

Grundlagen der Bildverarbeitung

Sommersemester 2022

Bilder

Visuelle Wahrnehmung:
Physik, Physiologie, Psychologie
Analoge und digitale Bilder

Wolfgang Heiden © 2014-22, mit Beiträgen von R. Herpers, F. Mannuß, B. Kahl

Wolfgang Heiden © 2014-22 wolfgang.heiden@h-brs.de

-- auf der Grundlage einer Lehrveranstaltung von Prof. Dr. Rainer Herpers sowie Folien
von Florian Mannuß 2011, Dr. Björn Kahl 2012 --

Fachbereich Informatik (Dpt. Computer Science)

Hochschule Bonn-Rhein-Sieg – University of Applied Sciences,

53754 Sankt Augustin

Germany

Was ist ein Bild?

- **Gesichtssinn = primärer menschlicher Sinn**
- **Aufnahme und Repräsentation von Information**
- **Komprimierte Information:**
 - „Ein Bild sagt mehr als tausend Worte.“

M.C. Escher



Beschreiben Sie
aussagekräftig folgendes Bild!

Beschreiben Sie das Bild so, dass es jemand, der es nicht sieht und nicht kennt, anhand Ihrer verbalen Darstellung möglichst originalgetreu zeichnen könnte.



M.C. Escher

Was ist ein Bild?

- Ein Bild ist eine **Menge von Informationen**, deren Inhalt **innerhalb eines 2-dimensionalen Datenraums** in mindestens einer **ortsabhängigen Größe** hinsichtlich deren **Intensität** variiert.

Wie entsteht ein Bild?

- Ein Bild entsteht durch das Bestrahlen einer Fläche mit Licht unterschiedlicher Intensität und Wellenlänge.
- Dieser Vorgang kann auch digital simuliert werden. Ein Bild ist dann allgemein eine **Sammlung von positionsbezogenen skalaren Werten in einem 2D-Datenraum**. Diese Werte können algorithmisch berechnet (**Computergrafik**) oder durch digitale Abtastung analoger Bildsignale (**A/D-Wandlung**) gemessen oder durch **bildgebende Apparate** (CT, MRT, etc.) aus physikalischen Messwerten errechnet werden.
- In der menschlichen **Wahrnehmung** entsteht ein Bild im Bewusstsein, indem Licht durch das optische System **Auge** in den Sinneszellen der Netzhaut elektrische Signale auslöst, die durch Transport- und Verrechnungsvorgänge in den **Nervenbahnen** zum und im **Gehirn** verändert, bewertet und interpretiert werden.

Komponenten der Bildwahrnehmung

• Auge

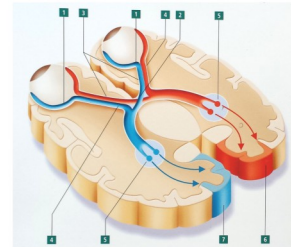
- Bündelung, Brechung und Filterung von Licht
- Umwandlung von Licht in Nervenimpulse
- Helligkeit, Wellenlängen, Ortsauflösung

• Nervenbahnen

- Bündelung, Separation und Weiterleitung der visuellen Nervenimpulse
- Kontraste, Muster (rezeptive Felder)

• Sehrinde (Gehirn)

- Interpretation der visuellen Signale
- Kanten, Farben, Schattierungen, Räumlichkeit, Objekterkennung, Emotionen, Aufmerksamkeit



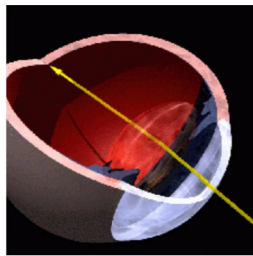
Bildquelle: Der menschliche Körper
(Atlas Verlag)

modif. nach: Lehmann et al., 1997

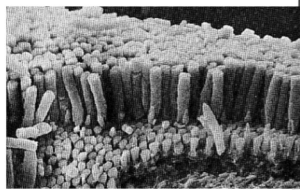
Das visuelle System des Menschen beinhaltet mehrere Komponenten:

1. Das Auge dient der Aufnahme visueller Information, indem elektromagnetische Signale innerhalb eines bestimmten Wellenlängenbereichs (d.h. „Licht“) physikalisch-optisch vorverarbeitet wird und schließlich in den Rezeptorzellen der Netzhaut durch Absorption elektrische Nervensignale auslösen. Physikalische Eigenschaften der Lichtsignale wie Intensität, Wellenlängen und örtliche Variation werden in den elektrischen Signalen codiert.
2. Die von den einzelnen Sehzellen ausgehenden Nervenimpulse werden über die Reizleitungsstrecke vom Auge zum Gehirn mehrfach gebündelt, wieder aufgetrennt und über Nachbarschaftsbeziehungen in Form von sog. rezeptiven Feldern miteinander in Beziehung gesetzt. Da die Sehnerven einander überkreuzen, werden (nach Informationsaustausch) Signale aus dem linken Auge in der rechten Hirnhälfte verarbeitet und umgekehrt.
3. In der Sehrinde im rückwärtigen Teil des Gehirns werden die visuellen Signale schließlich weiter verarbeitet, interpretiert und mit Erinnerungen bewertet an das Bewusstsein übermittelt. Dabei entstehen aus den physikalischen Größen Reizintensität, Wellenlänge und örtliche Variation (entsprechend) die Wahrnehmungsempfindungen Helligkeit, Farbe und Größe.

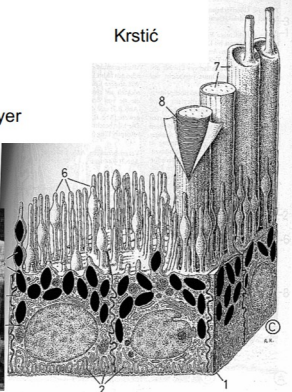
Das menschliche Auge



Mostafawy



Stryer



Krstić

- **Licht** trifft nach Fokussierung Sinneszellen in der Netzhaut (**Retina**).
- **Stäbchen**: hell/dunkel
- **Zapfen**: Farbsehen (3 Absorptionsmaxima)
- **elektrischer Reiz** in Nervenfasern.
- **Verschaltung** verschiedener **Nervenfasern** kombiniert Reize mehrerer Sehzellen.

Licht trifft die **Netzhaut** im Hintergrund des menschlichen Auges (Abb. Mostafawy). Dort befinden sich 2 Typen von Sehzellen: Zapfen und Stäbchen. (REM-Aufnahme, aus Stryer: Biochemistry; Handzeichnungen aus Krstić: Die Gewebe des Menschen und der Sägetiere)

Licht bestimmter Wellenlänge löst durch Energie-Übertragung eine **photochemische Reaktion** aus. Die chemische Veränderung des Sehpigments verursacht einen **elektrischen Reiz** in der ableitenden Nervenfasern. In den Zapfen kommen 3 verschiedene Sehpigmente mit jeweils eigenem Absorptionsmaximum vor, d.h. mit unterschiedlicher Empfindlichkeit für Licht verschiedener Wellenlänge. Die Maxima der (überlappenden) Absorptionsspektren liegen jeweils im roten, grünen bzw. blauen Bereich.

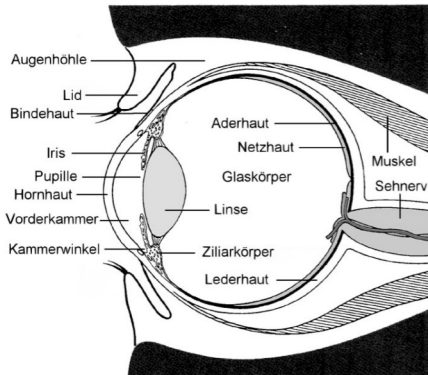
Die **Verschaltung** verschiedener **Nervenfasern** kombiniert die Reize mehrerer Sehzellen zu einem Bild, das sich aus Hell-/Dunkel-Werten, Farbqualität und Farbsättigung zusammensetzt.

Einzelninformationen werden im Gehirn zusammengesetzt, überarbeitet und als **Bild** ins Bewußtsein weitergeleitet.

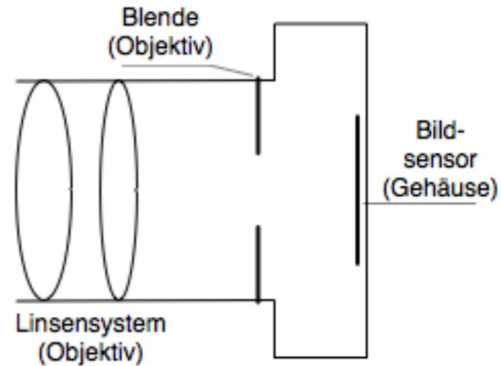
Der Bildvergleich beider Augen (binokulare Stereoskopie) liefert schließlich (zusammen mit anderen Parametern) eine **Tiefe**information.

Auge und Kamera als optische Systeme

Auge



Kamera



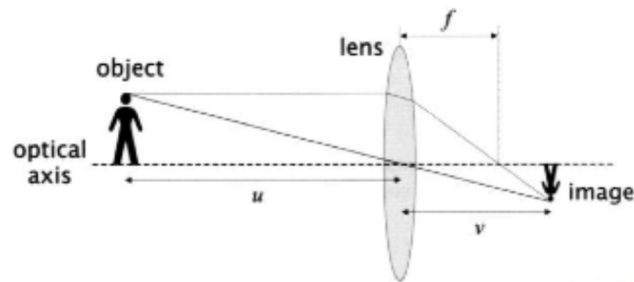
- Schärferegulierung
- Signalerzeugung
- Signalleitung

Aufbau bilderzeugender Systeme;

vgl. Auge – Kamera:

- Linsenverformung (A) vs. Linsenverschiebung (K)
- unterschiedliche Dichte unterschiedlicher Sehzellen (A) vs. gleichmäßige Körnung bzw. RGB-Sensoren
- invertierte Netzhaut → Blinder Fleck (A) vs. Fotoplatte/Film/CCD-Chip (K)

Aufnahmeoptik mit Linsensystem



• Abbildungsmaßstab

$$m = \frac{v}{u} \quad m = \frac{\text{image size}}{\text{object size}}$$

mit ...
und ...

• Aufnahmeoptik

$$\frac{1}{f} = \frac{1}{u} + \frac{1}{v}$$

mit ... Brennweite f [mm]

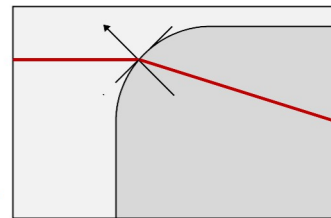
Linseneffekte im Auge ermöglichen durch die Bündelung von Lichtstrahlen ein erweitertes Sichtfeld.

Brechungsindex

n

- **Materialeigenschaft**
→ sammelnde od. streuende Eigenschaft einer Linse
- **Refraktion** basiert auf der Welleneigenschaft des Lichts.
- unterschiedliche Ausbreitungs-Geschwindigkeit in Materie
- Brechungsindex beschreibt relative Lichtgeschwindigkeit mit ...
 - n = Brechungsindex,
 - c = Lichtgeschwindigkeit im Vakuum,
 - v = Lichtgeschwindigkeit im Material
- Licht wird zur Oberflächennormale hin gebrochen, wenn es vom kleineren zum größeren Brechungsindex wechselt (d.h. vom optisch dünneren zum dichteren Medium).

$$n = c / v$$



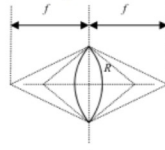
Im Bild: $n_{\text{links}} < n_{\text{rechts}}$

Der Brechungsindex ist proportional zur optischen Dichte.

Lintentypen

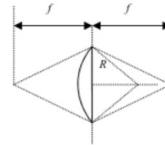
- **bikonvex**

- **sammelt** paralleles Licht im Brennpunkt



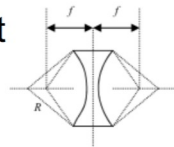
- **plankonvex**

- **sammelt** paralleles Licht im Brennpunkt



- **bikonkav**

- **streut** paralleles Licht so, als ob es vom Brennpunkt käme



vgl. Brillenstärken:

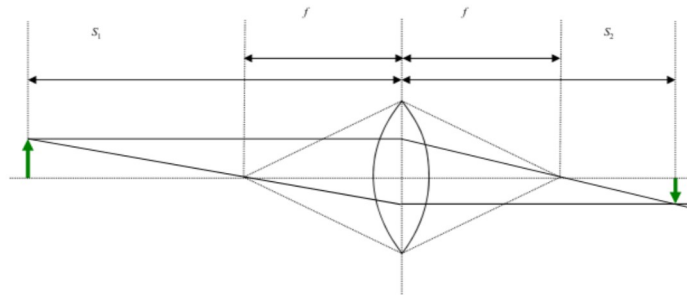
- kurzsichtig (Hilfe für Fernsicht) → konkav (Streuung, Bild verkleinert)
- weitsichtig (Hilfe für Nahsicht) → konvex (Sammlung, Bild vergrößert, vgl. Lupe)

Bildkonstruktion an Linsen

- Objekt bei S_1 außerhalb der Brennweite
- Bild bei S_2 auf anderer Seite der Linse

$$\frac{1}{S_1} + \frac{1}{S_2} = \frac{1}{f}$$

- Vergrößerung: $M = -\frac{S_2}{S_1} = \frac{f}{f - S_1}$

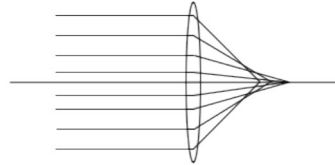


Nach den Gesetzen der Optik lassen sich Abbildungen an Linsensystemen genau berechnen.

● Sphärische Aberration

- Randstrahlen haben (bei kugelförmigen Linsen) eine geringere Brennweite als Strahlen nahe der optischen Achse.

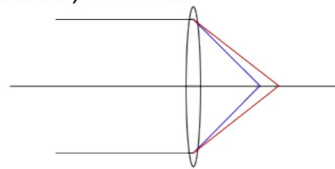
→ *Unschärfe*



● Chromatische Aberration

- Unterschiedliche Wellenlängen (Farben) werden unterschiedlich stark gebrochen.

→ *Farbige Kanten*

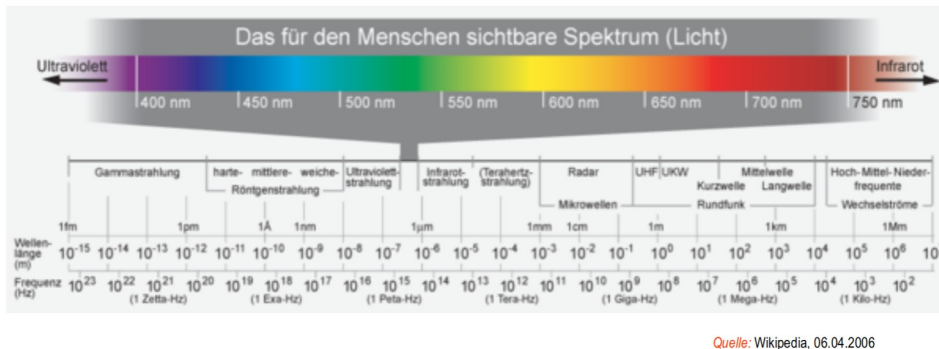


Linsenfehler können zu fehlerhaften Abbildungen führen. Solche Fehler nachträglich zu korrigieren, zählt zu den Aufgaben der digitalen Bildverarbeitung.

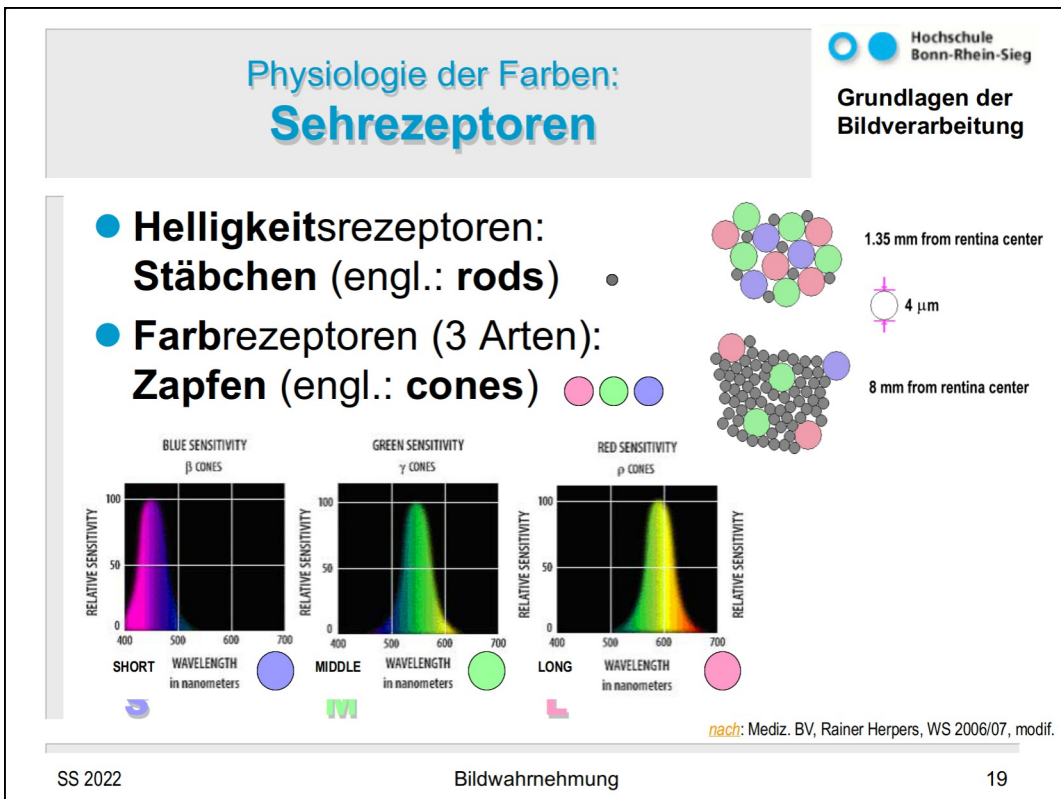
Bei Menschen mit fehlerhaften optischen Systemen im Auge kann eine Korrektur durch Brillen erfolgen;

bei digitalen Bildern dienen dazu entsprechende Algorithmen.

Physik des Lichts: Elektromagnetisches Spektrum



Das sichtbare Licht stellt nur einen kleinen Ausschnitt aus dem Spektrum elektromagnetischer Wellenlängen dar. Dies liegt darin begründet, dass nur in dem Wellenlängenbereich zwischen ca. 380 und 750 nm Sehpigmente in der menschlichen Netzhaut Energie absorbieren und damit ein elektrisches Signal auslösen können. Energieärmeres Licht größerer Wellenlängen reicht von Infrarot bis zu Wechselströmen, energiereicheres kurzwelligeres Licht von Ultraviolett bis zur Gammastrahlung.



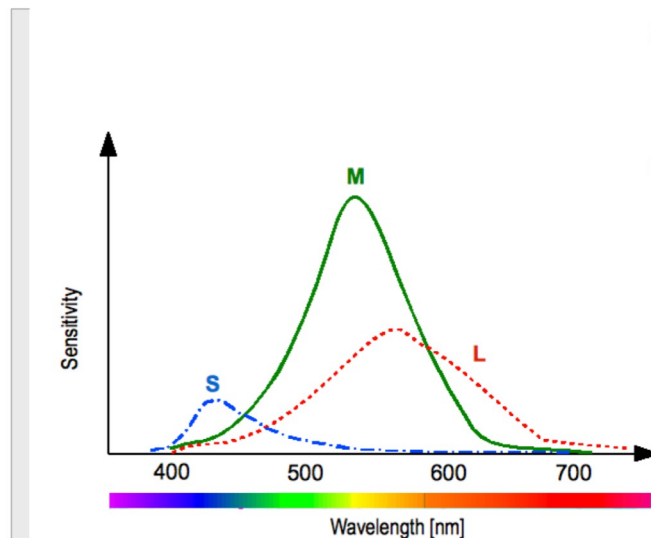
Im menschlichen Auge kommen zwei grundsätzlich unterschiedliche Arten von Rezeptorzellen vor, die sich in ihren Fähigkeiten, aber auch hinsichtlich ihrer Anzahl und Verteilung in verschiedenen Bereichen der Netzhaut unterscheiden:

ca. 10^8 Stäbchen (Intensität: hell/dunkel)

ca. 10^6 Zapfen (drei Typen mit spektral unterschiedlichen Absorptionsmaxima: Farben bzw. Wellenlängen)

Im Zentrum der Netzhaut (=Retina), der sog. Fovea (od.: gelber Fleck) stehen die Zapfen besonders dicht; hier ist die Sehschärfe (Ortsauflösung) am höchsten und die Farbwahrnehmung am intensivsten. In den äußeren Bereichen (Peripherie) dominieren die Stäbchen; hier reagiert das Auge empfindlicher auf schwache Lichtintensitäten und schneller auf zeitliche Veränderungen.

Physiologie der Farben: Farbwahrnehmung



• Stäbchen

- Helligkeits-Wahrnehmung
- ca. 100 Mio.

• 3 Zapfentypen

- rot (L) / grün (M) / blau (S)
- überlappende Empfindlichkeits-Kurven
- ca. 6,5 Mio.



L = long – M = medium – S = short (wavelength)

There is, however, another human sense of perception that has to deal with a similar difficulty – and it is performing quite well.

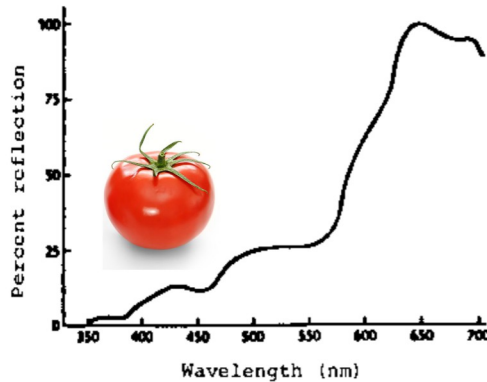
This is referring to color perception with three different types of retinal cone cells that show maximum electric response when stimulated with light of different wavelength, although their sensitivity curves overlap in a broad range of the visual spectrum. According to their maximum sensitivity, the cone types are named red, green, and blue sensor, respectively. (This is, by the way, the reason why the entire color range perceivable to humans can be composed by red, green, and blue light dots – which is how almost every electronic color display works.)

Despite the strong cross-sensitivity of the sensors human color perception is very accurate, as the combination of different sensor responses to an identical stimulus can be resolved to a unanimous color.

Psychologie der Farben: Farbtyp, Sättigung, Helligkeit

Spektralbereiche (R,G,B bzw. C, M, Y)

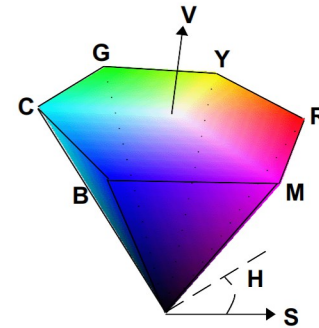
- Red, Green, Blue ; Cyan, Magenta, Yellow
- Wellenlänge, **Sinnesreiz**



nach: Mediz. BV, Rainer Herpers, WS 2006/07

Empfindungen (H, S, V)

- Hue (Farbtyp)
- Saturation (Farbsättigung)
- Value (Helligkeit)
- Wahrnehmung**



Wenngleich verschiedene Wellenlängen elektromagnetischer Signale die menschliche Farbwahrnehmung auslösen, ist zwischen Wellenlänge (als physikalische Größe) und Farbe (als Empfindung) zu unterscheiden. Dies wird besonders deutlich bei der als tiefrot wahrgenommenen Färbung einer reifen Tomate, die dennoch aus einer Kombination von Licht unterschiedlicher Wellenlängen, über praktisch das gesamte sichtbare Spektrum verteilt, entsteht; relevant für die Farbwahrnehmung ist aber die "dominante Wellenlänge", d. h. diejenige mit der höchsten Intensität – und die befindet sich im Rot-Bereich.

Spektralausschnitt der Wellenlängen (linear, 400-800 nm) → Farbkreis (Purpurbereich!)

Bildquelle Tomate:

http://commons.wikimedia.org/wiki/File:Bright_red_tomato_and_cross_section02.jpg?uselang=de, 08.01.2013, Creative Commons

Farbreiz vs. Farbempfindung

● Stimulus

bestimmt durch ...

- Dominante Wellenlänge, Wellenlängen-Reinheit, Helligkeit
- achromatisch vs. chromatisch

● Farbempfindung

bestimmt durch ...

- Farbtyp (engl.: *Hue*)
- Farbigkeit, Sättigung (engl.: *Saturation*: relative Farbigkeit im Vergleich zur Helligkeit) ()
- Helligkeit, Leuchtstärke (engl. *Lightness*: relative Helligkeit im Vergleich zu Weiß → Weißabgleich)

Farbe ist keine physikalische Eigenschaft, sondern eine subjektive biologische Empfindung!

Details zur Repräsentation von Farben in digitalen Bildern folgen später.

Intensitätswahrnehmung

- Wahrnehmung von **Reiz-Intensitäts-Unterschieden** (Reizunterscheidung)
- **logarithmische Beziehung** zwischen **Reizintensität** und **Sinnesempfindung**

- **Weber-Fechner'sches Gesetz**

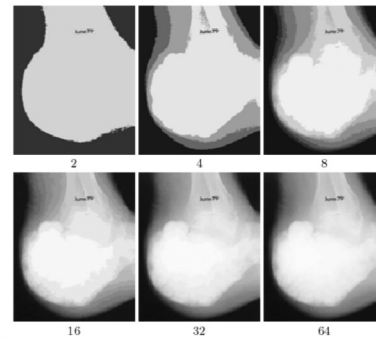
- $j = c \log I_j + C$

- j = empfundene Intensität

- c = Konstanter Faktor

- I_j = physikalische Intensität

- C = konstanter Offset



→ Konsequenzen für
Helligkeitsabstufungen in Grauwertbildern

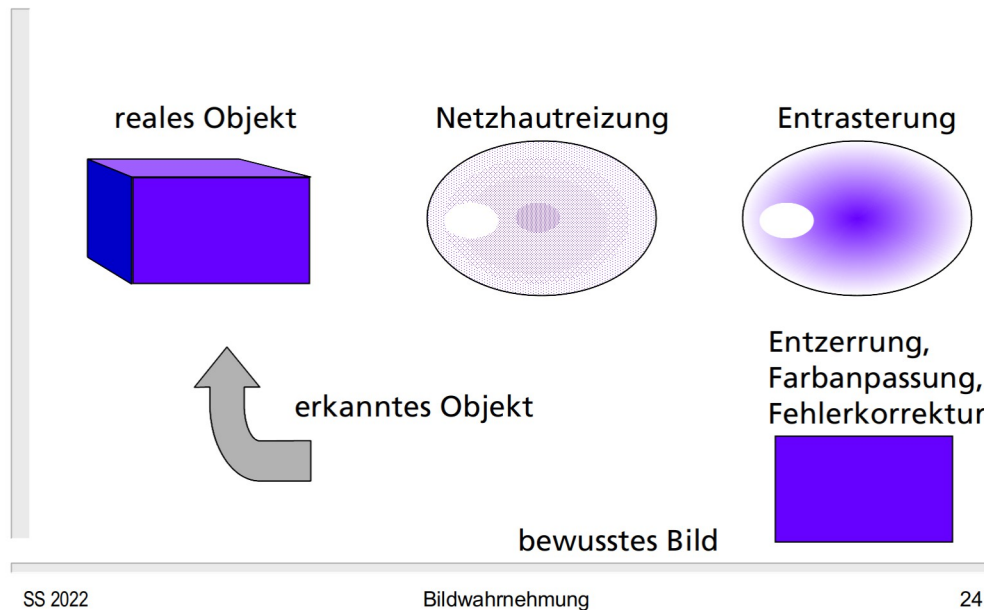
Quelle: Lehmann et al., 1997

Konsequenzen: z.B. Grauwertspreizung, weil bei insgesamt hohen Helligkeitswerten Unterschiede erst bei sehr großen Betragsdifferenzen noch wahrgenommen werden.

→ Abbildung einer logarithmischen Signalskala auf eine lineare Wahrnehmungsskala

HyMN-Version: Zoom auf Bild implementieren!

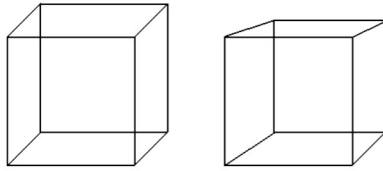
Neuronale Bildverarbeitung (unbewusste Vorfilterung)



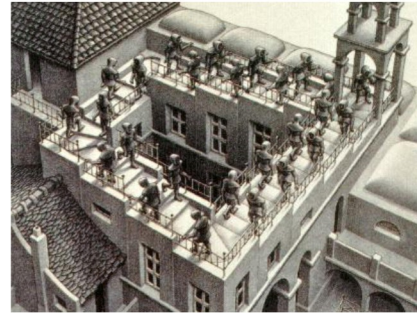
Wie stark die Wahrnehmung eines visuellen Eindrucks von neuronaler Vorverarbeitung beeinflusst wird, soll diese schematische Abbildung illustrieren.

Ein rechteckiges Objekt (blauer Quader) wird als solches identifiziert, obwohl die Reizung der unterschiedlich dicht verteilten Sehzellen auf der gerundeten, gewölbten Netzhaut als primäres „Bild“ nicht viel damit gemein haben.

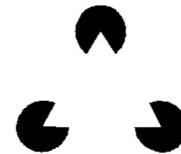
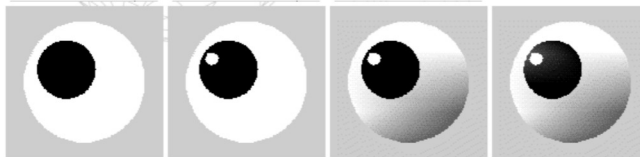
Optische Täuschungen



Hohes Risiko von intuitiven Fehlinterpretationen außerhalb des natürlichen Erfahrungskontexts!



Quelle: M.C. Escher



Die neuronale Vorverarbeitung von visuellen Signalen trägt in der Regel dazu bei, dass wir Objekte als solche erkennen, ohne dass es dazu einer kognitiv intellektuellen Anstrengung bedürfte. Diese Vorbewertung kann uns aber auch einen Streich spielen. Diverse optische Täuschungen können das verdeutlichen.

1. Oben links sind 2 Würfel zu sehen, die sich hinsichtlich der gewählten **Projektionsmethode** unterscheiden. Während bei der **orthogonalen** Projektion (links) Kantenlängen und rechte Winkel in der Projektionsebene erhalten bleiben, erscheint die Rückwand des rechten Würfels in der **perspektivischen** Projektion mit exzentrischem Fluchtpunkt kleiner. Dass das menschliche visuelle System unterbewusst eine perspektivische Projektion erwartet, kann dazu führen, dass der linke Würfel so wirkt, als würde er sich nach hinten verbreitern.
2. Der Grafiker M. C. Escher spielte gern mit unmöglichen räumlichen Objekten – so z.B. eine in sich geschlossene Treppe, die – je nach Betrachtungsrichtung – endlos aufwärts oder endlos abwärts zu führen scheint.
3. Wenige einfache Kreise erscheinen in geeigneter Anordnung als dreidimensionales Objekt. Einfache, lineare Helligkeitsverläufe tun ein Übriges, dass man sich des Eindrucks kaum erwehren kann, ein Auge zu sehen. Wenn Erwartungen des Unterbewusstseins für ein bekanntes Objekt erfüllt werden, kommt der Sinneseindruck bereits kategorisiert im Bewusstsein an.
4. Was zeigt das Bild rechts unten? – drei schwarze Kreisausschnitte vor weißem Hintergrund oder ein weißes Dreieck vor drei schwarzen Kreisen?

- **Analoge Bilder**

- Mathematisch: kontinuierliche 2D-Abbildung einer Ortsfunktion
- beliebig hohe Genauigkeit
- Digital „analog“: Vektorgrafiken

- **Digitale Bilder**

- Mathematisch: diskrete Abbildung einer Ortsfunktion
- Gerastert: finite Elemente: **Pixel** = *Picture Elements*
- Physiologisch „digital“:
 - | begrenzte Dichte der Sehzellen in der Retina → Ortsauflösung
 - | Aktionspotenziale leitender Nervenzellen → Intensitätsschwellen, Weber-Fechner
 - | Anzahl unterscheidbarer Farben → Farbauflösung

- **Bildwahrnehmung wird analog empfunden.**

Die eigentliche Unterscheidung zwischen analogen und digitalen Bildern liegt zwischen mathematisch **kontinuierlich** und **diskret**, nicht zwischen Mensch und Computer.

Wenn wir von „**Digitalen** Bildern“ sprechen, meinen wir aber i.d.R. **gerasterte Bilder mit begrenzter örtlicher Auflösung und quantisierten Signalintensitäten mit begrenzter Bittiefe**.

- Abbildung einer zweidimensionalen, kontinuierlichen Intensitätsfunktion $I(x,y)$ auf diskrete Werte
 - Räumliche Diskretisierung:
kontinuierliche Ortskoordinaten (x,y)
→ diskrete Ortskoordinaten (u,v)
 - Quantisierung der Intensitätswerte mit begrenzter Bittiefe:
z.B. 8 bit pro Pixel pro Farbkanal ($0 \rightarrow 255$)
- diskrete Funktion

$$I(u, v) \in \mathbb{P}, \quad \text{mit } u, v \in \mathbb{N}$$

Wenn wir von „Digitalen Bildern sprechen“, meinen wir i.d.R. gerasterte Bilder mit begrenzter örtlicher Auflösung und quantisierten Signalintensitäten mit begrenzter Bittiefe.

Gebräuchliche Dateiformate für (gerasterte) digitale Bilder

• GIF

- indizierte Farben, i.d.R. 8 bpp
- Transparenz binär

• TIFF

- RGBA, 32 bit (od. mehr)
- Layers
- keine bis geringe (verlustfreie) Kompression
- höchste Qualität

• PNG

- RGBA, 32-48 bit
- verlustfrei komprimiert

• JPEG

- $Y C_r C_b$ (\rightarrow RGB, 24 bit)
- keine Transparenz
- verlustbehaftet (!)
- variabel komprimiert
- +/- deutliche Kompressionsartefakte

• JPEG 2000

- vgl. JPEG
- bessere Kompression

• BMP

- RGB, 24 bit
- verlustfrei, kaum komprimiert
- Keine Transparenz
- (Windows-Relikt)

GIF-Bilder entsprechen seit einiger Zeit nicht mehr aktuellen Qualitätsstandards. Trotzdem finden sie noch verschiedentlich Anwendung als einfache Icons mit transparentem Hintergrund oder kleine Animationen auf Webseiten.

TIFF gilt nach wie vor als Standard für qualitativ hochwertig gespeicherte Bilddaten, insbesondere wenn diese zur Weiterverarbeitung bestimmt sind.

PNG ist das aktuell am weitesten verbreitete Bildformat mit verlustfreier Kompression und abgestufter Transparenz.

JPEG setzt als verlustbehaftet komprimiertes Bildformat – trotz der deutlich höheren Qualität bzw. Kompressionsrate von JPEG 2000 – nach wie vor den Standard schlechthin für die meisten Anwendungen mit digitalen Standbildern.

BMP ist ein veraltetes Bildformat ohne verbliebene praktische Vorteile und wird hier nur aus historischen Gründen erwähnt (und weil es in zahlreichen älteren Datensammlungen immer noch anzutreffen ist).

Kompetenzcheck

- **Generische Definition eines „Bildes“**
- **Menschliche visuelle Wahrnehmung**
 - Sinnesorgan Auge
 - Physik, Physiologie und Psychologie der visuellen Wahrnehmung
 - Spektralausschnitt, Strahlengänge an Linsen
 - Sehzellen (Typen, Spektralkurven)
 - Beziehung zwischen Signalreiz und Empfindung
 - Wahrnehmungsgrenzen
- **Analoge und digitale Bilder**
- **Gebräuchliche digitale Bildformate**
 - Name/Endung, Charakteristika, Eignung



Die hier noch einmal zusammengefassten Kompetenzbereiche sollten nach erfolgreicher Beschäftigung mit der vorliegenden Lehreinheit abgedeckt werden.