-Abtastrate nächste Zeile in Relation zu R
(Chrominanzpixel/Chrominanzauflösung in nächster Zeile, vertikal)
- **Beispiele:**
 - 4:4:4 – gleiche Abtastrate für alle Komponenten
 - 4:2:2 – Farbkomponenten mit 1/2 Abtastrate
 - 4:1:1 – Farbkomponenten mit 1/4 Abtastrate (eine Spalte auslassen)
 - 4:2:0 – Farbkomponenten mit 1/4 Abtastrate (eine Zeile auslassen)
 - 4:2:1 – (obsolet, entspricht nicht der obigen Systematik) Rote Komponenten mit halber, blaue mit 1/4 Abtastrate (vgl. YIQ/NTSC)



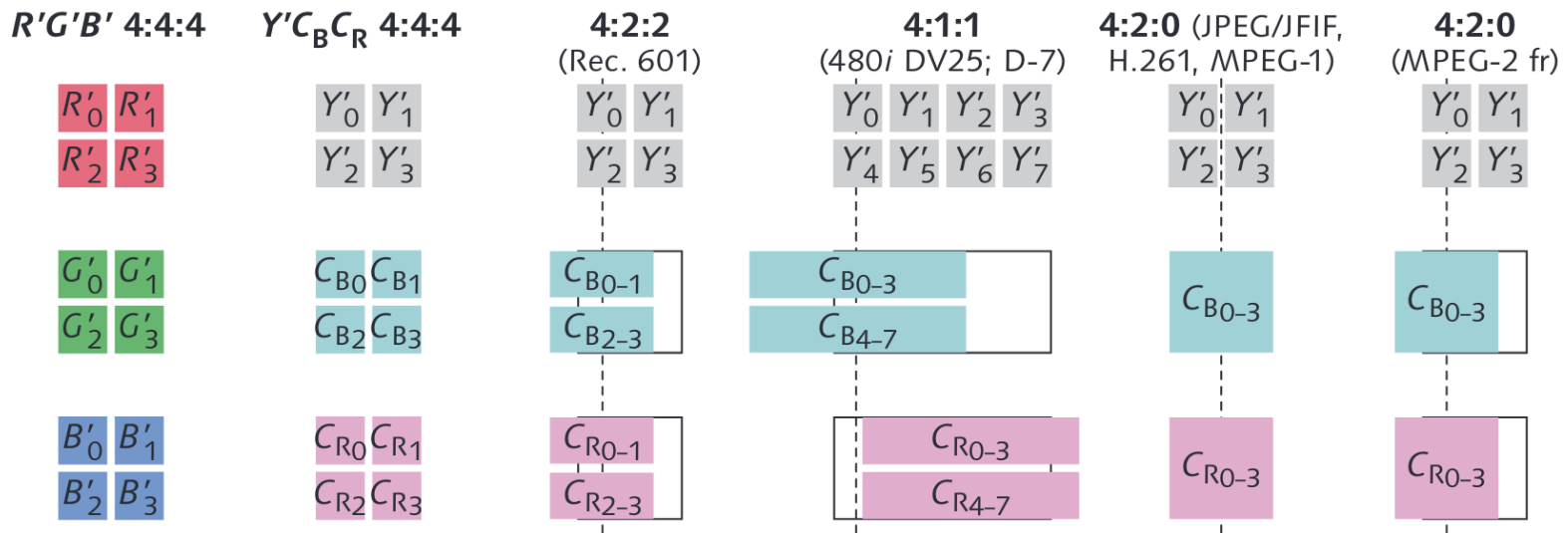
Komponentenkodierung mit unterschiedlicher Auflösung



- 4:4:4 beschreibt den Farbwert pro 1x1 Block (Pixel)
- 4:2:2 beschreibt den Farbwert pro 2x1 Block
- 4:2:0 beschreibt den Farbwert pro 2x2 Block
- 4:1:1 beschreibt den Farbwert pro 4x1 Block



Komponentenkodierung mit unterschiedlicher Auflösung



Verschiebung der Samples bezieht sich auf die im Standard festgelegten Orte des Farbsamplings.

Bildquelle: „Chroma subsampling notation“, Charles Poynton

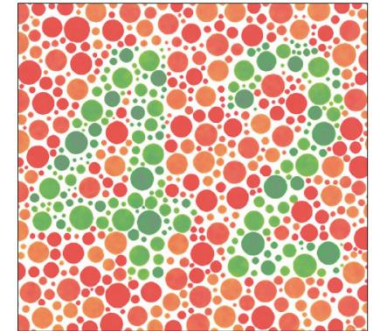


Komponentenkodierung u. Leistungsanforderungen

Signale	f_{sample} [MHz]	Samples/ Zeile	# Zeilen	Datenrate [MBit/s]	Gesamt [MBit/s]	Format
R	13.5	864	625	108	324	4.4.4
G	13.5	864	625	108		
B	13.5	864	625	108		ITU 601
Y	13.5	864	625	108	216	4.2.2
Cr	6.75	432	625	54		
Cb	6.75	432	625	54		ITU 601
Y	13.5	720	576	83	166	4.2.2
Cr	6.75	360	576	41.5		
Cb	6.75	360	576	41.5		
Y	13.5	720	576	83	124.5	4.2.0
Cr/Cb	6.75	360	576	41.5		
Y	13.5	720	576	83	124.5	4.2.0
Cr/Cb	6.75	360	576	41.5		SIF

Farbwahrnehmung

- Die Wahrnehmung ist abhängig von
 - Betrachter – Rot-Grün-Schwäche etc.
 - Situation – Position und Beleuchtung
 - Medium
 - Alter des Beamers
 - Monitoreinstellung
 - Digital oder Druck
 - Umgebung – z.B. Kontraste



Blauer Himmel durch Lichtstreuung, steiler Lichteinfall



Abendrot durch Lichtstreuung, flacher Lichteinfall

Bildquelle: J. Böhringer, P. Bühler, P. Schlaich:
Kompendium der Mediengestaltung; Springer, 2008.

Semiotik: Lehre von der Bedeutung der Zeichen

• Zeichenkategorien

- **Ikone (Icons):** Zeichen, die dem dargestellten Objekt ähneln



- **Index:** Zeichen, das direkt mit einem Objekt verknüpft ist



- **Symbol:** kein direkter Bezug zwischen Zeichen, Objekt und Bedeutung



• Zeichendimensionen

- **Syntaktik** – Wie: Formale Gestaltung eines Zeichens
- **Semantik** – Was: Bedeutung und die Botschaft eines Zeichens
- **Pragmatik** – Warum: Zweck und Einsatzgebiet eines Zeichens





A

B

C

12

13

14

Bildquelle: J. Böhringer, P. Bühler, P. Schlaich:
Kompendium der Mediengestaltung; Springer, 2008.

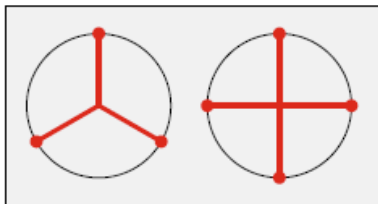
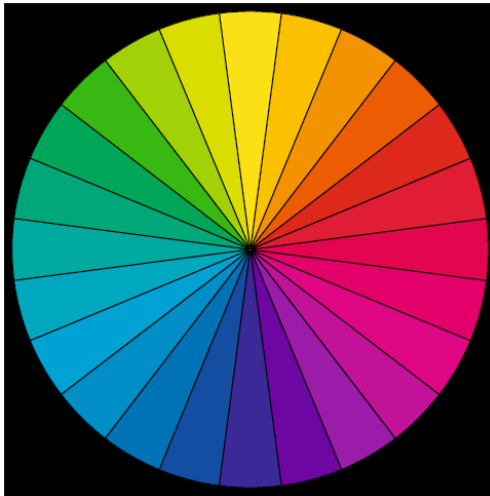


Farbgestaltung

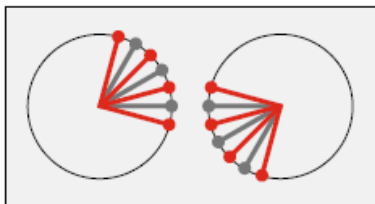
- Unterschiedliche Wahrnehmungen ausnutzen
 - Angenehme Farbkombinationen
 - Werbung
 - Wohnungseinrichtung
 - Corporate Design
 - Unangenehme Farbkombinationen
 - Warnschilder
 - Aggressive Werbung
- Was ist angenehm?



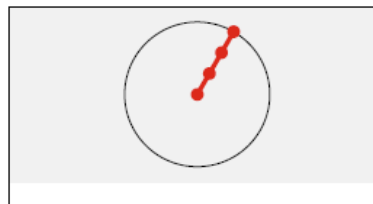
Farbgestaltung



Auswahl gleichabständiger Farben



Auswahl benachbarter Farben mit einer Farbe Abstand zur besseren Unterscheidung



Abstufung der Helligkeit des gleichen Farbtons



Farbauswahl als Dreiklang



Farbauswahl aus dem gelb-roten Farbbereich



Abstufung der Helligkeit eines Farbtons nach Weiß



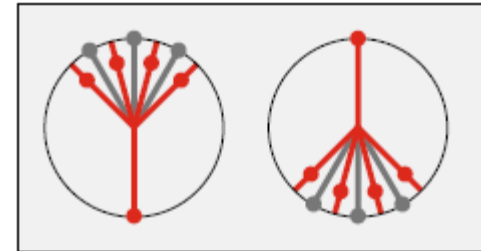
Farbauswahl als Vierklang



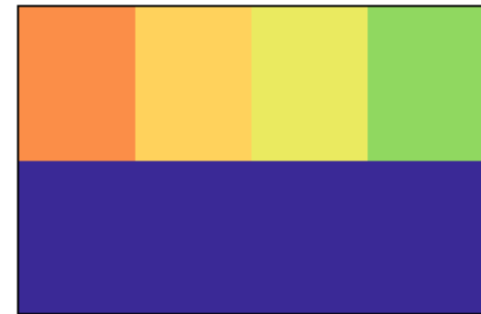
Farbauswahl aus dem blau-grünen Farbbereich



Abstufung der Helligkeit eines Farbtons nach Schwarz



Auswahl gegenüberliegender Farben als Variation des Komplementärkontrastes



Die Hauptfarbe Blau ist gesättigt, die vier Gegenfarben sind aufgehellt.



Die Hauptfarbe Gelb ist gesättigt, die einzige Gegenfarbe ist aufgehellt.

Bildquelle: J. Böhringer, P. Bühler, P. Schlaich:
Compendium der Medientgestaltung; Springer, 2008.



- Beispiel: Universität Tübingen

Die Farben der Eberhard Karls Universität Tübingen

Primärfarben



KARMINROT



GOLD



ANTHRAZIT

[zu den Primärfarben]

Die Sekundärfarben



[zu den Sekundärfarben]

Farbwerte der Hausfarben

Karminrot

Pantone 201

CMYK 35 | 100 | 70 | 10

RGB 165 | 30 | 55

Gold

Pantone 873

CMYK 30 | 30 | 60 | 10

RGB 180 | 160 | 105

Anthrazit

Pantone 431

CMYK 30 | 0 | 0 | 85

RGB 50 | 65 | 75

Die Pantone®-Farbpalette hat Leitfunktion, sie ist im Druck den anderen Systemen vorzuziehen und an ihr sind alle Darstellungen in anderen Farbsystemen auszurichten.

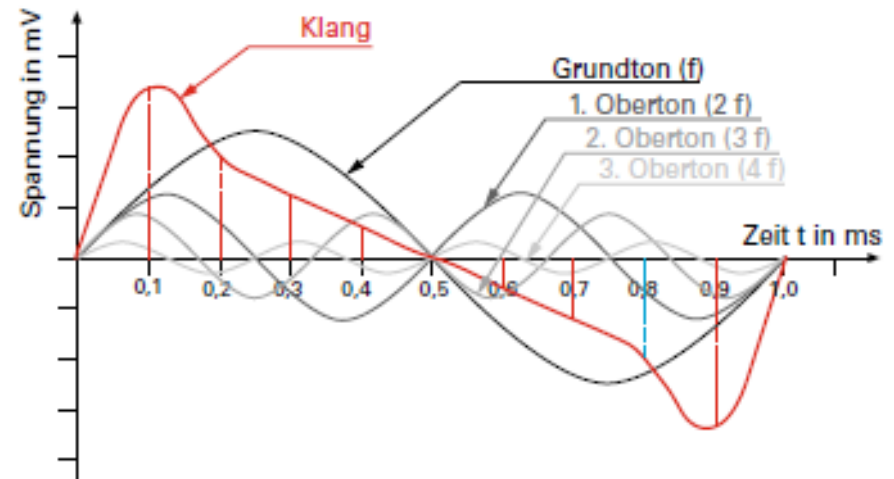
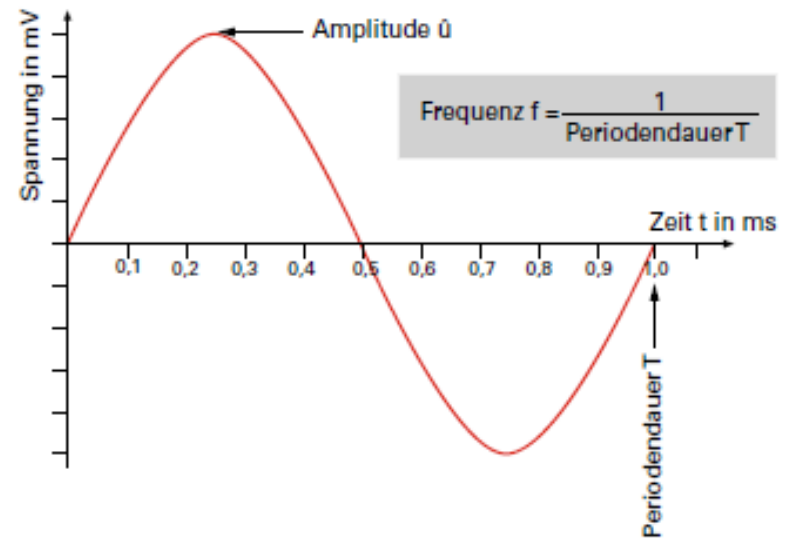


AKUSTISCHE WAHRNEHMUNG



Was ist ein Ton?

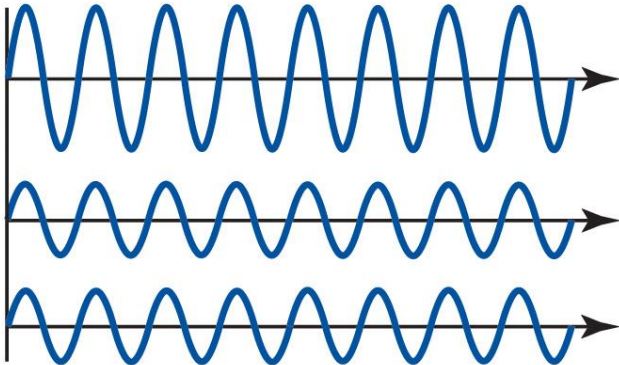
- Ein **Ton ist eine Schallwelle**
- **Amplitude** gibt **(Laut-)Stärke** an
- **Frequenz** die **Höhe** des Tons
- Signale bzw. Töne können durch Überlagerung/Interferenzen verändert werden.
 - Im Gegensatz zu Licht können die Bestandteile nicht direkt extrahiert werden



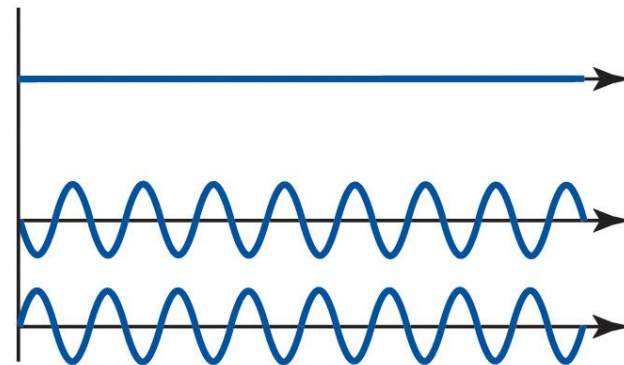


Interferenzen und Schwebungen

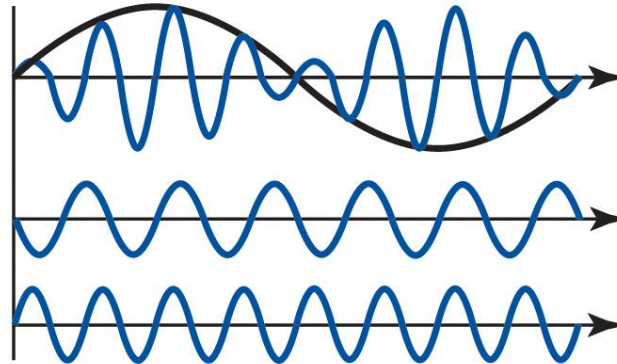
Konstruktive Interferenz



Destruktive Interferenz



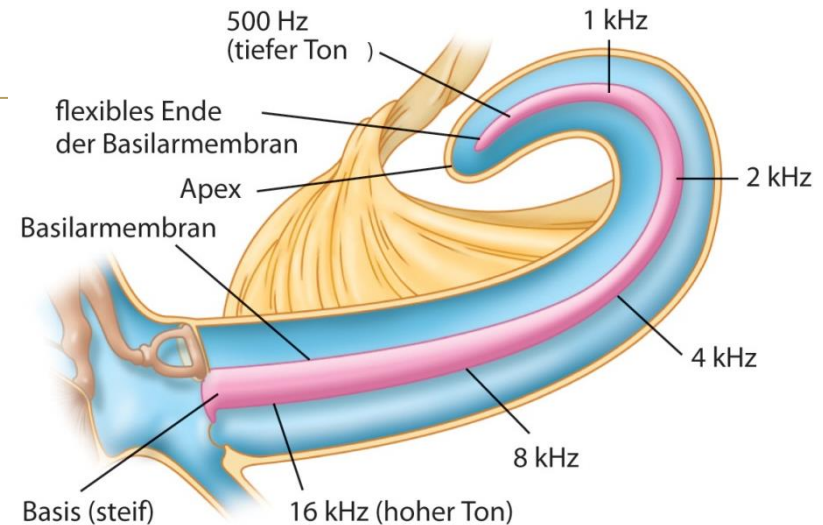
Schwebung





Akustische Wahrnehmung

- **Tiefe Töne** werden **am Ende** und **hohe Töne am Anfang** der Basilmembran **detektiert**
→ **Spektralanalyse**



- **Einschränkung:** Erregung nahe beieinander liegender Haarzellen mit stark unterschiedlicher Reizintensität

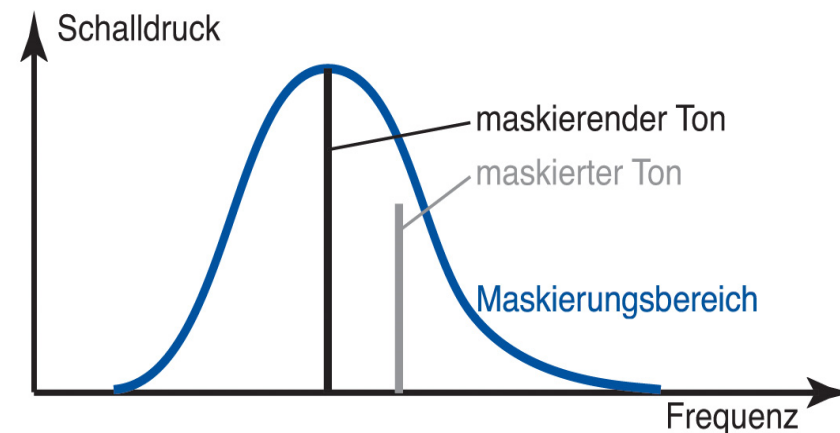
• Maskierungseffekte

- Frequenzmaskierung

- Stark unterschiedliche Reizintensität in benachbarten Frequenzen

- Zeitliche Maskierung

- Stark unterschiedliche Reizintensität in kurzen zeitlichen Abständen





Amplitude (Schalldruckpegel)

- **Logarithmische Kenngröße gemessen in Dezibel (dB)**
 - Je 10dB entsprechen wahrgenommener Lautstärkeverdopplung
 - Hörschwelle bei 2kHz: 0dB
 - normale Unterhaltung: ca. 40-60dB
 - Gehörschäden bei langfristiger Einwirkung: 85dB
 - Straßenverkehr: ca. 90dB
 - Diskothek: >100dB
 - Gehörschäden bei kurzfristiger Einwirkung: 120dB
 - Schmerzschwelle: 134dB
 - Düsenflugzeug: 140dB



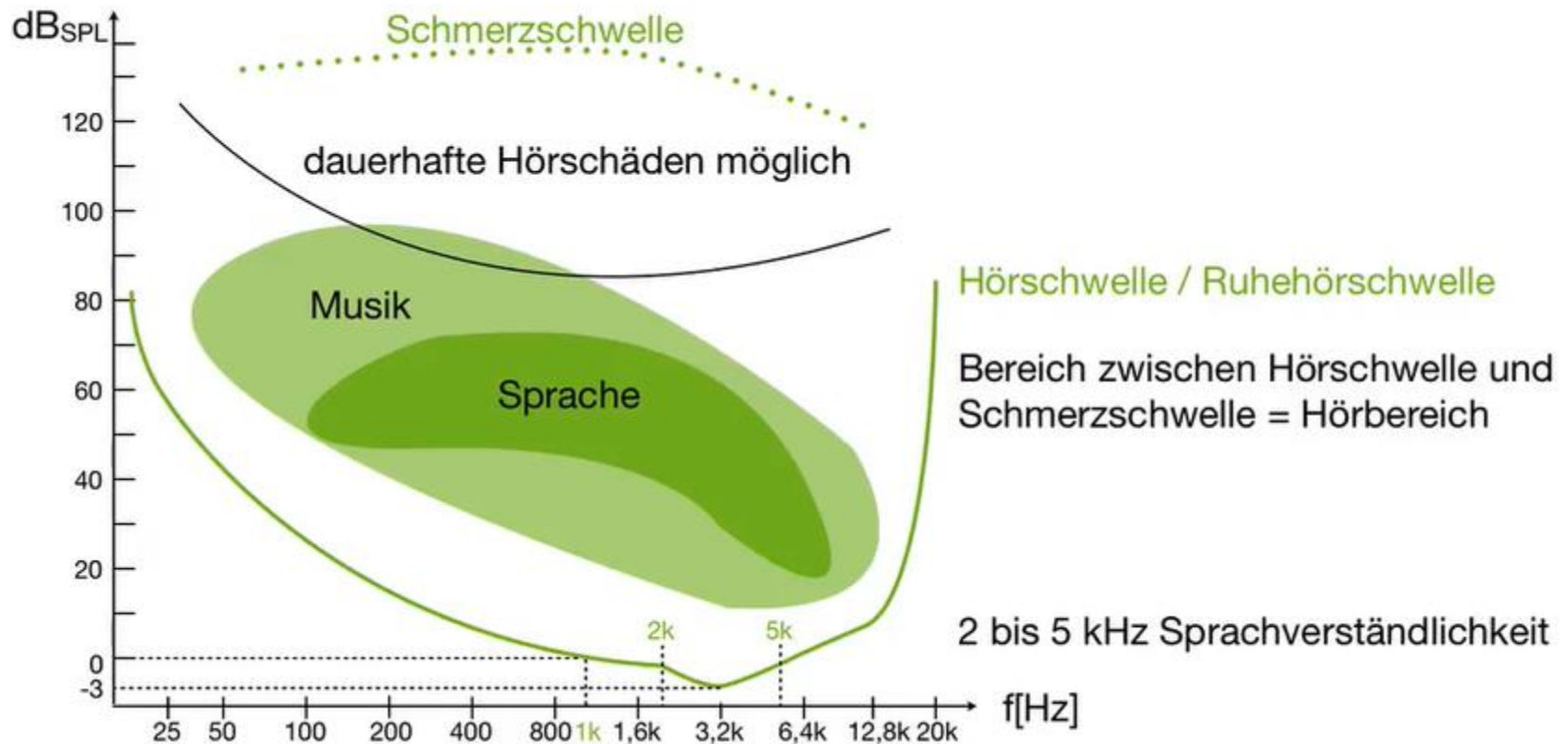
Gefühlte Lautstärke

- **Physikalisches Maß: Schalldruckpegel**
 - wird in Dezibel (db) gemessen
- **Phon** beschreibt die **Wahrnehmung eines Schallereignisses** als Hörereignis als Vergleich mit Sinuston mit 1000Hz
 - jedoch: fehlende Proportionalität
 - 10 → 20phon → 6-mal so laut
 - 50 → 60phon → doppelt so laut
 - 40 → 80phon → 16-mal so laut
- **Sone** liefert **subjektiv wahrgenommene Lautstärke linear skaliert auf den Lautheitseindruck eines Sinustons von 40 dB** unter Berücksichtigung von psychoakustischen Effekten
 - 1 sone = 40 phon → n sone = n-fach so laut

sone	1	2	4	8	16	32	64	128	256	512	1024
phon	40	50	60	70	80	90	100	110	120	130	140



Der Hörbereich

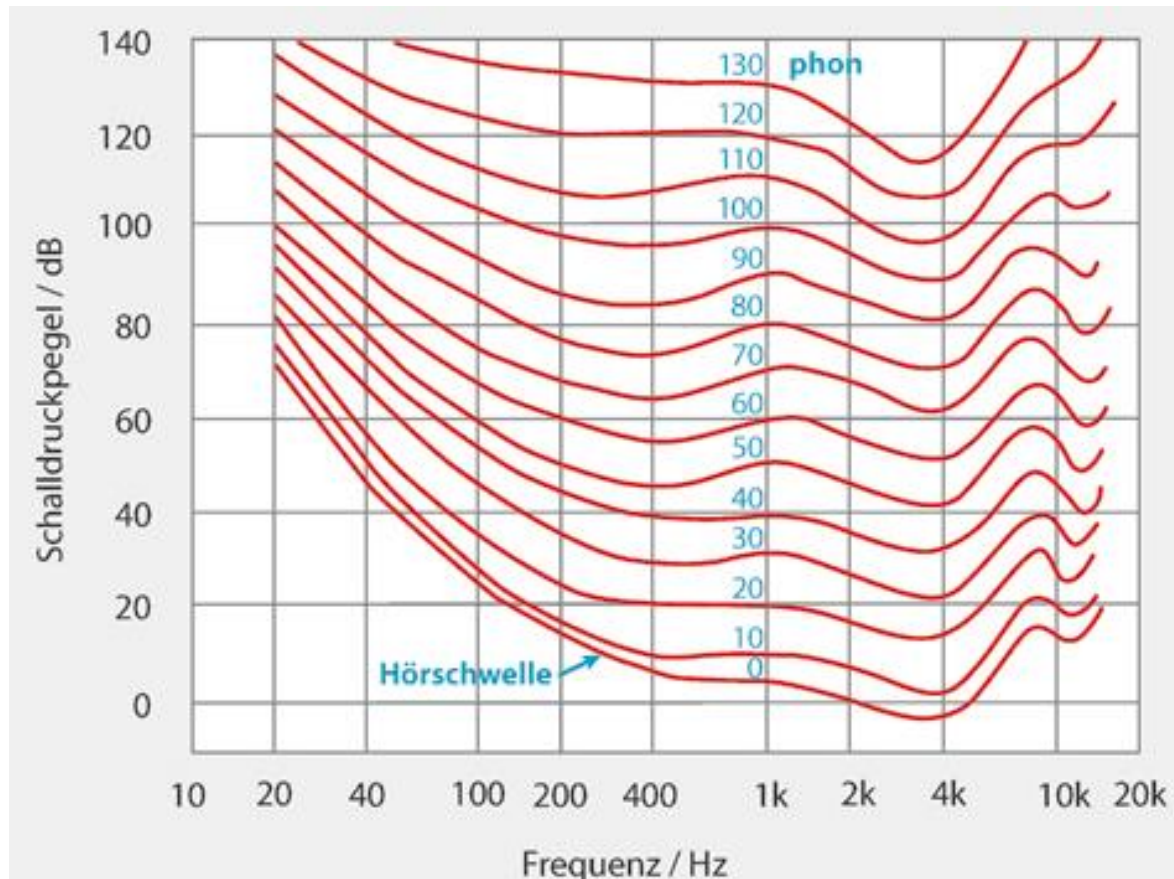


Bildquelle: <https://www.youtube.com/watch?v=glvx0u0aQ1M>



Gefühlte Lautstärke: Kennlinien [Phon]

Isophone Kurven stellen Abhängigkeit wahrgenommener Lautstärke von Frequenzen dar





Psychoakustisches Modell

- Menschliches Ohr erfasst nur Klänge von **20Hz** bis **20kHz**
 - **Erfassung nicht linear** über gesamtes Frequenzspektrum
→ vgl. Sone
 - Unschärfe im Randbereich
- **Verarbeitung** der **Phasenlage** (**Laufzeitunterschiede**) zwischen rechtem und linkem Ohr zur Ortung
- Einbeziehung der **Maskierungseffekte**
 - **Zeitliche Maskierung**:
 - Aufeinanderfolgende Töne werden bei starkem Wechsel der Lautstärke „überhört“
 - **Frequenzmaskierung**:
 - Beeinflussung der simultanen Frequenzwahrnehmung
→ dominante Frequenzanteile vermindern Wahrnehmung für schwächer ausgeprägte Frequenzanteile



Zusammenfassung

- **Menschliche Wahrnehmung** hat **großen Einfluss** auf die **technische Repräsentation** multimedialer Daten
 - Aufnahme
 - Kompression
 - Speicherung
- **Visuelle Wahrnehmung**
 - Mensch mischt Farben im Auge (Metamerie)
 - Mensch kann Licht nicht in seine Spektralfarben zerlegen
 - Helligkeit wichtiger als Farbe
- **Akustische Wahrnehmung**
 - Menschliche Ohr kann Klänge in Spektrum zerlegen
 - Beschränkter Frequenzbereich: 20Hz bis 20kHz
 - Keine lineare Frequenzerfassung
 - Maskierungseffekte: Frequenzmaskierung u. zeitliche Maskierung