



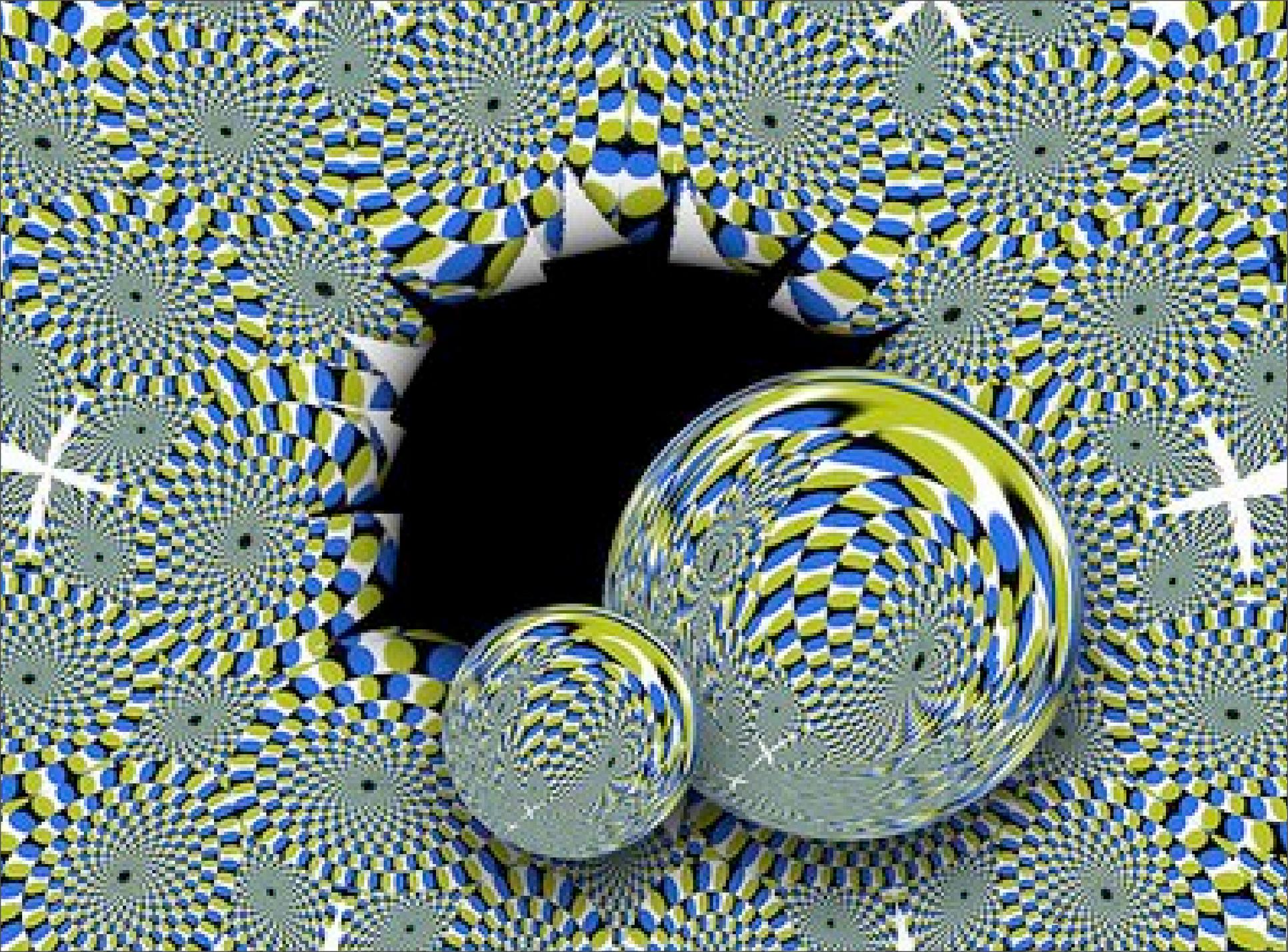
KR II: Visuelle Verarbeitung I

(Abbildung, Farben, Kamera)

Prof. Dr. Verena V. Hafner,
Humboldt-Universität zu Berlin, Institut für Informatik,
WS 2008/2009

Ein Großteil dieser Folien wurde von Prof.
Dr. Hans-Dieter Burkhard und seinen
Mitarbeitern und Studenten übernommen.

Vielen Dank!



Bilder

Was ist das?



V.V. Hafner, HU Berlin Winter-Semester 2008/09

Bilder

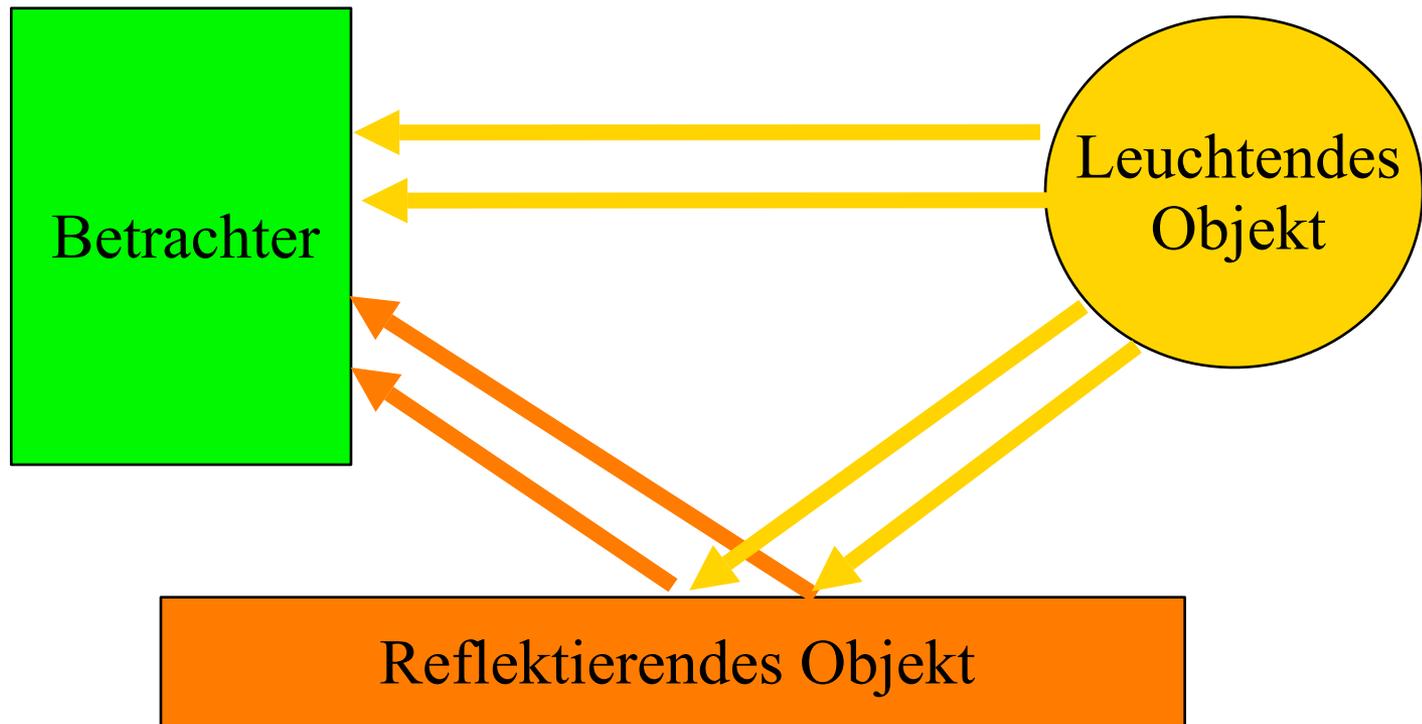
Was ist das?



V.V. Hafner, HU Berlin Winter-Semester 2008/09

Abbildung

3 Beteiligte:



Betrachter: Mensch

SZENE

Licht von Objekten: Strahlung, Reflexion

Ausbreitung (Welle/Teilchen)

Erfassung (Auge)

Verarbeitung (Gehirn)

ABBILD

Betrachter: Technisch/Chemisch

SZENE

Licht von Objekten: Strahlung, Reflexion

Ausbreitung (Welle/Teilchen)

Erfassung (Kamera: Film Belichten)

Verarbeitung (Entwickeln, Drucken)

Techn. ABBILD

Betrachter: Technisch/Elektronisch

SZENE

Licht von Objekten: Strahlung, Reflexion

Ausbreitung (Welle/Teilchen)

Erfassung (Kamera: Pixel)

Verarbeitung (Bildschirm, Drucken)

Techn. ABBILD

Abbildung

SZENE

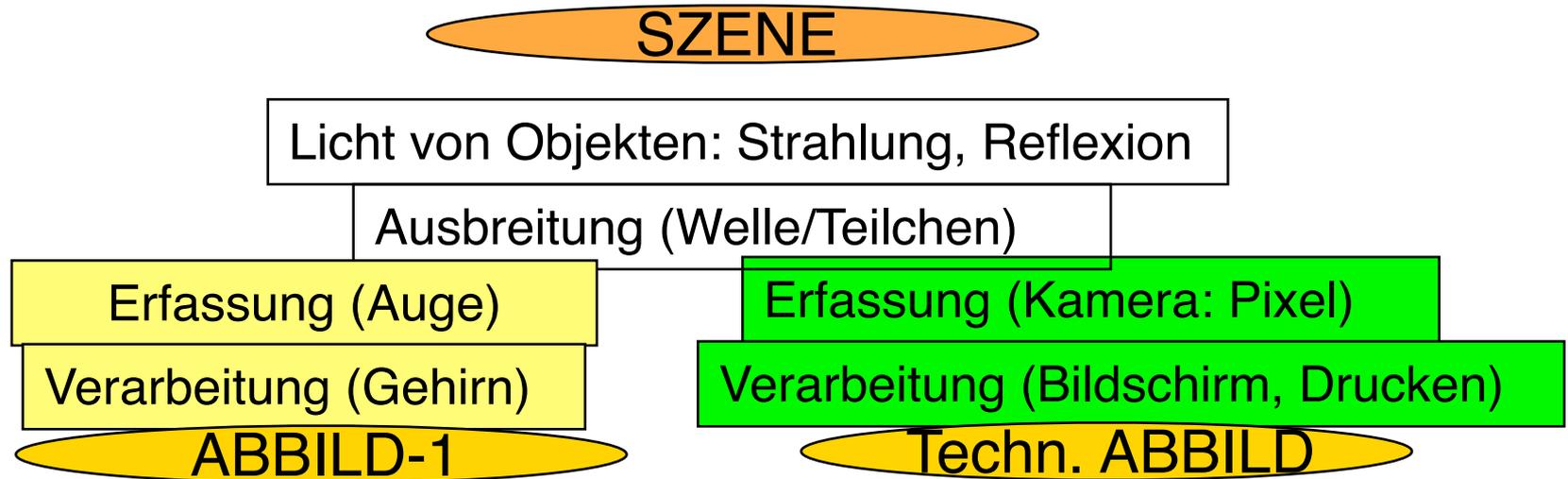
Was passiert tatsächlich?
Wieweit stimmen
SZENE und ABBILD überein?

(Techn.) ABBILD

Abbildung



Abbildung



Abbildung

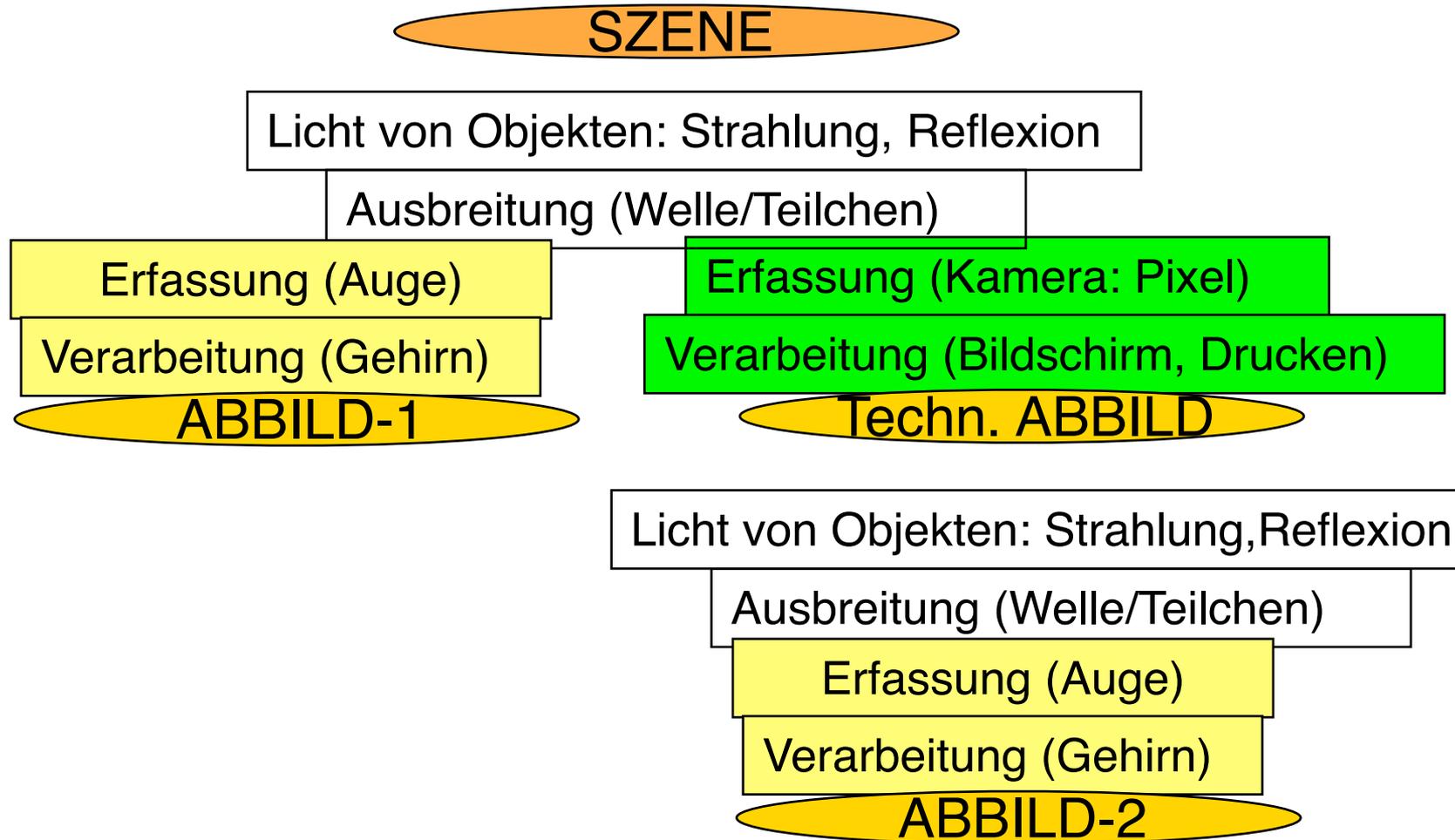


Abbildung: Anforderungen

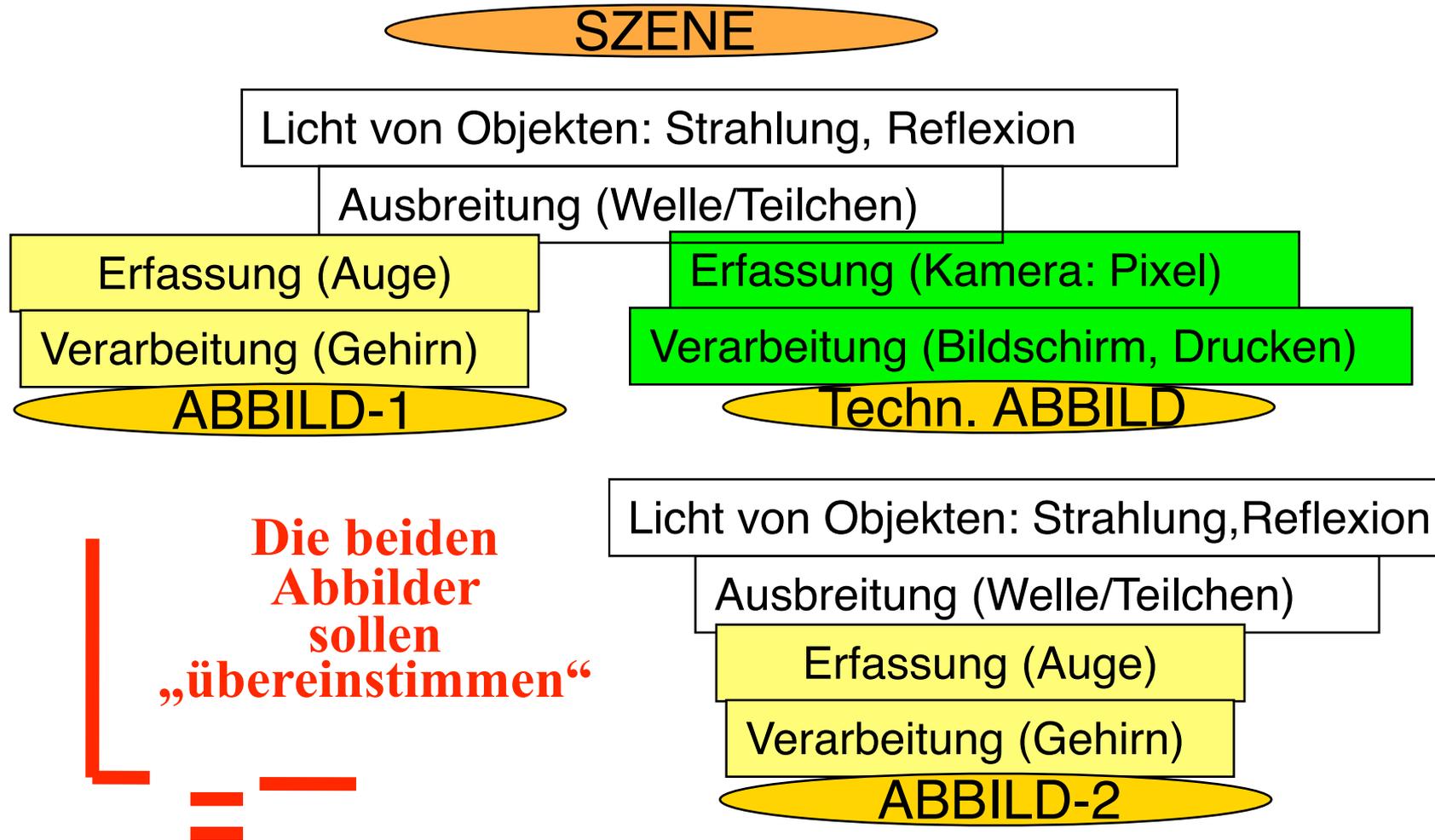


Abbildung: Anforderungen

„Übereinstimmung“ mit unterschiedlichen Bezügen

- Naturgetreue Wiedergabe
- Schematisch
- Emotional
- ...

**Die beiden
Abbilder
sollen
„übereinstimmen“**

Abbildung: Anforderungen

Maßstab für „Übereinstimmung“ ist die menschliche Betrachtung.

Technik soll Bild wiedergeben, das vom Menschen in Bezug auf bestimmte Aspekte „ähnlich“ wahrgenommen wird wie die Szene.

Darüber hinaus wird i.a. keine Übereinstimmung gefordert.

Ein Roboter könnte mit anderen Sehfähigkeiten ausgestattet sein.

Künstler

SZENE

Licht von Objekten: Strahlung, Reflexion

Ausbreitung (Welle/Teilchen)

Erfassung (Auge)

Verarbeitung (Gehirn)

Verarbeitung (Zeichnen, ...)

BILD



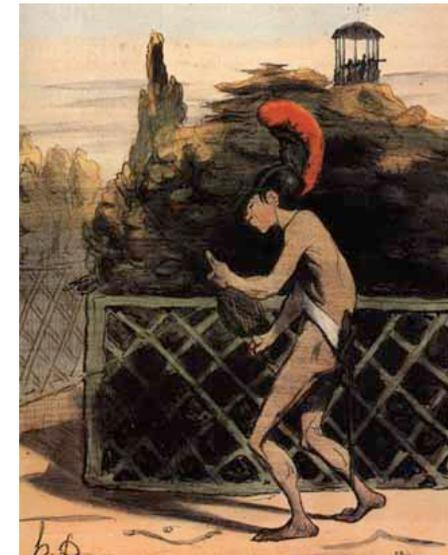
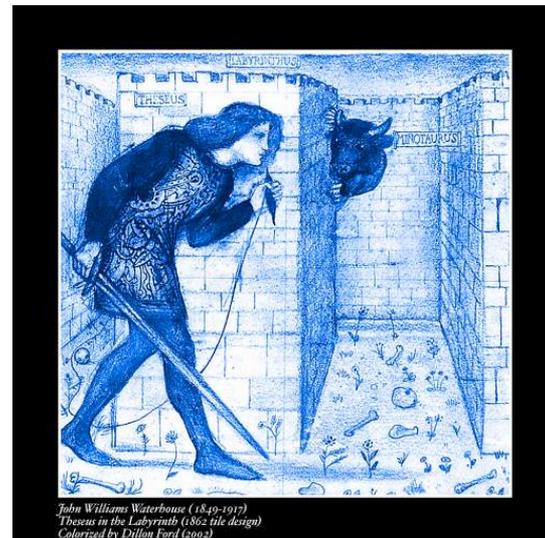
Künstler und Betrachter



**Unterschiedliche
Möglichkeiten für
Beziehungen zwischen
Abbild 1 und Abbild 2**

Künstler und Betrachter

Theseus und der Minotaurus

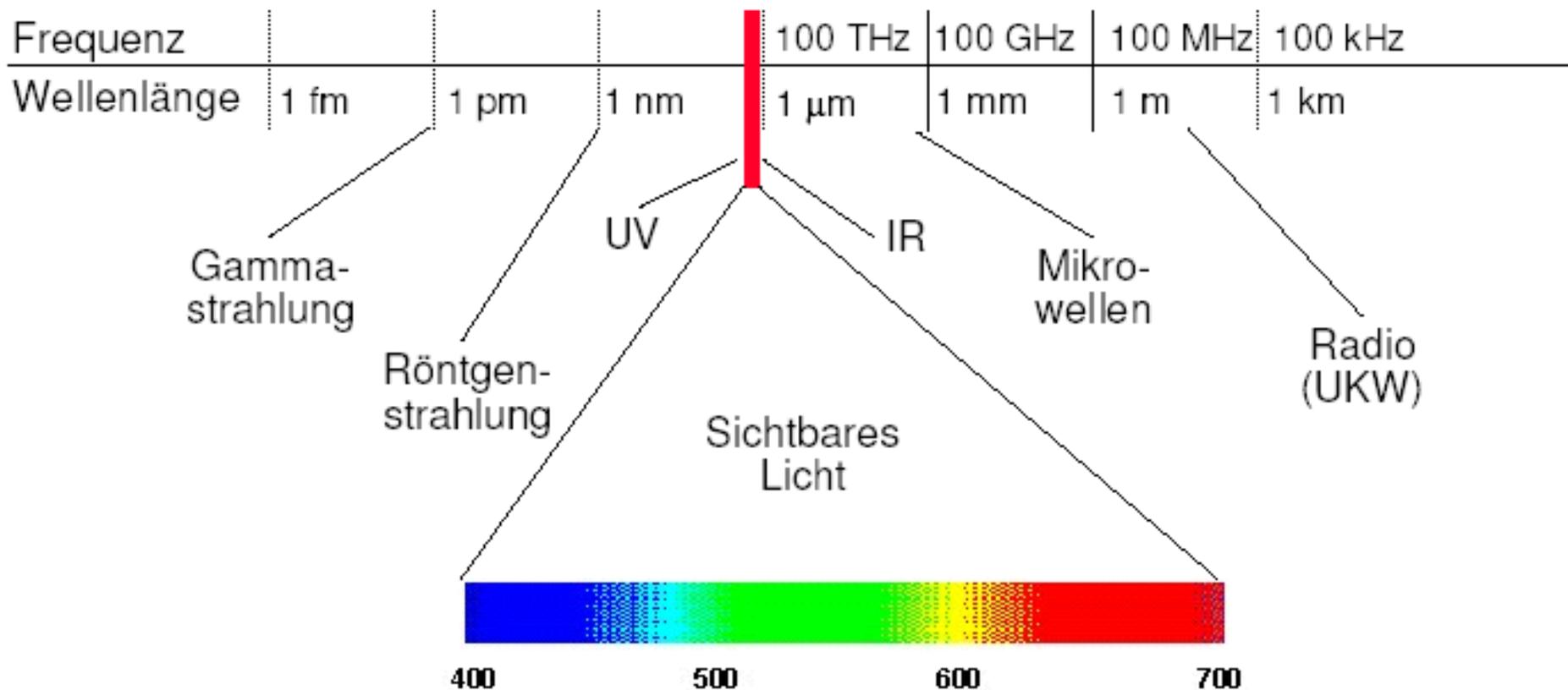


Licht: Elektromagnetische Strahlung

Für Menschen sichtbares Licht:

Wellenlängen zwischen 380 nm und 780 nm

Spektrum der elektromagnetischen Strahlung



Licht

Dualer Charakter (2 Modelle):

Welle:

Brechung, Beugung, Dispersion, Streuung

Teilchen (Photonen):

Absorption, Emission

Wellenlänge λ , Frequenz f , Periodendauer T

$$f = 1 / T \text{ [Hz]}$$

$$T = \lambda / c \text{ [s]}$$

$$f = c / \lambda \text{ [Hz]}$$

$$c = 2,998 \cdot 10^8 \text{ m/s (Lichtgeschwindigkeit)}$$

Licht

Geradlinige Ausbreitung

Oberflächen reflektieren unterschiedlich:

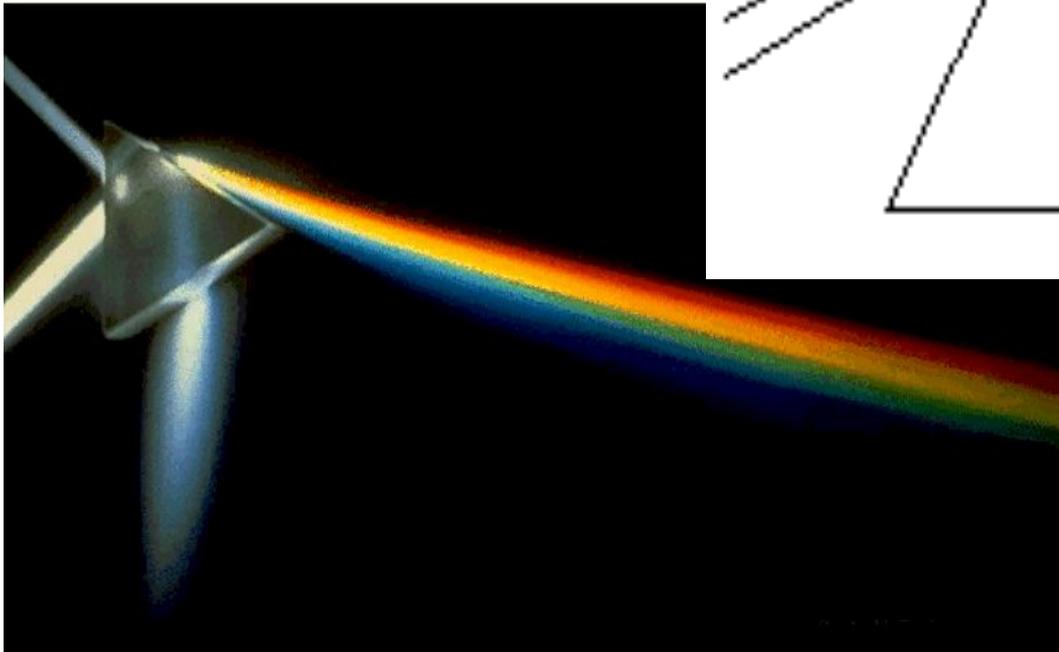
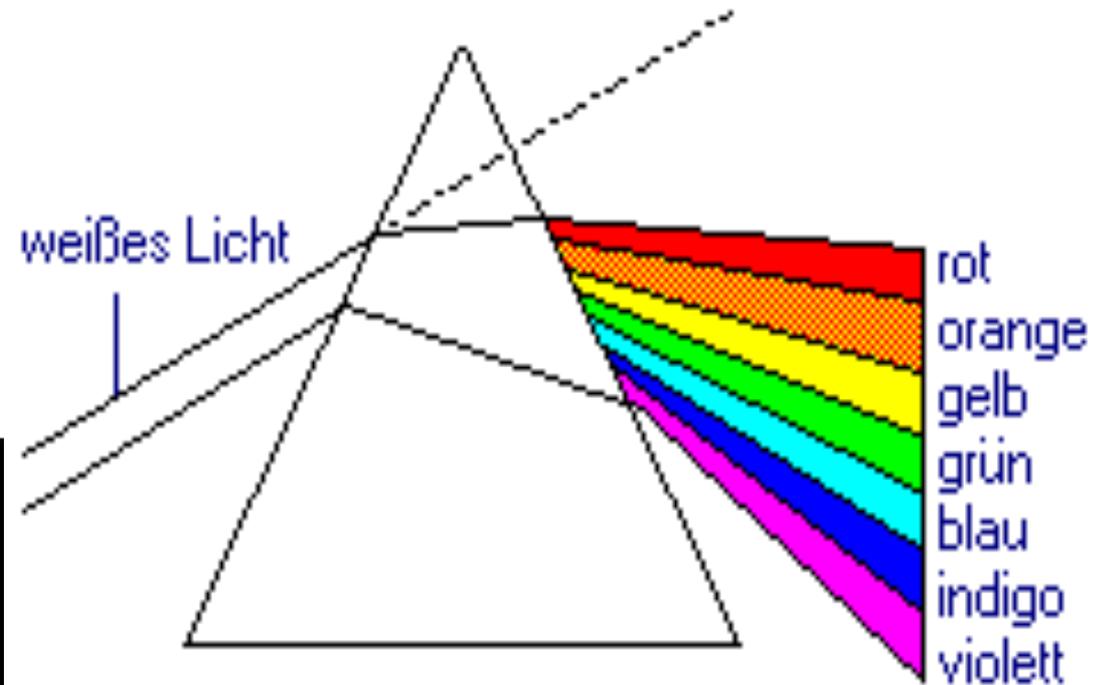
- Reflexionskoeffizient (abhängig von der Wellenlänge)
- Rauigkeit

Optische Dichte von Materie:

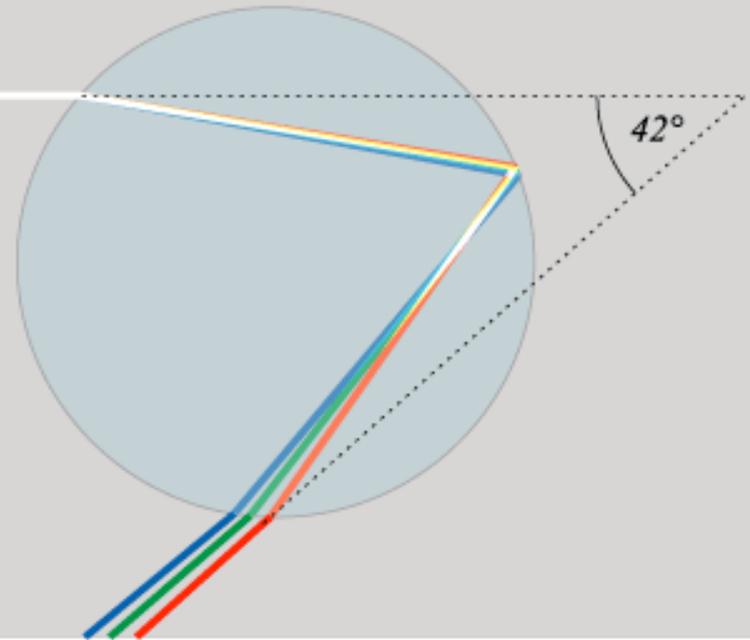
- Brechungsindex beschreibt niedrigere Ausbreitungsgeschwindigkeit gegenüber Lichtgeschwindigkeit
- Bei Eintritt in optisch dichtere Materie erfolgt Beugung des Lichtstrahls zur Senkrechten
- Dispersion: Abhängigkeit des Brechungsindex von der Wellenlänge (Regenbogeneffekt)

Licht

Zerlegung von weißem Licht in Spektralfarben



Regenbogen



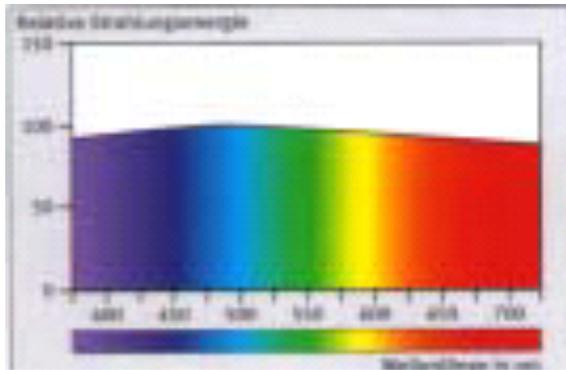
Licht

Reale Strahlungsquellen strahlen fast immer eine Mischung verschiedener Frequenzen aus.

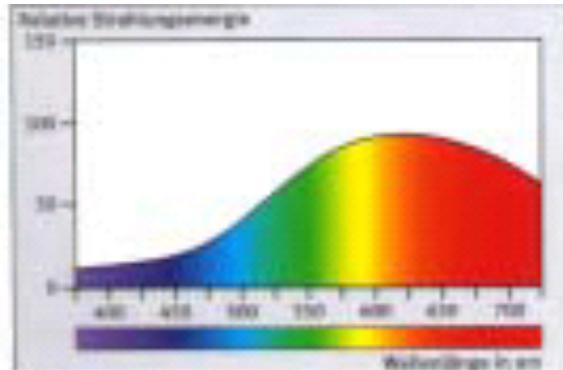
Das von Körpern reflektierte Licht hängt ab vom eingestrahlten Licht und von den Absorptionseigenschaften (nicht alle Frequenzen werden reflektiert).

Licht

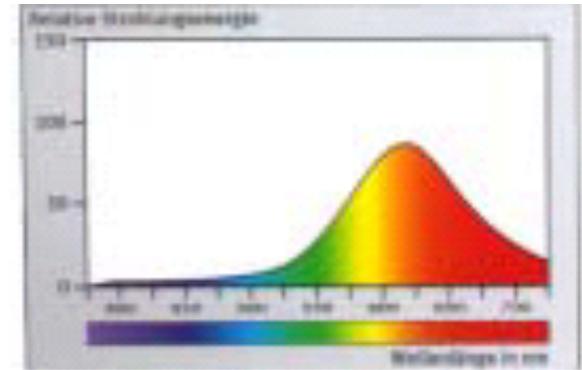
Intensität als Funktion der Wellenlänge: $I(\lambda)$



Tageslicht



Glühlampe



Rotes Licht

Unendlich-dimensionaler Raum L von Möglichkeiten.
Basis von L : Wellenlängen des Lichts

Sensor, Rezeptor

Unterschiedliche Farbempfindlichkeit

Technisch:

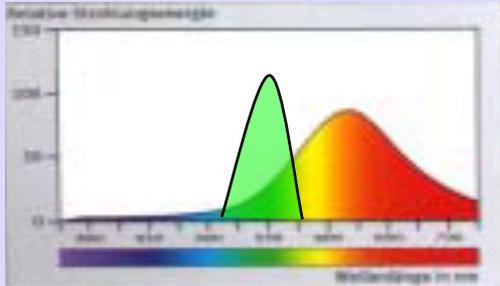
z.B. CCD (Ladung proportional zur Lichtstärke) + Filter
(nur durchlässig für eingeschränkten Frequenzbereich)

Auge:

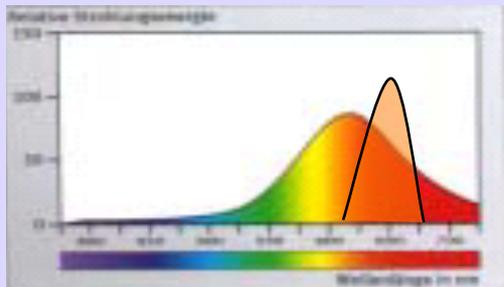
Zapfen, Stäbchen (durch Licht ausgelöste chemische
Prozesse an Nerven übermitteln)

Reduktion des unendlich dimensionalen Lichtraums auf
eine Dimension (Erregungsskala des Sensors)

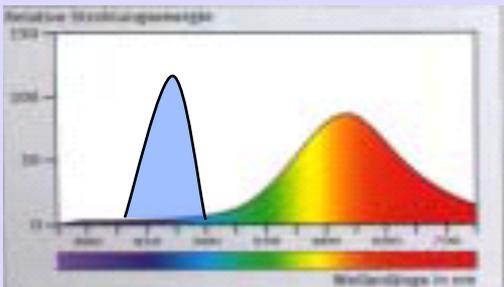
Sensor, Rezeptor: Beispiele



Erregung: Grünempfindlicher Sensor



Erregung: Rotempfindlicher Sensor



Erregung: Blauempfindlicher Sensor

Sensor, Rezeptor

Erregung (Messung am Sensor):

Reaktion auf Licht in bestimmten Frequenz-Bereichen

Empfindlichkeit abhängig von Frequenz: $\phi(\lambda)$

Erregung e insgesamt abhängig von

eingestrahlttem Licht $I(\lambda)$ und Empfindlichkeit $\phi(\lambda)$

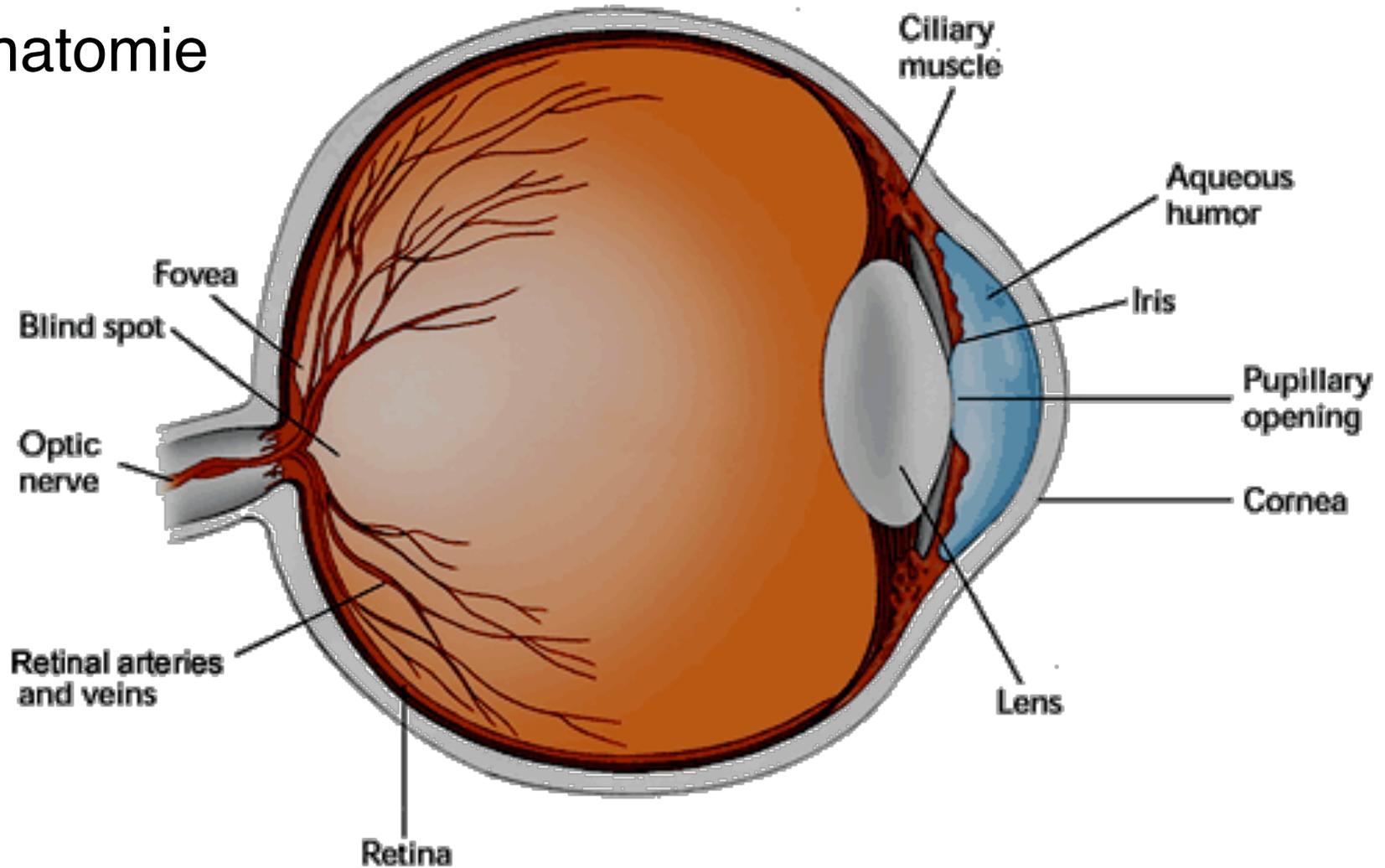
für unterschiedliche Frequenzen λ .

Bei Annahme linearer Beiträge und ohne Berücksichtigung von Schwelle/Sättigung realer Sensoren näherungsweise beschreibbar durch

$$e = \int \phi(\lambda) I(\lambda) d\lambda$$

Menschliches Sehen: Das Auge

Anatomie



Aufbau der Netzhaut (Retina)

Stäbchen (*Rods*)

Schwarzweiß-Sehen

Sehr empfindlich

ca. 120 Millionen

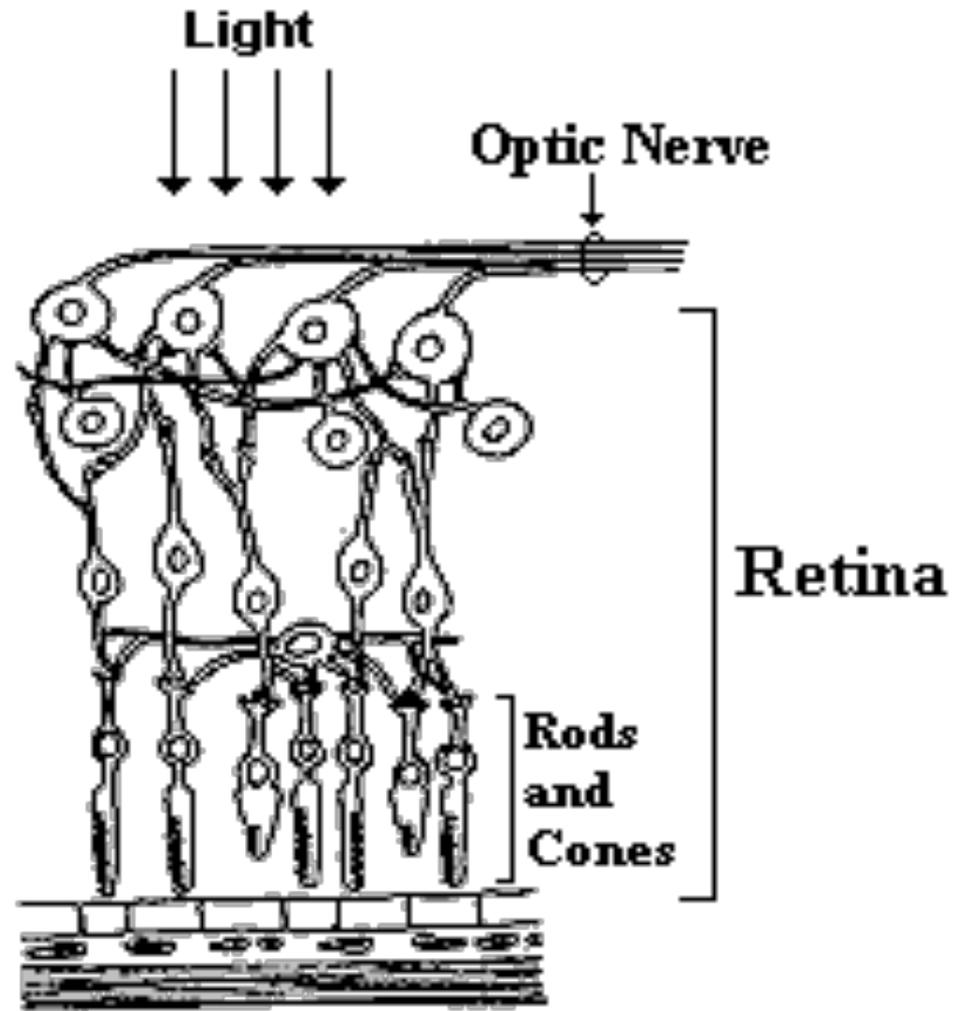
Zapfen (*Cones*)

Farben-Sehen

ca. 7 Millionen

Nerven

Blutgefäße

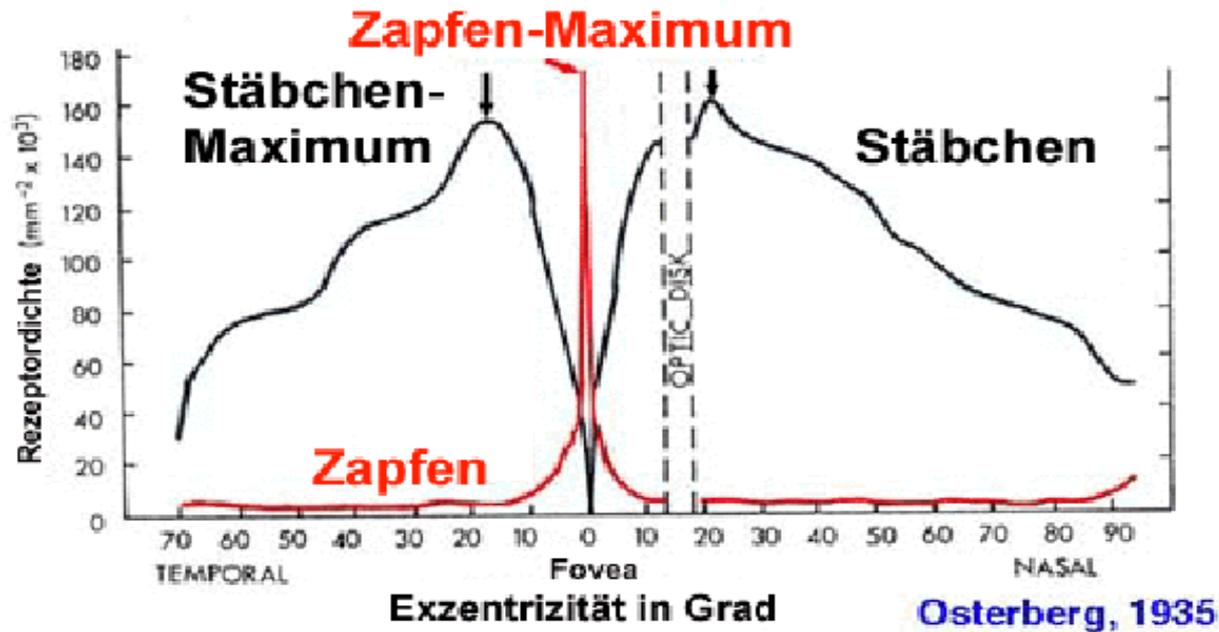


Verteilung Stäbchen/Zapfen

Stäbchen (SW) in der Peripherie

Zapfen (Farbe) im Zentrum (fovea centralis)

Dämmerungssehen in der Peripherie besser als im Zentrum



Verteilung der Zapfen

Blaurezeptoren:

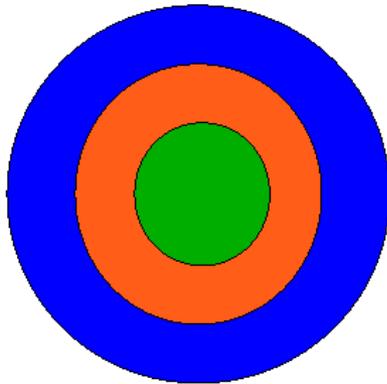
Ca. 4%, nur peripher, nicht im Zentrum

Grünrezeptoren:

Ca. 32 %, im zentralen Bereich konzentriert

Rot-(bzw. Gelb-)rezeptoren:

Ca. 64%, mittlerer Abstand vom Zentrum



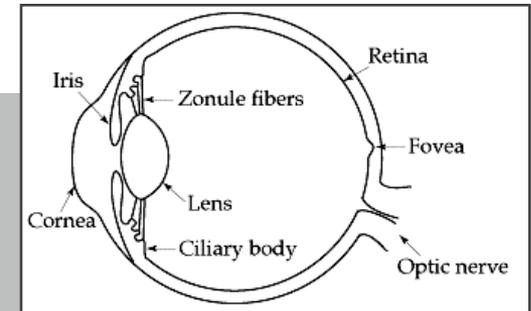
Detailreiche Information (Text) in reinem Blau ist schwer wahrzunehmen.

In der Peripherie sind rot und grün schwerer wahrzunehmen als blau.

Blinder Fleck: Austritt des Sehnervs

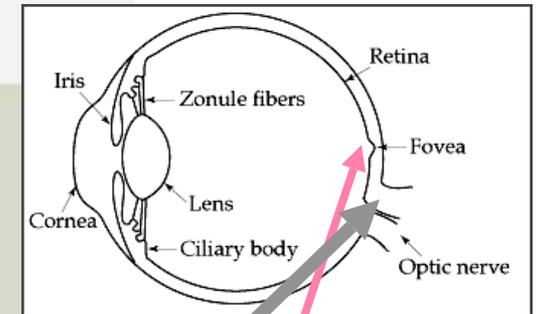
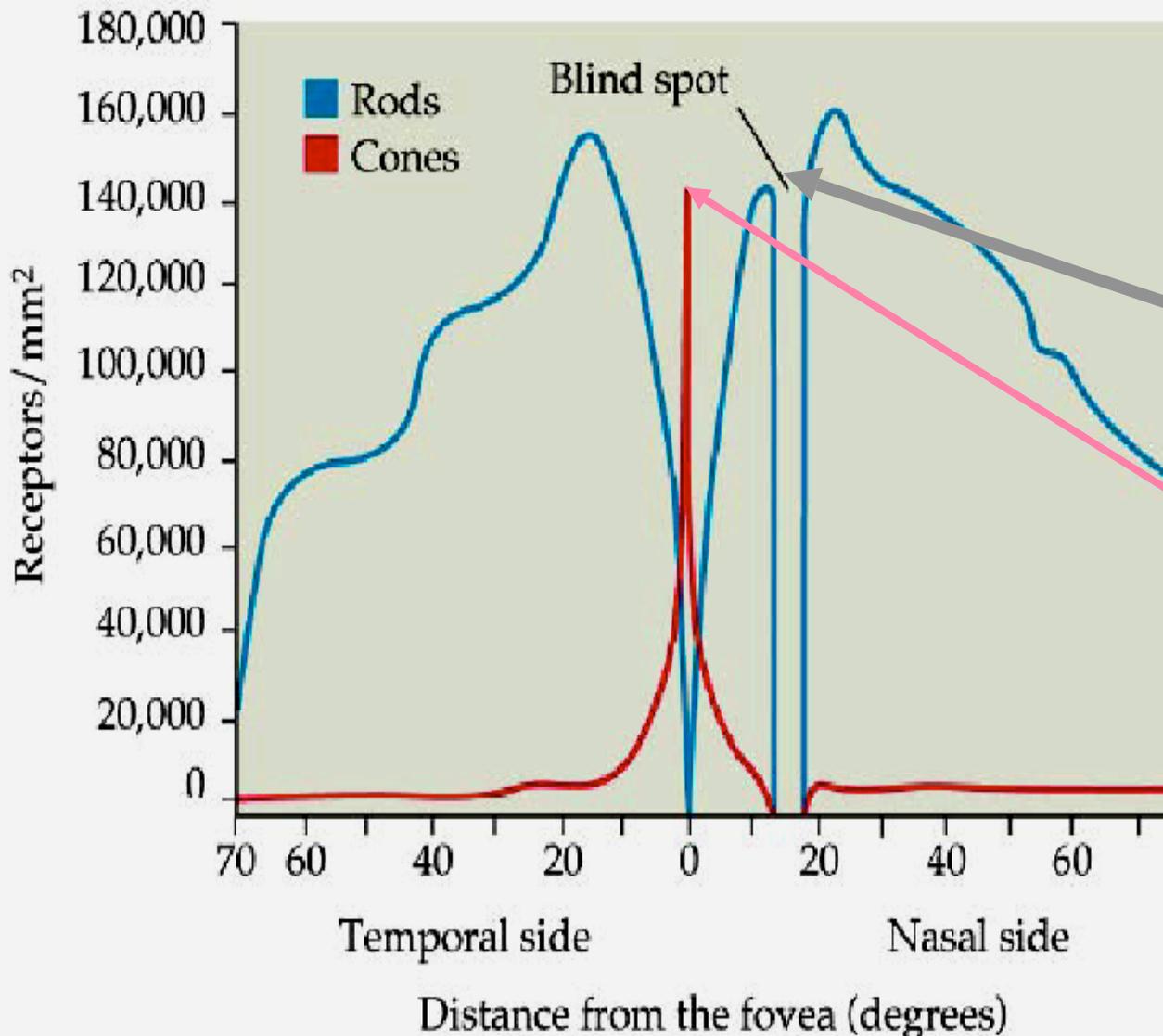


- Kreis mit dem rechten Auge fixieren
- Abstand zum Diagramm variieren
- Bei einem bestimmten Abstand verschwindet das Kreuz



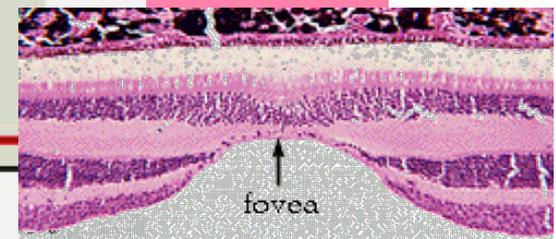
Fovea: Bereich höchster Auflösung

(a) Distributions of rods and cones across the retina



Blinder Fleck

Fovea



Anpassung

Anpassung der Augen an die Helligkeit der Umgebung
(mit Verzögerung)

Helladaptation:

- Pupillenverengung
- Übergang vom Stäbchen- zum Zapfensehen

Dunkeladaptation:

- Übergang vom Zapfen- zum Stäbchensehen
- Stäbchenadaptation
- Stäbchen außerhalb Zentrum: Bei Dunkelheit im Zentrum schwächere Unterscheidung als daneben

Brennweite der Augenlinse

abhängig von der Wellenlänge

groß im roten Bereich

klein im blauen Bereich

Bild mit roten und blauen Bereichen: Auge ermüdet.

Rot vor blauem Hintergrund und umgekehrt vermeiden!

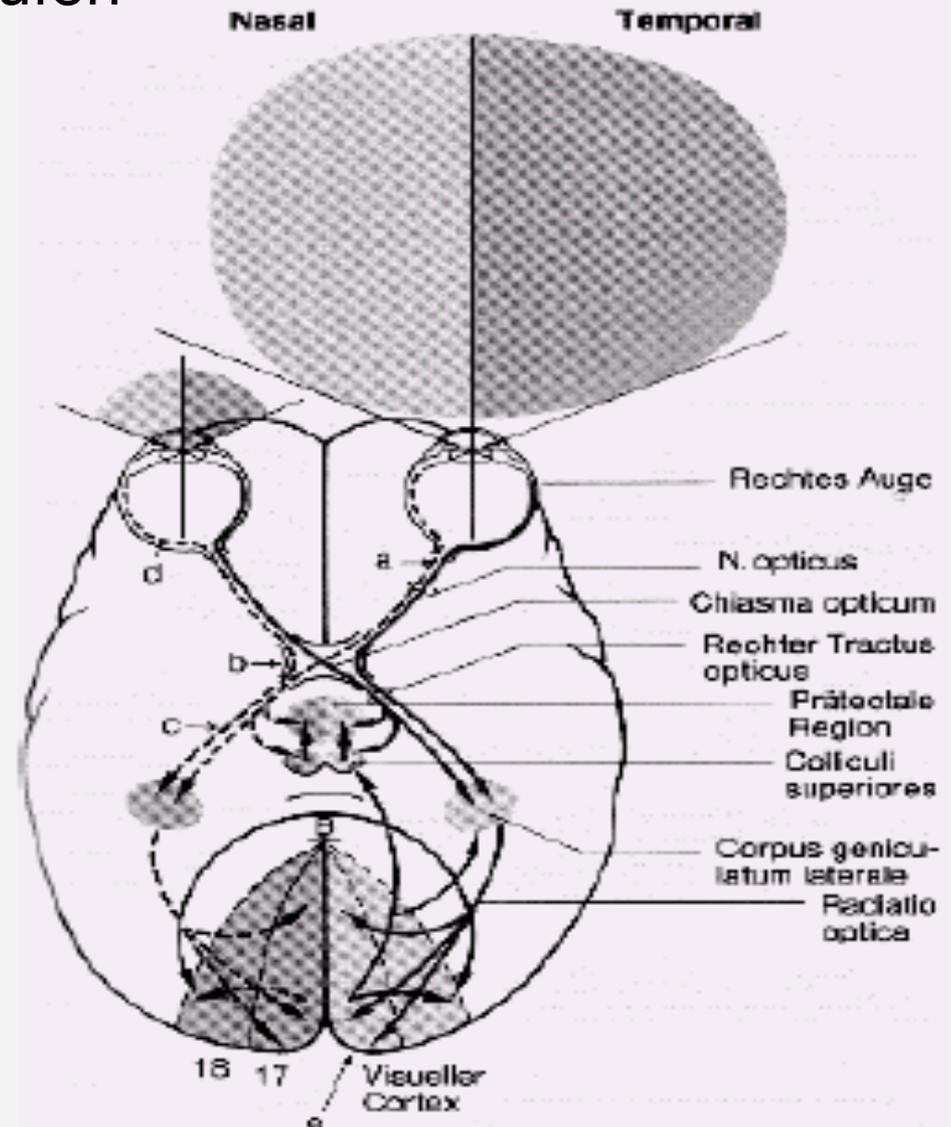
Räumliche Wahrnehmung: Rot vorn, blau hinten



Menschliches Sehen: Das Gehirn

Verarbeitung in mehreren Stufen

Steuerung der Bewegungen



Verarbeitung im menschlichen Gehirn

Im Zwischenhirn werden die separaten Signale für Rot (R), Grün (G), Blau (B) in drei neue Signale umgewandelt:

Luminanzsignal: $Y = R + G$

Helligkeit (Gelb)

Differenzsignal: $R - G$

Rot/Grün-Unterscheidung

Differenzsignal: $Y - B$

Blau/Gelb-Unterscheidung

Gelb-Anteil wesentlich für Helligkeitswahrnehmung

Blau-Anteil unwesentlich für Helligkeitswahrnehmung

Farbkontraste Rot/Grün und Blau/Gelb besonders klar



"Komplementärfarben"

Sakkaden

Rasches Anvisieren von interessierenden Merkmalen:
(Korrektur der Sichtlinie)

Während der Augenbewegung keine Wahrnehmung
Unwillkürlich (schnell) und willkürlich möglich



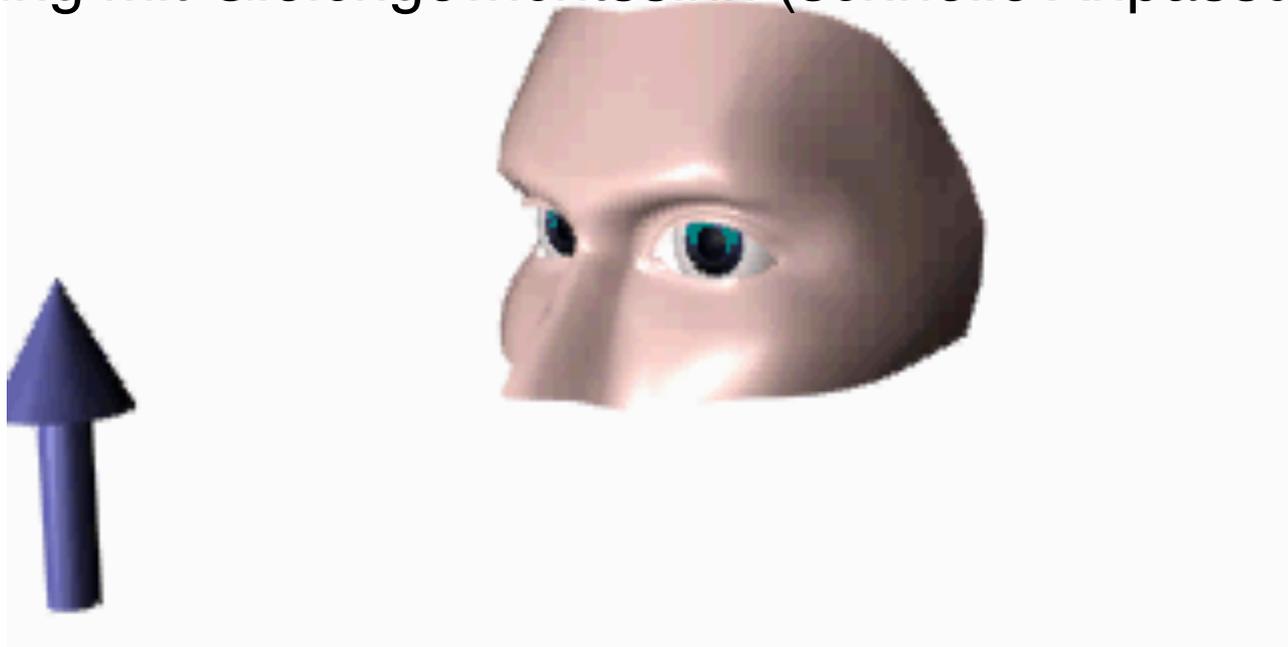
Vestibulo-Okulärer Reflex

Kompensation von Eigenbewegungen

Augen werden entgegen der Kopfbewegung bewegt

Blickrichtung bleibt annähernd konstant

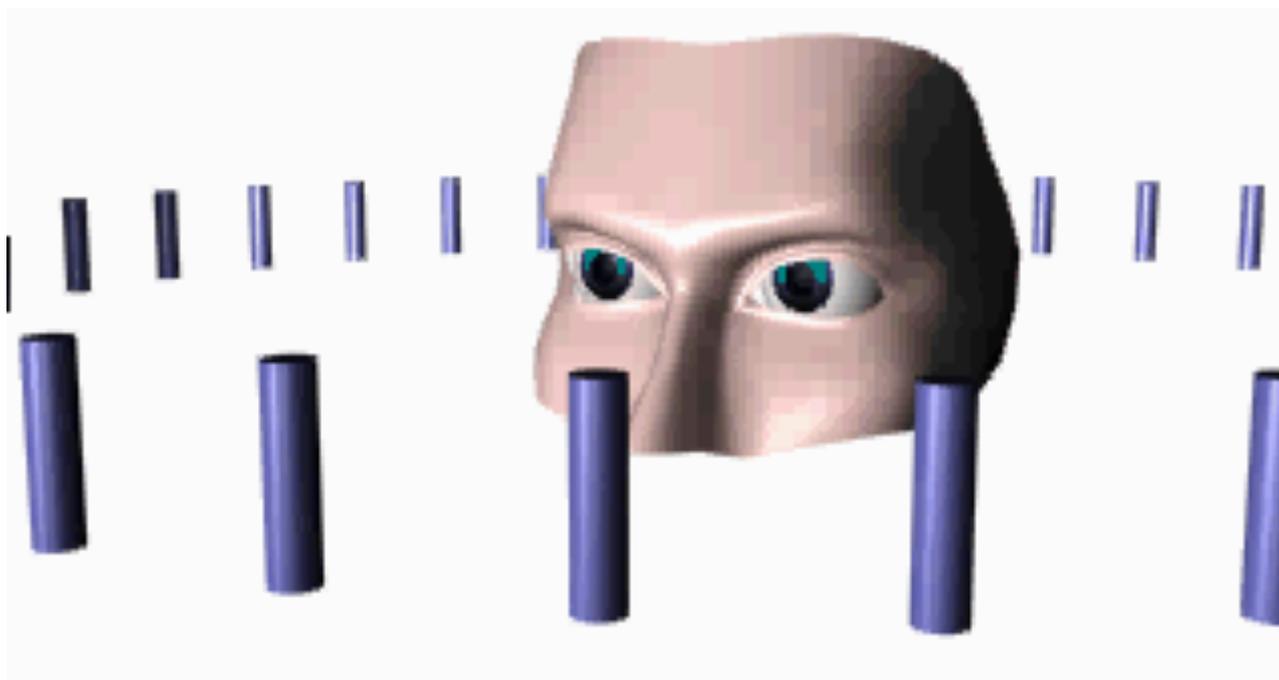
Rückkopplung mit Gleichgewichtssinn (schnelle Anpassung)



Optokinetischer Reflex

Ausgleich langanhaltender Bewegungen (wenn Gleichgewichtsänderung abklingt)

Rückkopplung über visuelles System

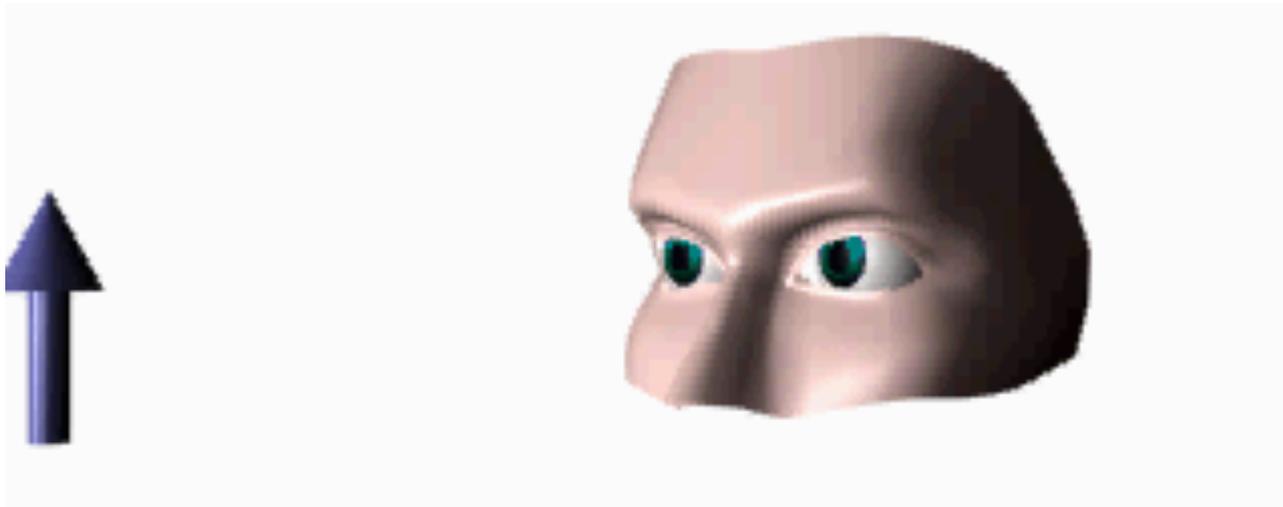


Augenfolgebewegung (*Smooth Pursuit*)

gleitende Nachführung des Auges

dadurch wird das Bild von bewegten Objekten dauerhaft auf einer Stelle der Netzhaut gehalten

Rückkopplung mit unterschiedlichen Bereichen

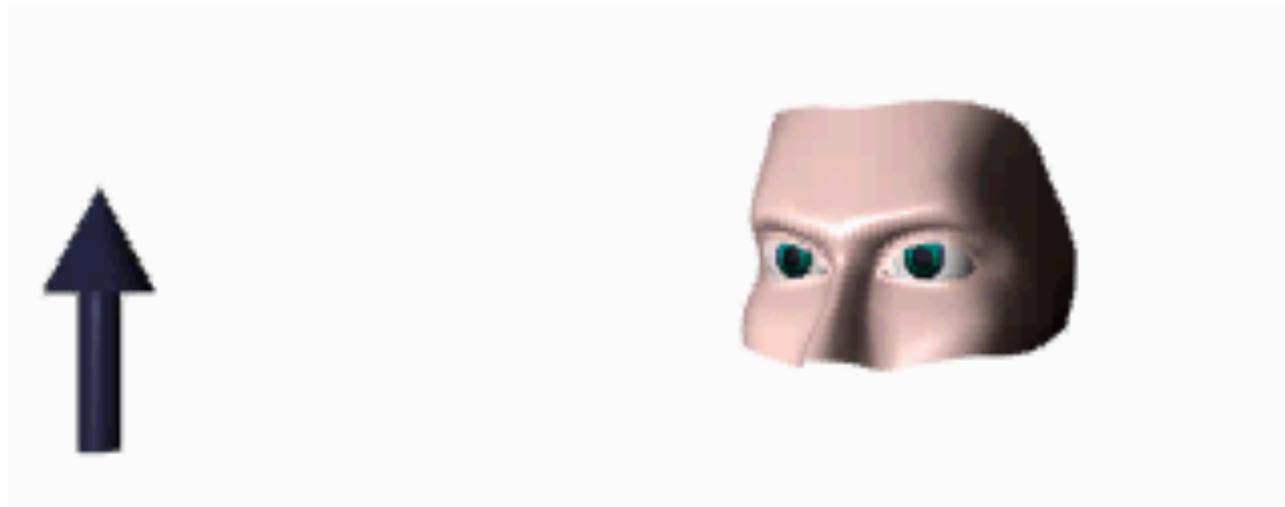


Stereo-Sehen, Vergenz

Koordination zwischen den Bewegungen beider Augen
binokuläres Erfassung von Objekten

(Korrespondenz-Problem, entfernungsabhängig)

räumliches Sehen



Menschl. Auge: Aufbau der Netzhaut

Stäbchen (*Rods*)

Schwarzweiß-Sehen (Sehr empfindlich!)

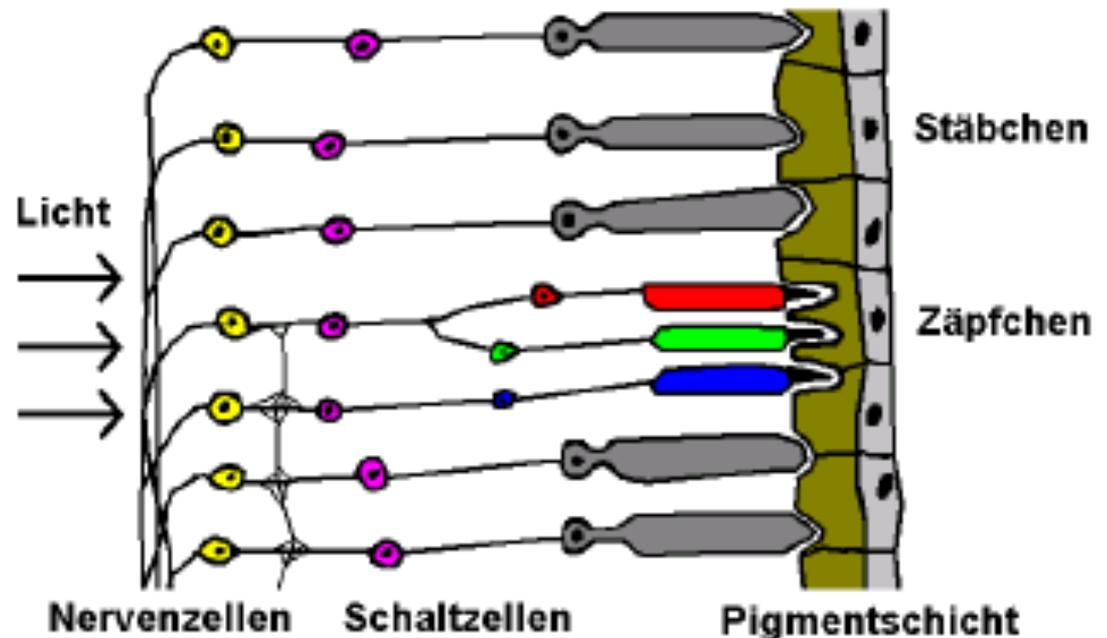
Zapfen (*Cones*)

Farben-Sehen

Helligkeit

Nerven

Blutgefäße



Helligkeitsempfinden

1-dimensionaler Empfindungsraum

Helligkeit abhängig von Intensität (Energie) des Lichts.

Unterschiedliche Empfindlichkeit der Sehzellen im Auge.

Zapfen: Hohe Empfindlichkeit bei grün, geringe bei blau

Stäbchen: Dämmerungssehen

Helligkeitsempfinden

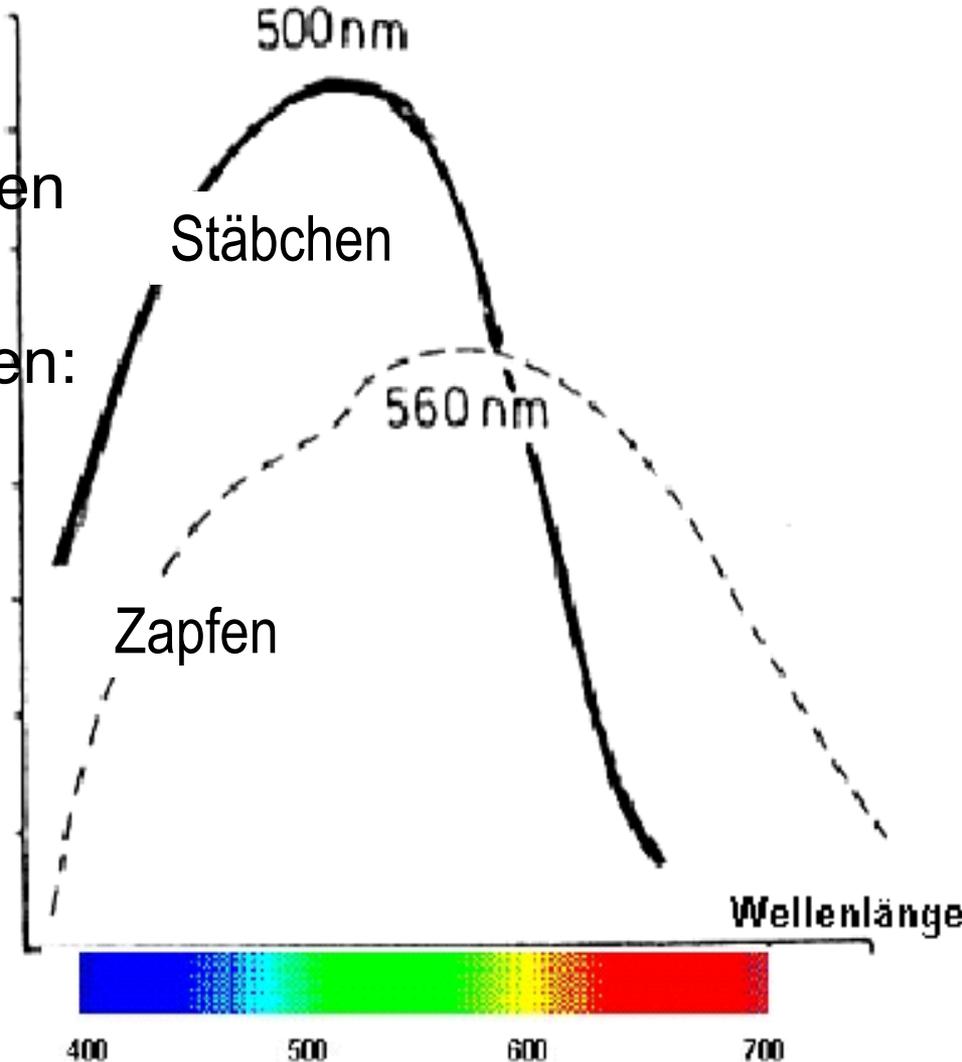
Stäbchen sind "farbenblind",
aber lichtempfindlicher als Zapfen

Dämmerungssehen mit Stäbchen:

Blau = hellgrau,
rot = schwarz

„Nachts sind alle Katzen grau.“

Empfindlichkeit

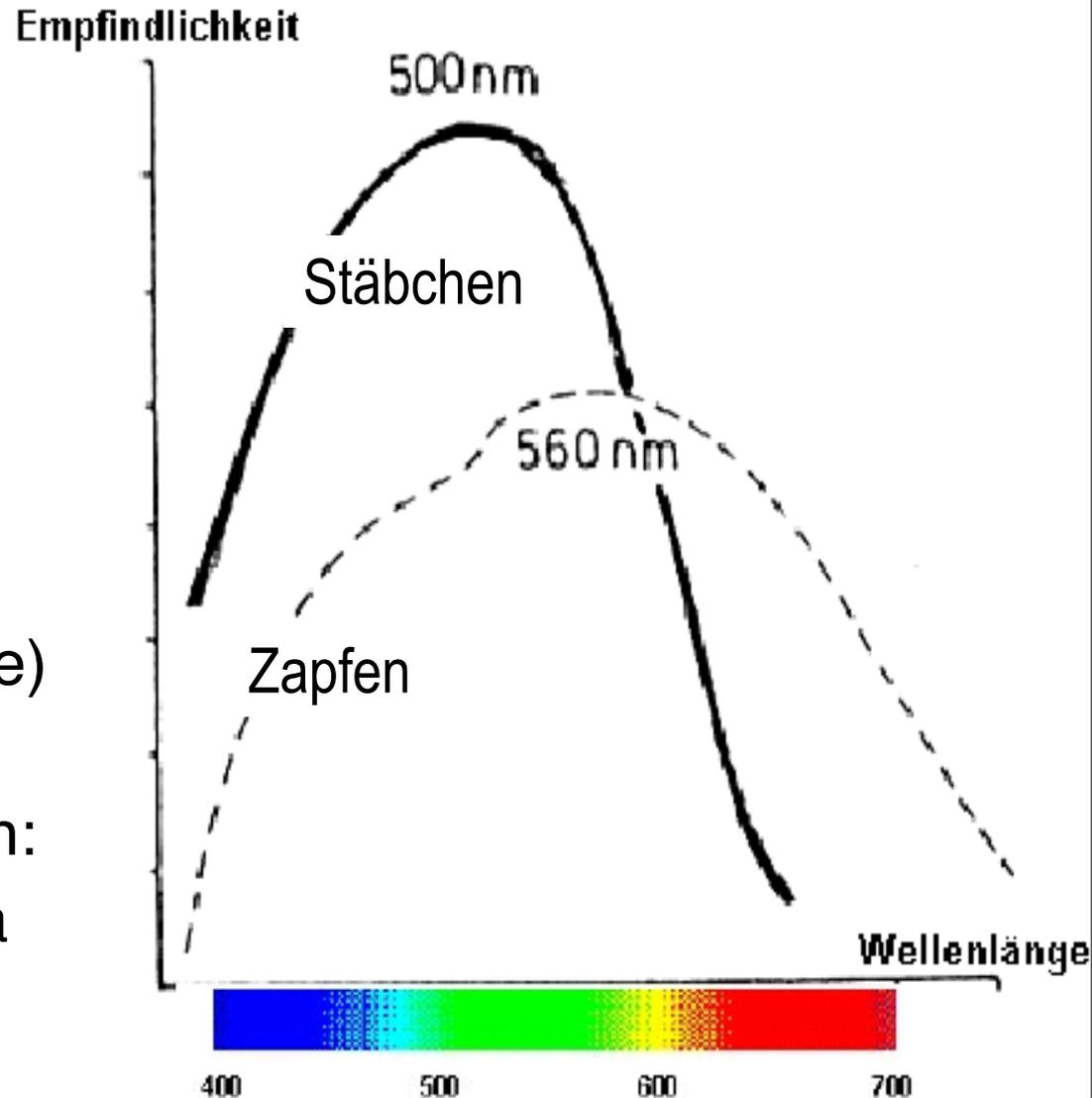


Helligkeitsempfinden

Zapfen

Tagsehen mit Zapfen:
Größte Empfindlichkeit
für Gelb-Grün-Töne (Sonne)

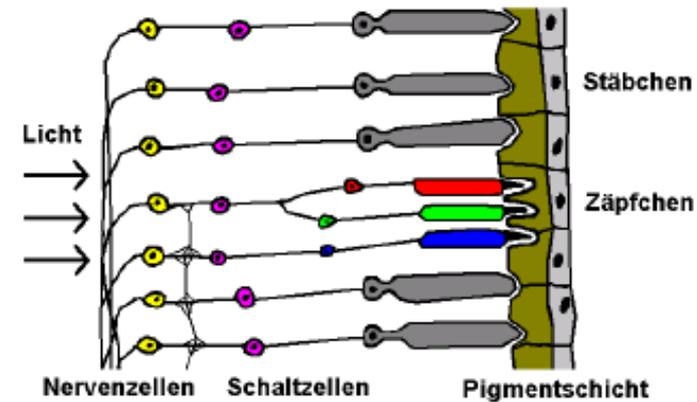
Scharfes Sehen mit Zapfen:
Größte Häufigkeit in Fovea



Farbsehen mit Zapfen

3-dimensionaler Empfindungsraum

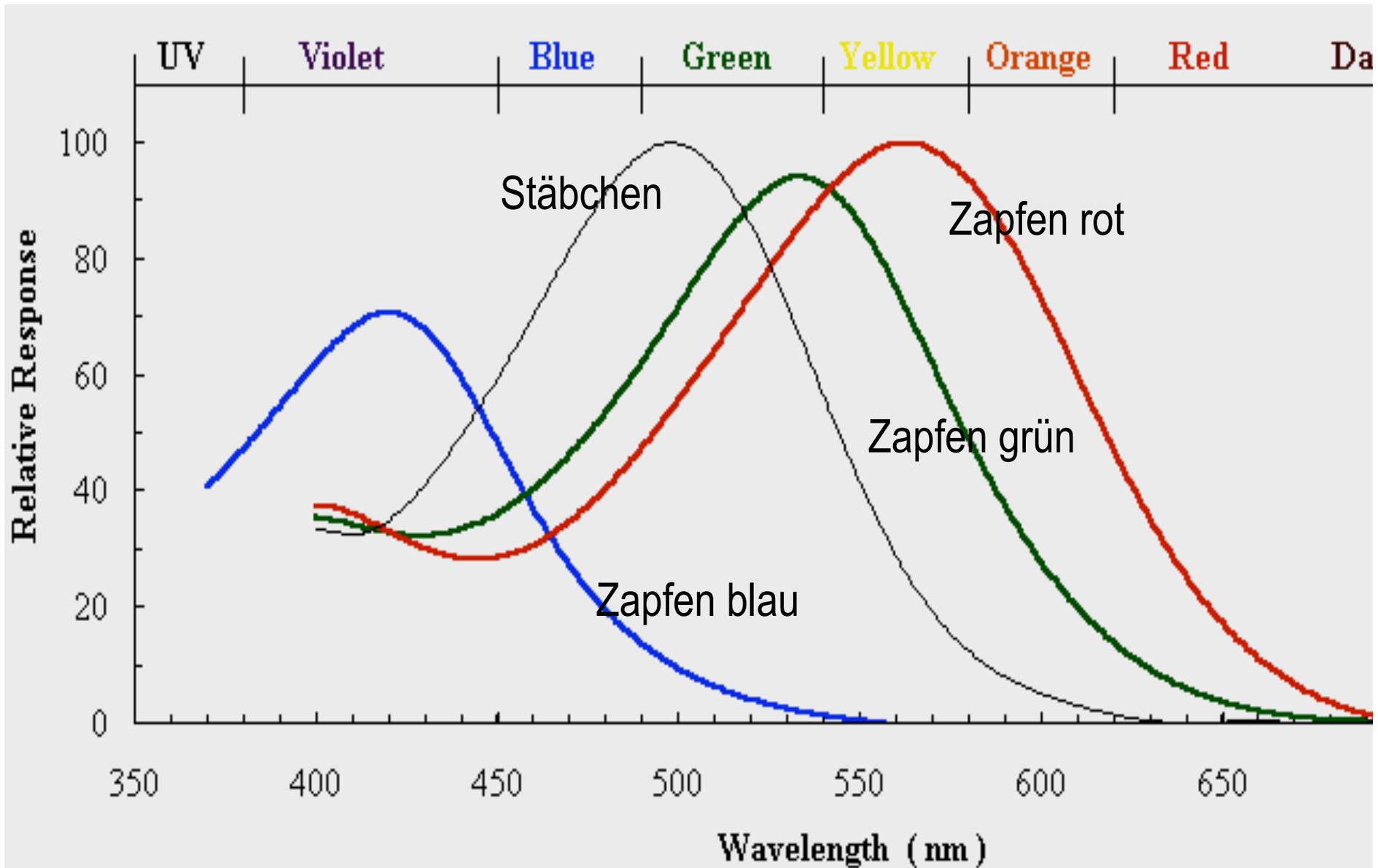
- Rot
- Grün
- Blau



Retina enthält 3 Typen farbsensitiver Zapfen mit maximaler Sensitivität bei:

- 430 nm (ca. 4% der Zapfen, blausensitiv)
- 530 nm (ca. 32% der Zapfen, grün-sensitiv)
- 560 nm (ca. 64% der Zapfen, rot-sensitiv)

Farb-Empfindlichkeit



Farb-Empfindlichkeit

Objektive Bewertung der Erregung problematisch.

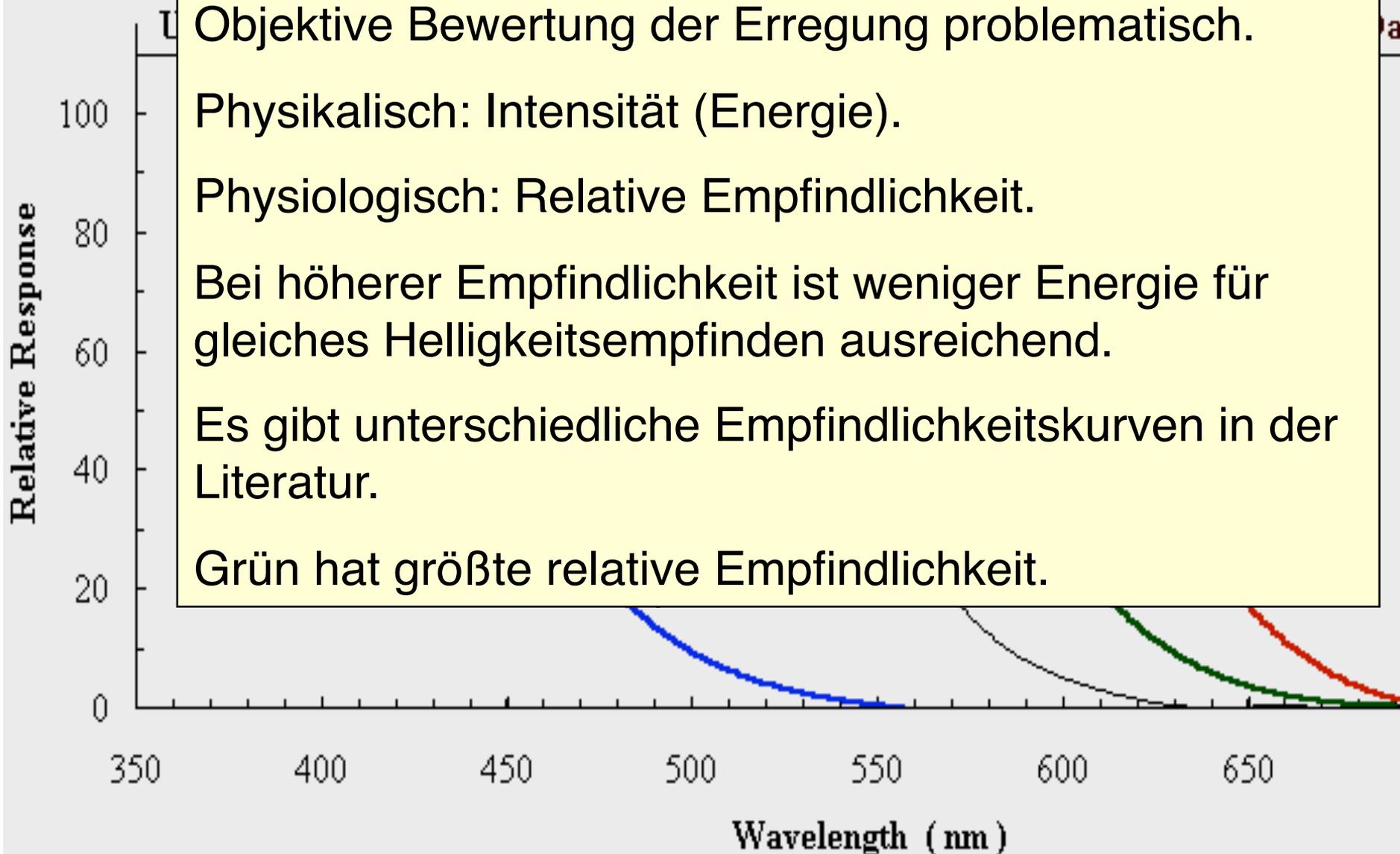
Physikalisch: Intensität (Energie).

Physiologisch: Relative Empfindlichkeit.

Bei höherer Empfindlichkeit ist weniger Energie für gleiches Helligkeitsempfinden ausreichend.

Es gibt unterschiedliche Empfindlichkeitskurven in der Literatur.

Grün hat größte relative Empfindlichkeit.



Farbsehen mit Zapfen

Reduktion der Dimension:

3-Dimensionaler Sensor-Raum für Erfassung des unendlich-dimensionalen Raums L des Lichts.

Unterschiedliche Spektralverteilungen können gleiche Reaktion der Zapfen hervorrufen.

Unterschiedliches Licht kann gleiche Farbempfindung verursachen: Metameres Licht

Farbwahrnehmung

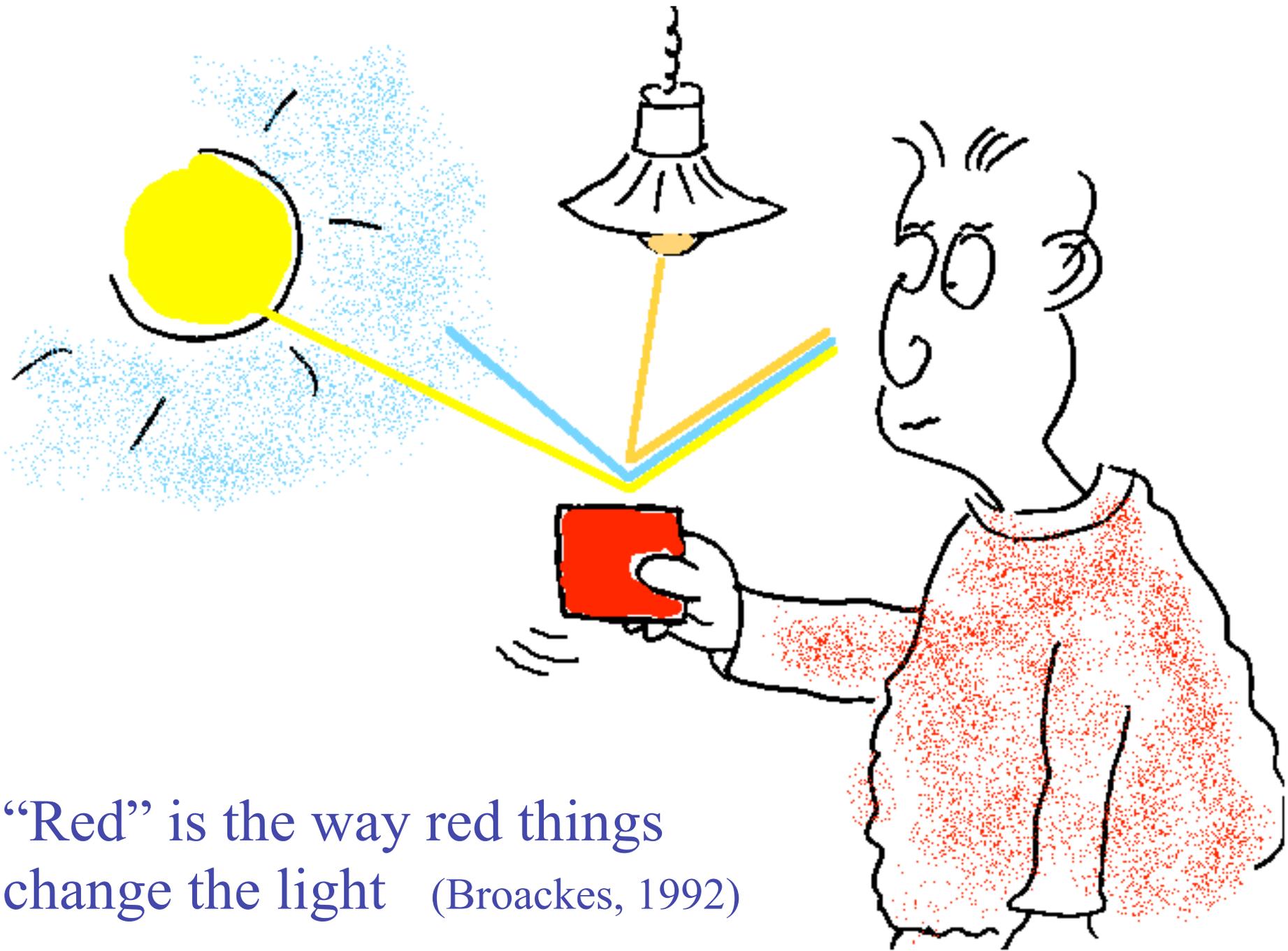
Farbwahrnehmung wird bestimmt durch die Rezeptoren für rot, grün, blau.

Deren Reizung wird im Gehirn weiterverarbeitet und führt zu unterschiedlichen Wahrnehmungen.

Bei der Weiterverarbeitung werden die ursprünglichen Empfindungen miteinander verrechnet (bessere Ausnutzung der Informationskanäle).

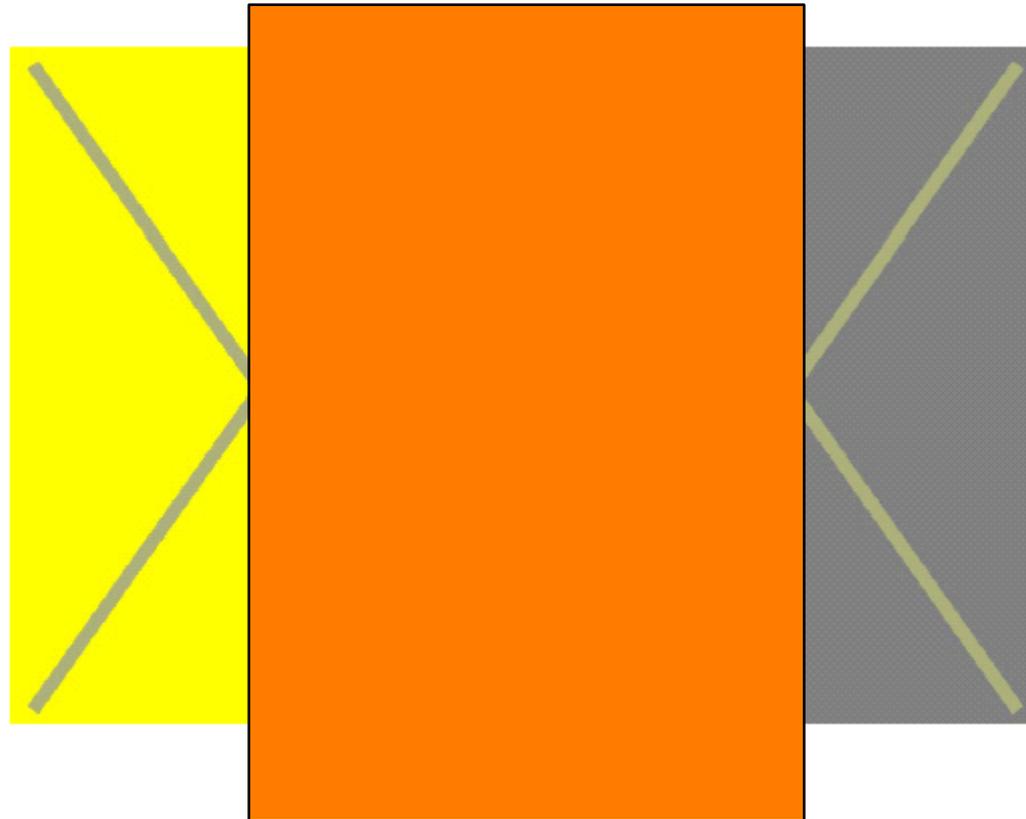
Außerdem gibt es „Ergänzungen“ (können zu Farbillusionen führen).

Die **Farbwahrnehmung** ist ein physiologisches Phänomen.
(Farben existieren nur im Gehirn, nicht in der Natur.)



“Red” is the way red things
change the light (Broackes, 1992)

Farbwahrnehmung



Die **Farbwahrnehmung** ist ein physiologisches Phänomen.
(Farben existieren nur im Gehirn, nicht in der Natur.)

Farbwahrnehmung

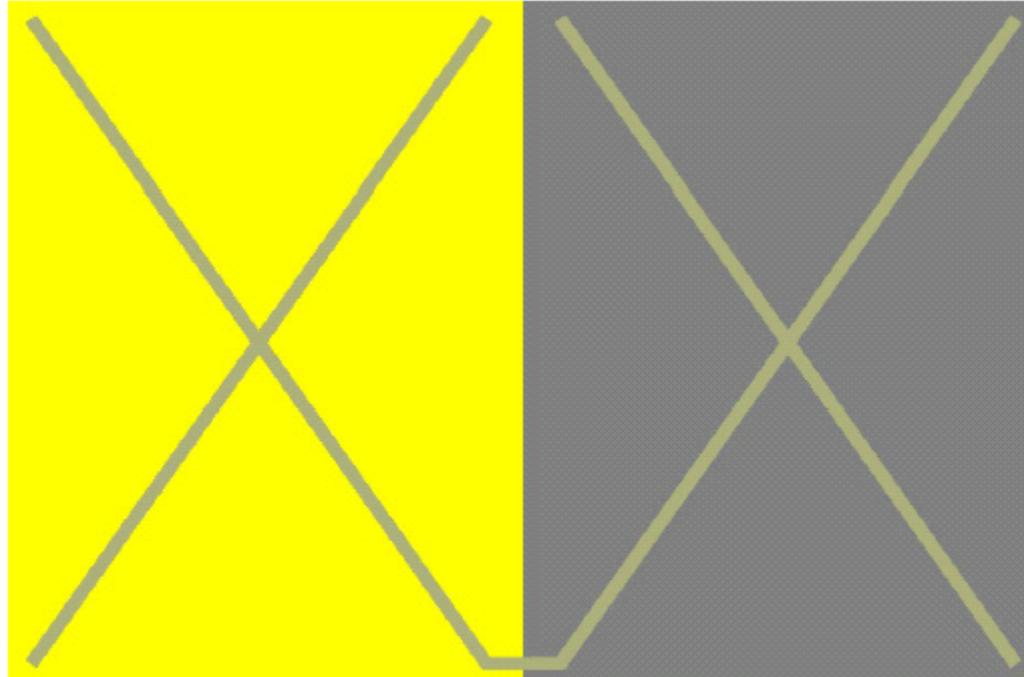


Bild: Joseph Albers

Die **Farbwahrnehmung** ist ein physiologisches Phänomen.
(Farben existieren nur im Gehirn, nicht in der Natur.)

Farbwahrnehmung

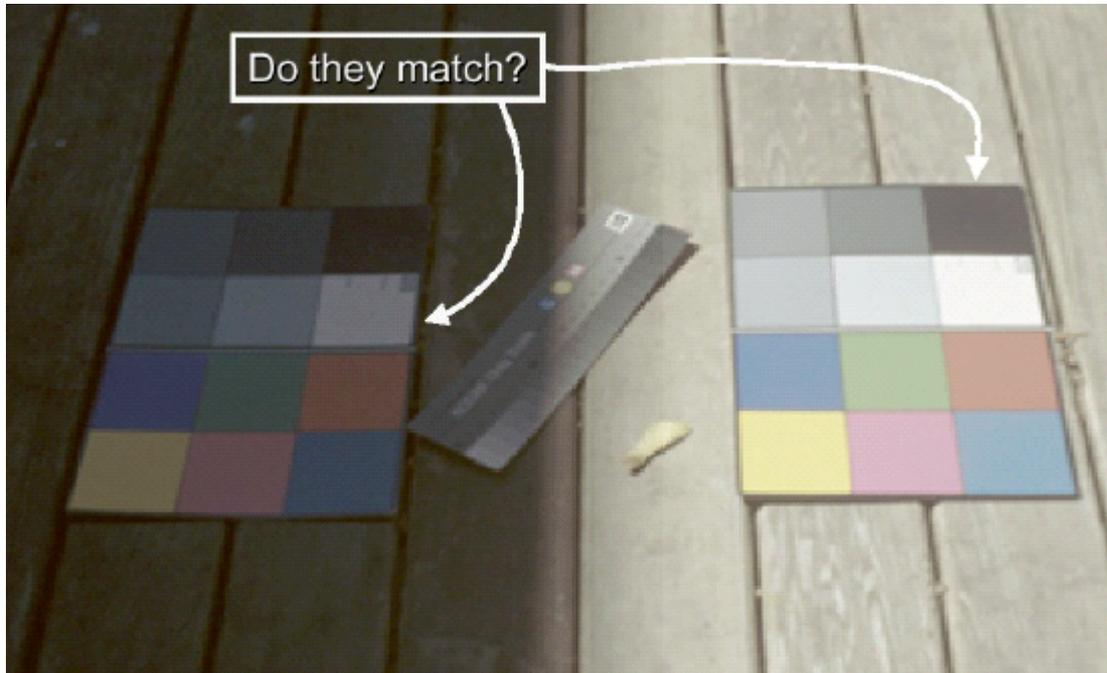
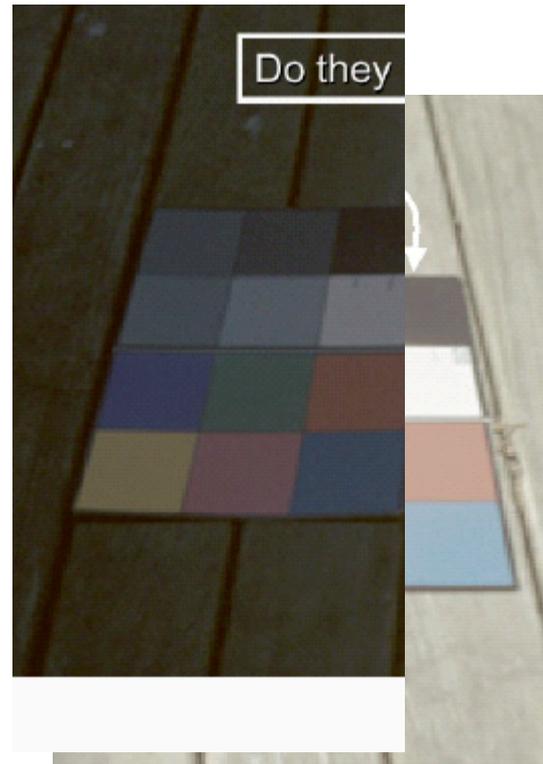


Bild: John M. Cann

Farbwahrnehmung



Cann

Farbwahrnehmung

Farbwahrnehmung entsteht durch (unterschiedliche) Erregung der Zapfen.

„Reine Farben“:

Monochromatisches Licht (gleiche Farbwahrnehmung aber auch durch Mischfarben möglich).

„Unbunte“ Farben:

In dem Maße, wie die Zapfen gleichermaßen erregt sind, entsteht ein Farbeindruck auf der Skala schwarz-grau-weiß (abhängig vom Helligkeitseindruck).

Physikalisch: „weiß“ als Mischung aller Spektralfarben.

Farbwahrnehmung

„Bunte“ Farben:

Bei ungleicher Erregung entsteht ein Farbeindruck, der physiologisch beschrieben wird durch:

- **Helligkeit** (Value, Brightness)
- **Farbton** (Hue)
- **Sättigung** (Saturation)

Subjektive Farbmerkmale

Helligkeit (Value, Brightness):

Maß für die Stärke der wahrgenommenen Intensität.

Farbton (Hue):

Begriff zur Unterscheidung verschiedener charakteristischer Spektralmuster; unterscheidet zwischen reinen Farben (rot, gelb, grün, blau, usw.).

Sättigung (Saturation):

Maß für den Grad, in dem der Farbton eines gegebenen Lichtes von dem Farbton eines weißen Lichtes gleicher Intensität abweicht. Gesättigte Farbe enthält maximal 2 Grundfarben: Hinzufügen der dritten Farbe bewirkt (zusammen mit den anderen beiden Farben) einen **Unbunt-Anteil: schwarz bzw. grau bzw. weiß.**

Menschliches Farbempfinden

Gegründet auf sensorischen Fähigkeiten des Auges:
Kombination der Reizstärke der drei Zapfen-Sorten

Verarbeitung im Gehirn unter Einfluss von Erfahrung/Kultur
Unterschiedlicher Namensvorrat in verschiedenen Sprachen

Verbindung von Farben mit

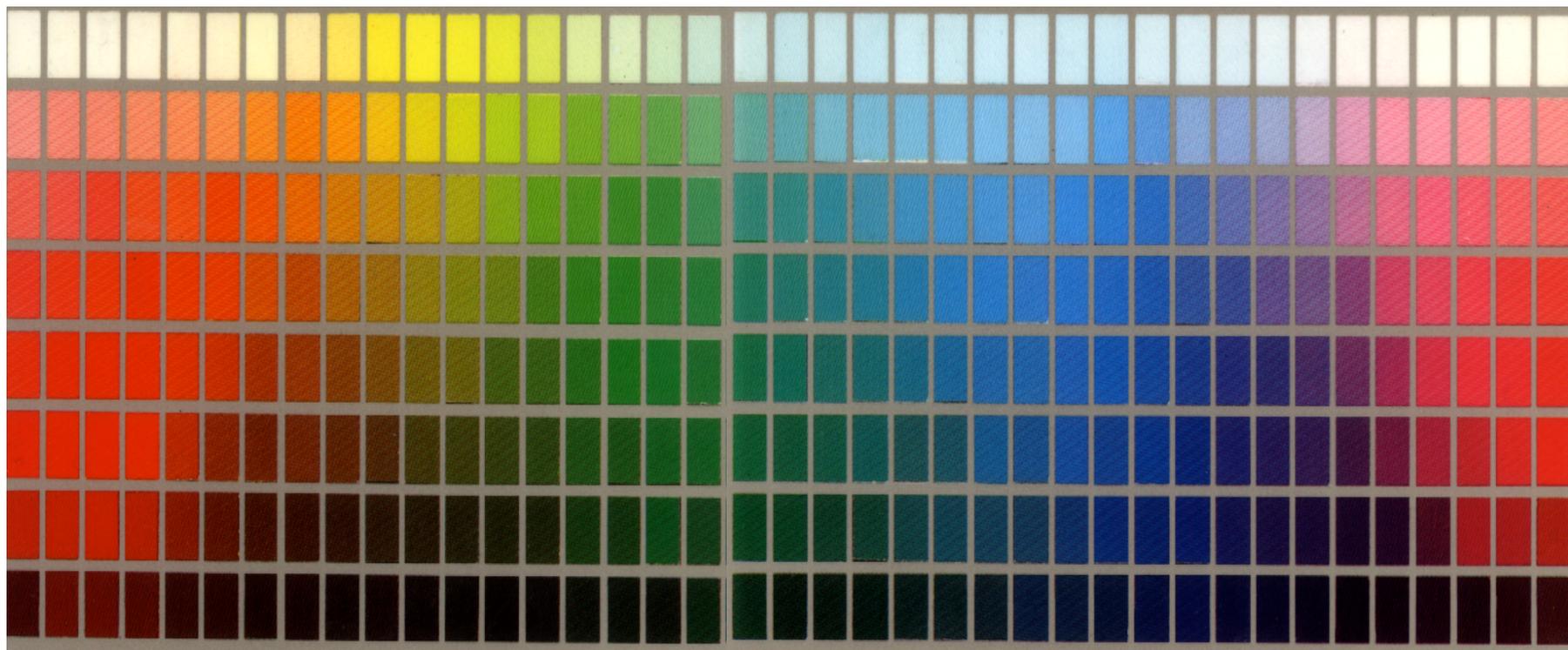
- Stimmungen
- Situationen
- Politik
- Signalfarben
- ...

Menschliches Farbempfinden

Unterschiedlicher Namensvorrat in verschiedenen Sprachen

- Inuits Farben für Schnee (urban legend?)
- Unterschied Männer - Frauen?
- Farbkategorien

Diskrete Wortkategorien für Farben



Munsell chips

	5R	10R	5YR	10YR	5Y	10Y	5GY	10GY	5G	10G	5BG	10BG	5B	10B	5PB	10PB	5P	10P	5RP	10RP
9	Pink	Pink	Pink	Yellow	Yellow															
8	•	Pink	Orange	•	Yellow															
7	Pink		•	Yellow	Yellow														Pink	Pink
6	Pink	Orange	Orange	Yellow																
5	Red	Red	Orange	Brown				•					Blue	Blue	Blue	Purple	Purple	Purple	Pink	Pink
4	•	Red	Brown	Brown			Green						•	Blue	Blue	Blue	Purple	Purple	Purple	Red
3	Red	Red	Brown	Brown													•	Purple	Purple	Red
2	Red	Brown	Brown	Brown																Red

English

	5R	10R	5YR	10YR	5Y	10Y	5GY	10GY	5G	10G	5BG	10BG	5B	10B	5PB	10PB	5P	10P	5RP	10RP
9	3	2	5	Wap	2				1	1			Wap	1	5	12	6	3		2
8				9	6	2	3										1			
7			2	5	4	4	1	1		2										
6		Mehi		2	1			2	3		1									
5	6	2			Wor	1		6	7	4	Nol	2	2							3
4	19								5			3								11
3	2			Kel							1		1		1					1
2		1	1	3	4	6	12				2	1	4	3	4	4	Kel	2		

Dani

(Davidoff, Roberson, et.al. 1999)

Wahrnehmbare Farben

Metamerie: Unterschiedliches Licht kann gleiche Wahrnehmung hervorrufen

Anzahl wahrnehmbarer Farben

Der menschliche Sehapparat kennt näherungsweise

- 128 verschiedene Farbtöne (hues)
- 130 verschiedene Farbsättigungen (Farbreinheit)
- 16 (im gelben Bereich) – 26 (im blauen Bereich) verschiedene Helligkeitswerte

Insgesamt ca. 380 000 verschiedene Farben

- Sichere Unterscheidung gleichzeitig dargestellter Farben in Experimenten nur bei ca. 15 Farben

Objektive Farbbeurteilung

Physiologische Ausgangssituation:

- Drei Rezeptortypen
- Licht unterschiedlicher Eigenschaften erzeugt Farbwahrnehmung
- Die (chemischen) Prozesse der Reizaufnahme und die (kognitiven) Prozesse der Verarbeitung sind subjektiv

Hilfsmittel: Farbtabellen

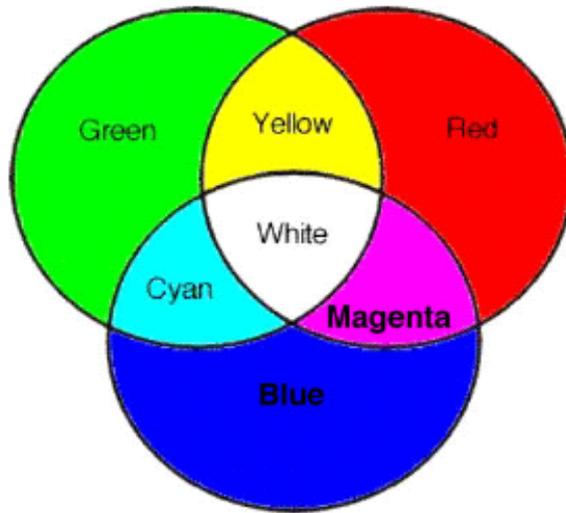
Nachmischung von Farben zielt auf gleiche Farbempfindung

Objektive Farbbeurteilung

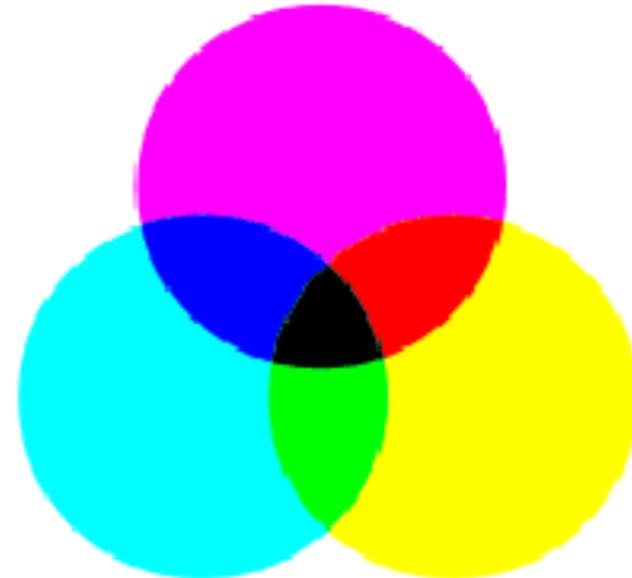
Physikalische Ausgangssituation

- Farbmischung:
 - Additive Mischung (Strahlungsquellen)
 - Subtraktive Mischung (Absorption, Filter)
- Farbmodelle auf der Basis von drei Rezeptoren
- Farbräume:
 - Technische Basis für Farbverarbeitung
 - Druck, Foto, TV, ...
- Festlegung von Standards:
CIE (Commission Internationale de l'Éclairage)
 - Internationale Organisation mit Sitz in Wien
 - <http://www.cie.co.at>

Farbmischung



Additive Mischung
(Strahlungsquellen)



Subtraktive Mischung
(Absorption, Filter)

Farbmodelle

Ziel: Eindeutige Beschreibung von Farben

Allgemeine Farbmodelle:

- CIE-Farbraum, CIE-L*a*b

Hardwarebezogene Farbmodelle:

- RGB, CMY, CMYK, YUV, YIQ

Physiologisch orientierte Farbmodelle:

- HLS, HSV, HSI

Ein Farbmodell muss nicht alle wahrnehmbaren Farben enthalten.

Die Kombination von Farben soll dem Mischen von Farben entsprechen.

Modellierung

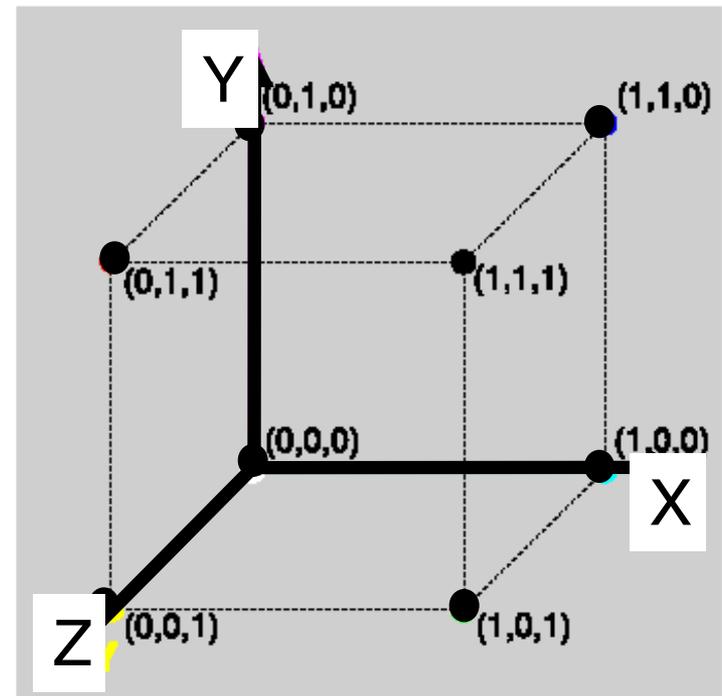
Wegen der Eigenschaften des Auges reichen 3 Parameter:
3-dimensionaler Farbraum mit Koordinaten X, Y, Z .

Beim Mischen kommen nur Farben hinzu:
Kombinationen mit positiven Koeffizienten.

Untergrenzen und Obergrenzen:

Normierung

z.B. auf Intervall $[0, 1]$



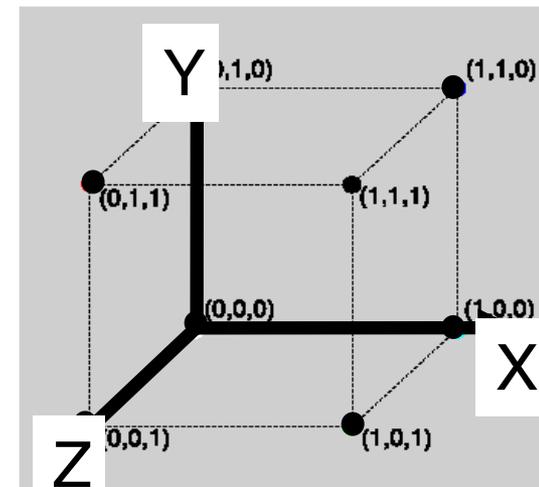
Modellierung aller Farben

Ausgangssituation entsprechend Farbwahrnehmung.

Die Rezeptorempfindlichkeiten überschneiden sich:

nicht der ganze Würfel ist besetzt.

Alle Farben aus Spektralfarben kombinierbar mit positiven Komponenten: Farben liegen in der konvexen Hülle der Spektralfarben.



Modellierung aller Farben

Beschreibung der Reizaufnahme (näherungsweise)
für drei Sensoren X,Y,Z mit gegebenen Empfindlichkeiten:

$$e_x = \int \phi_x(\lambda) I(\lambda) d\lambda$$

$$e_y = \int \phi_y(\lambda) I(\lambda) d\lambda$$

$$e_z = \int \phi_z(\lambda) I(\lambda) d\lambda$$

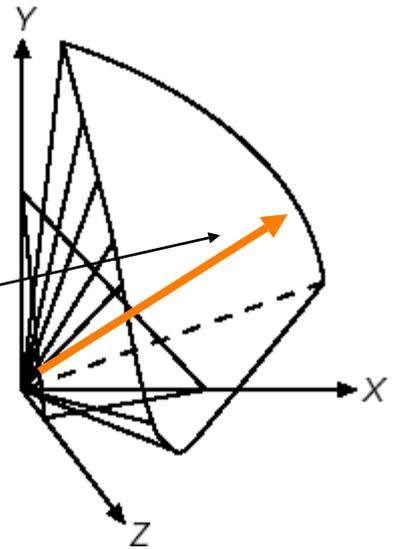
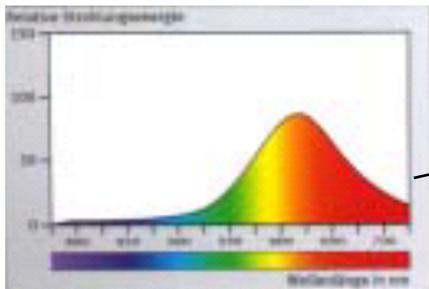
Erregung e jeweils abhängig von Intensität $I(\lambda)$ des Lichts
und Empfindlichkeit $\phi(\lambda)$ des Sensors für die
unterschiedlichen Frequenzen λ .

Modellierung aller Farben

$$e_x = \int \phi_x(\lambda) I(\lambda) d\lambda$$

$$e_y = \int \phi_y(\lambda) I(\lambda) d\lambda$$

$$e_z = \int \phi_z(\lambda) I(\lambda) d\lambda$$



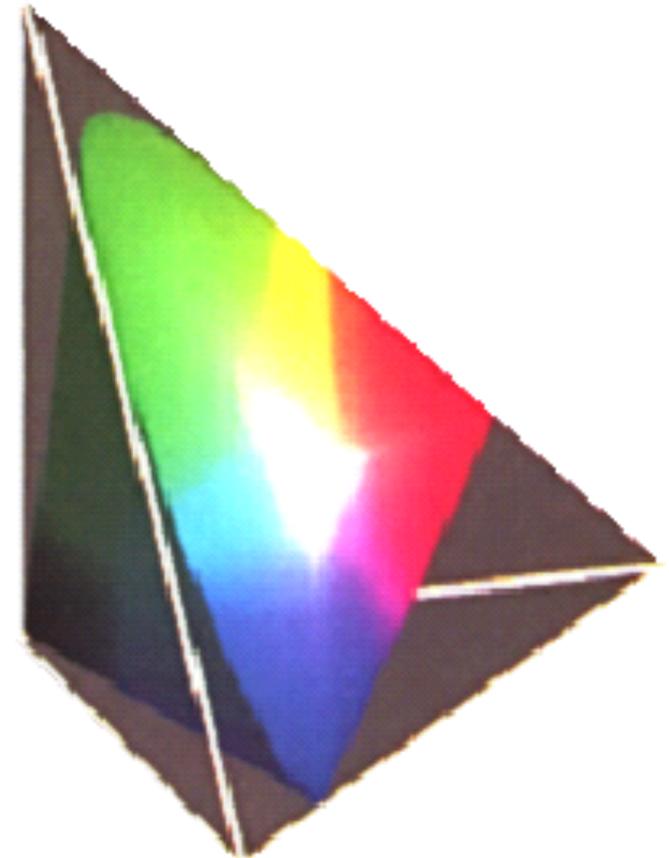
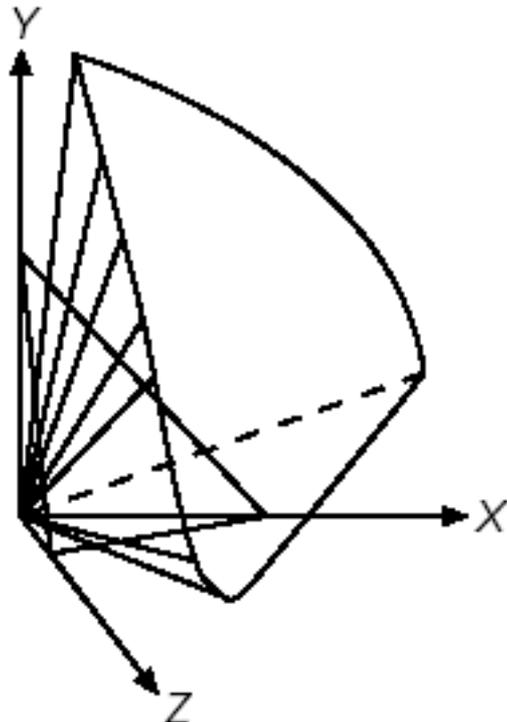
Bei gegebenen Φ_x , Φ_y , Φ_z kann für jede Intensitätsverteilung $I(\lambda)$ der darstellende Vektor (e_x, e_y, e_z) im Farbraum berechnet werden.

Modellierung aller Farben

Die Spektralfarben liegen am rechten und linken Rand.

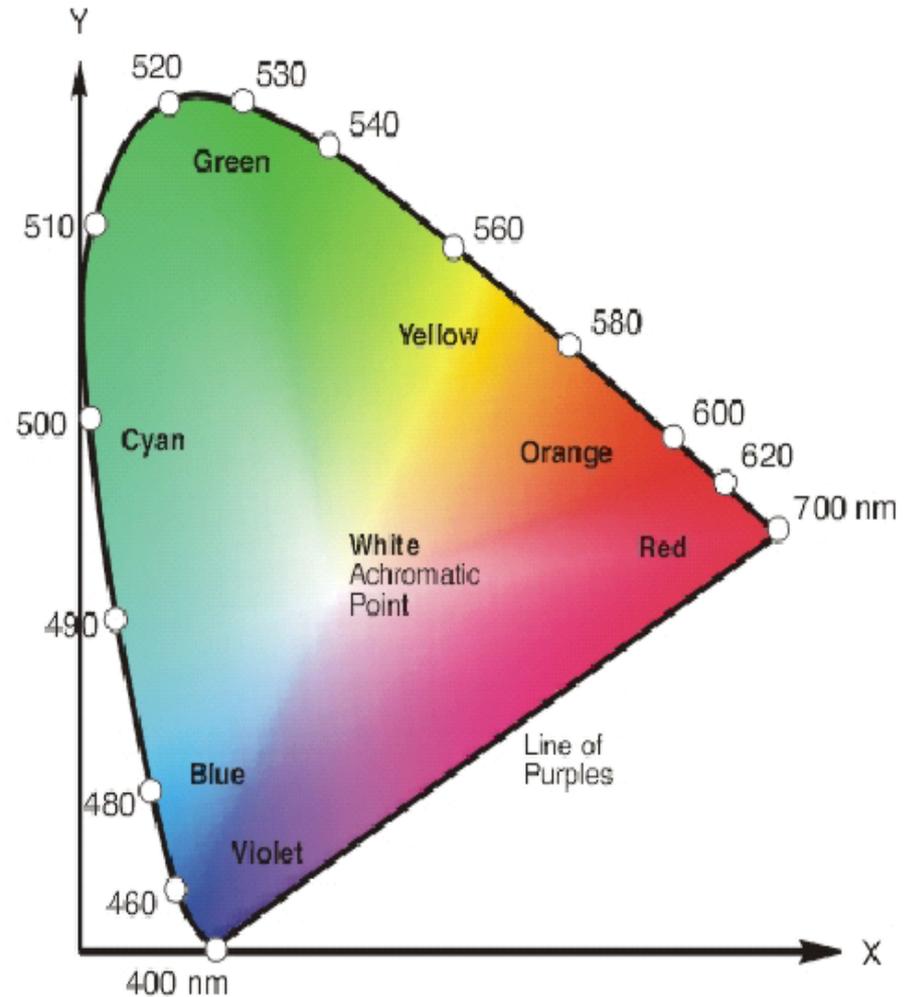
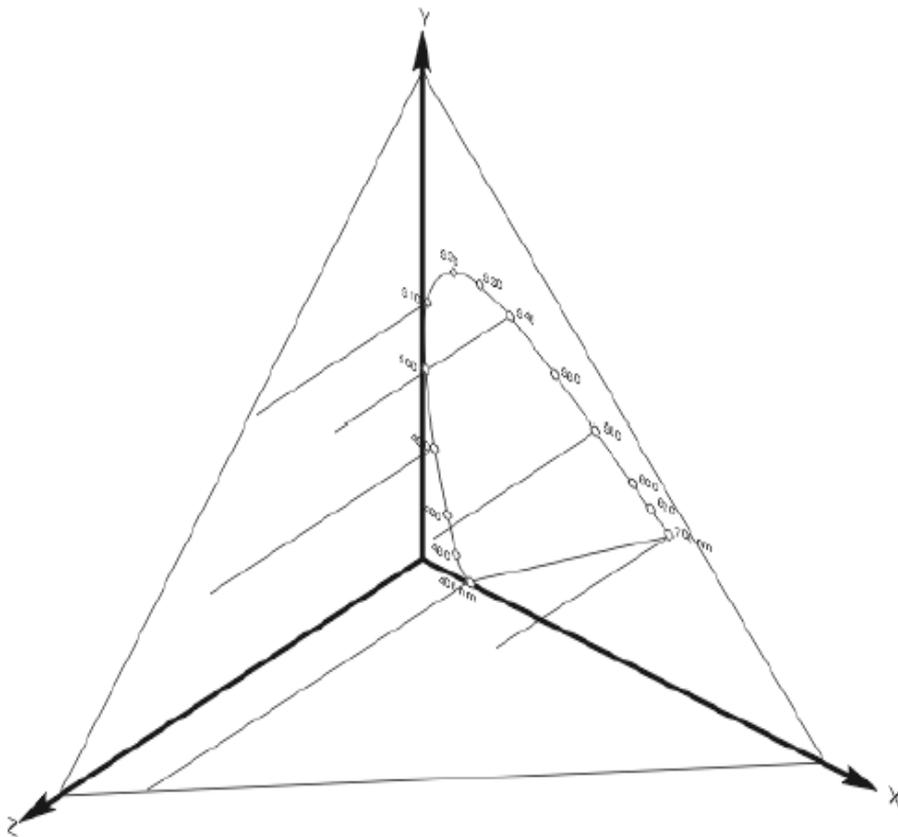
Die anderen Farben innerhalb der konvexen Hülle.

Untere Begrenzung: „Purpurlinie“



Farbtafel (*chromaticity diagram*)

Verzicht auf Darstellung der Helligkeit
durch Beschränkung auf die Ebene $X+Y+Z = 1$



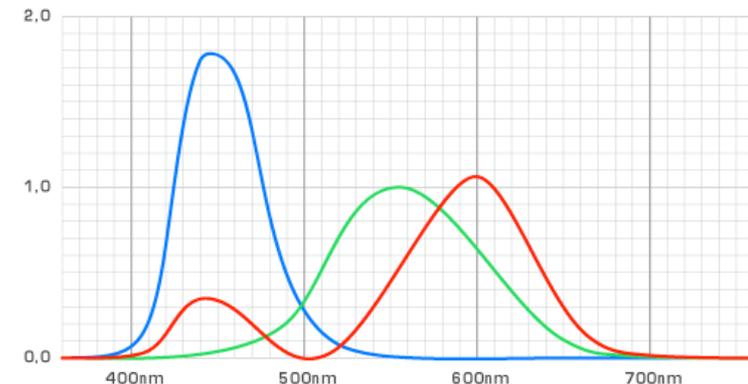
CIE-Standard (1931, 1976)

Festlegung von X,Y,Z („Standard-Beobachter“):

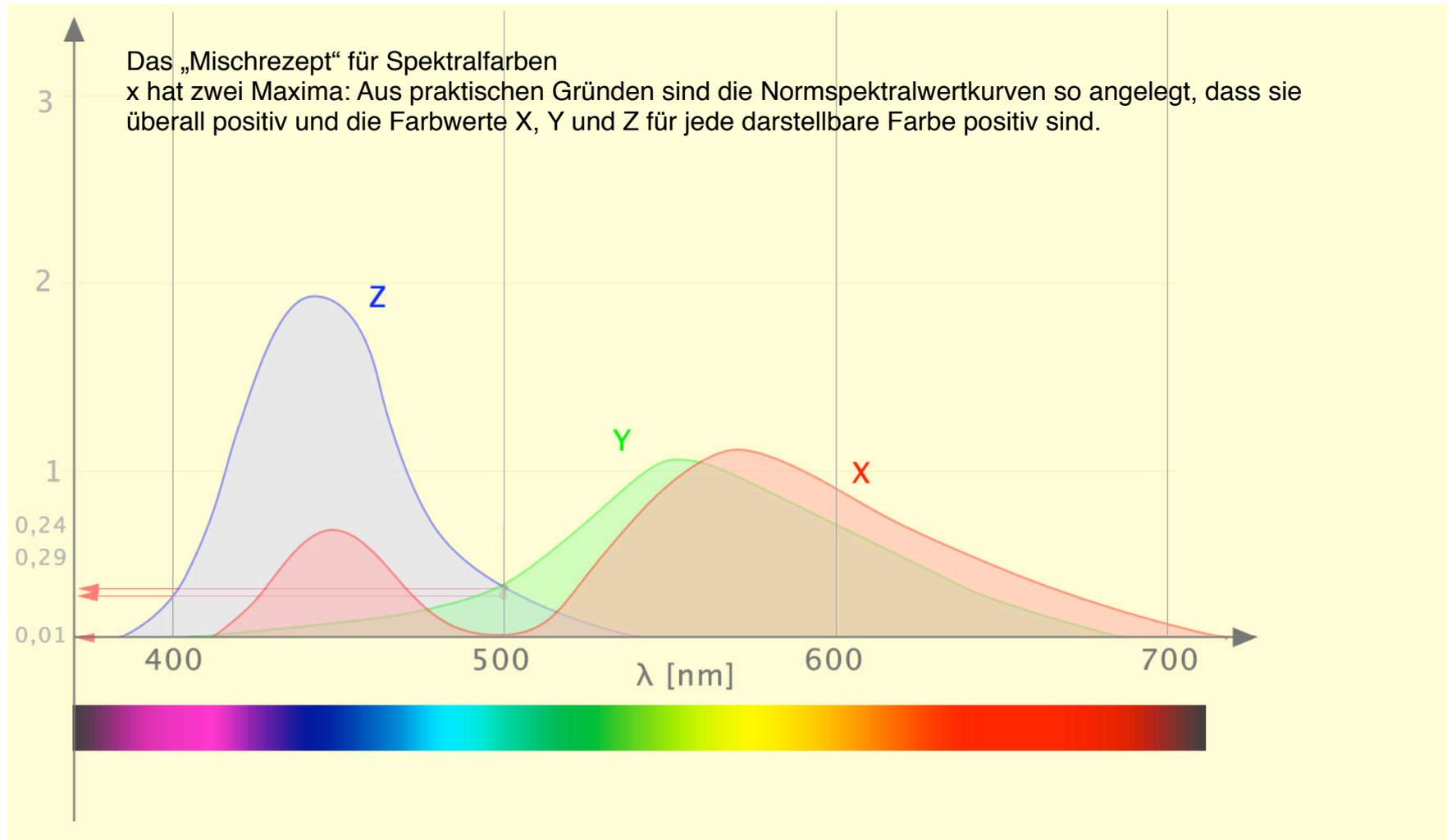
Normspektralfunktionen x,y,z in Anlehnung an RGB-System.

Koordinatentransformation zur besseren Normierung:

- Empfindlichkeitskurven entsprechen nicht dem natürlichen Sehen
- x und y orthogonal
- y entspricht Helligkeitssehen (Zapfen)
- in x-Richtung ist Helligkeit konstant
- Helligkeit nicht durch Summierung
- Weißpunkt bei $(1/3, 1/3)$
- x hat 2 Maxima (analog rot/violett)



Norm-Beobachter



RGB-Modell

R=700 nm

G=546,1 nm

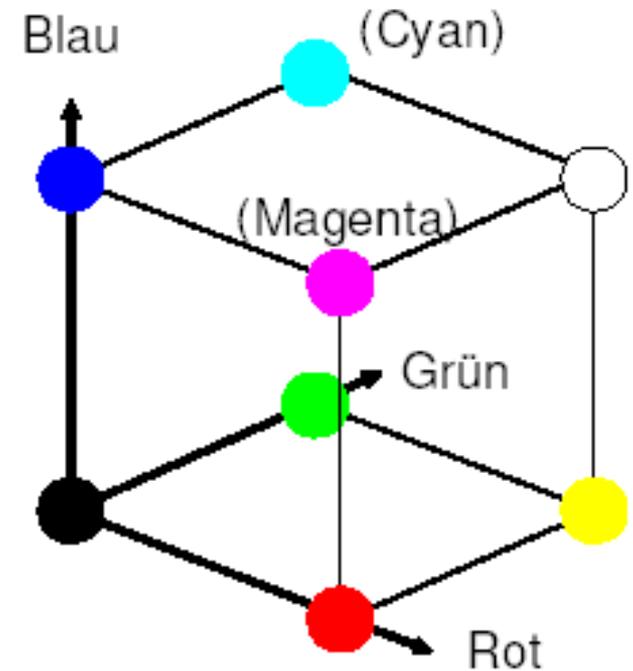
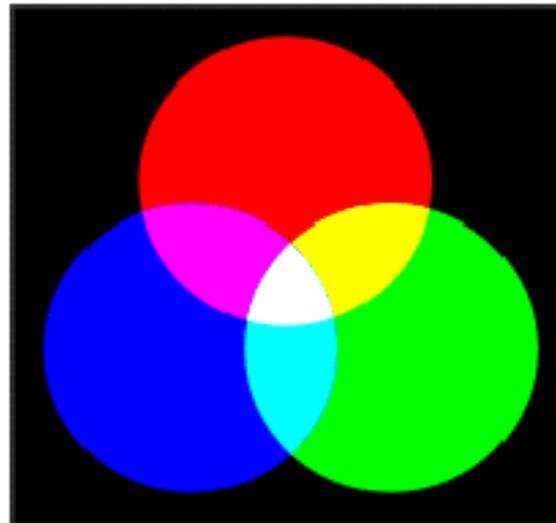
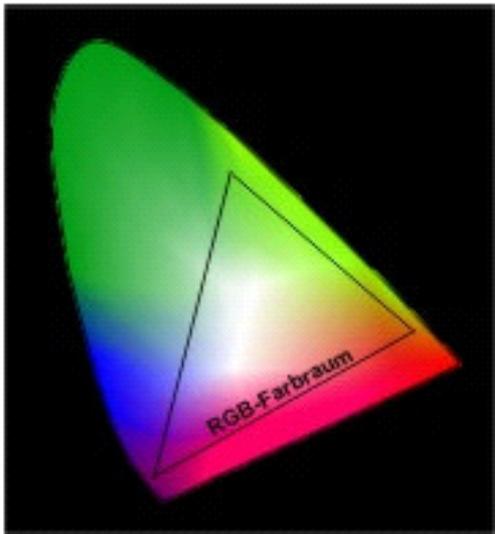
B=435,8 nm

Additives Modell

Meistverwendetes Modell für aktiv lichterzeugende
Ausgabemedien (z.B. Displays)

Spektrale Intensitäten der Komponenten werden addiert

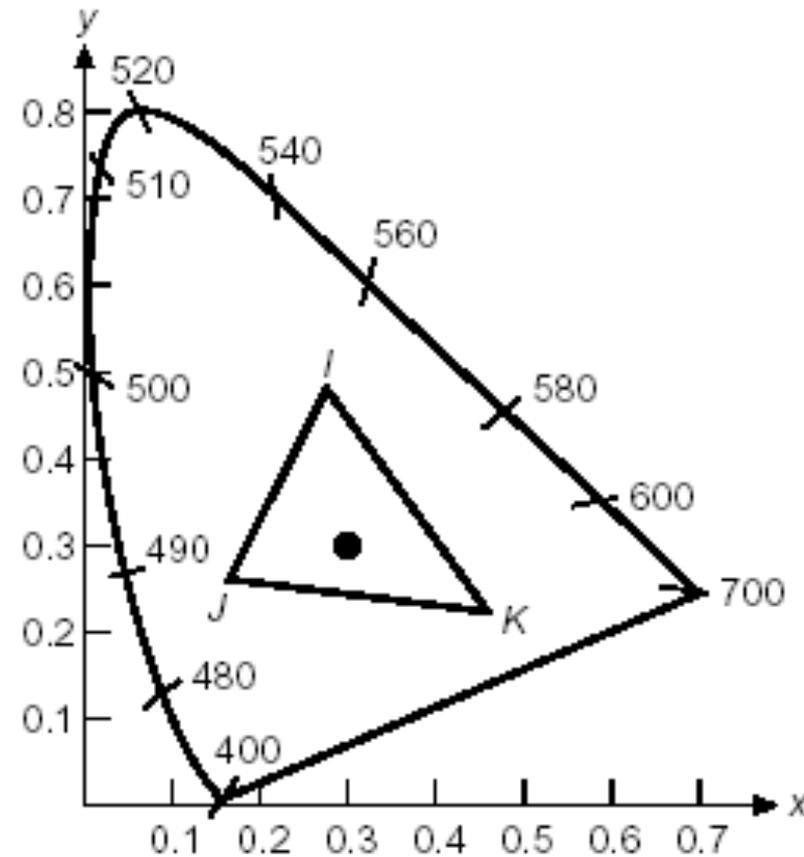
Bestimmte Farben nicht im RGB-Modell darstellbar.

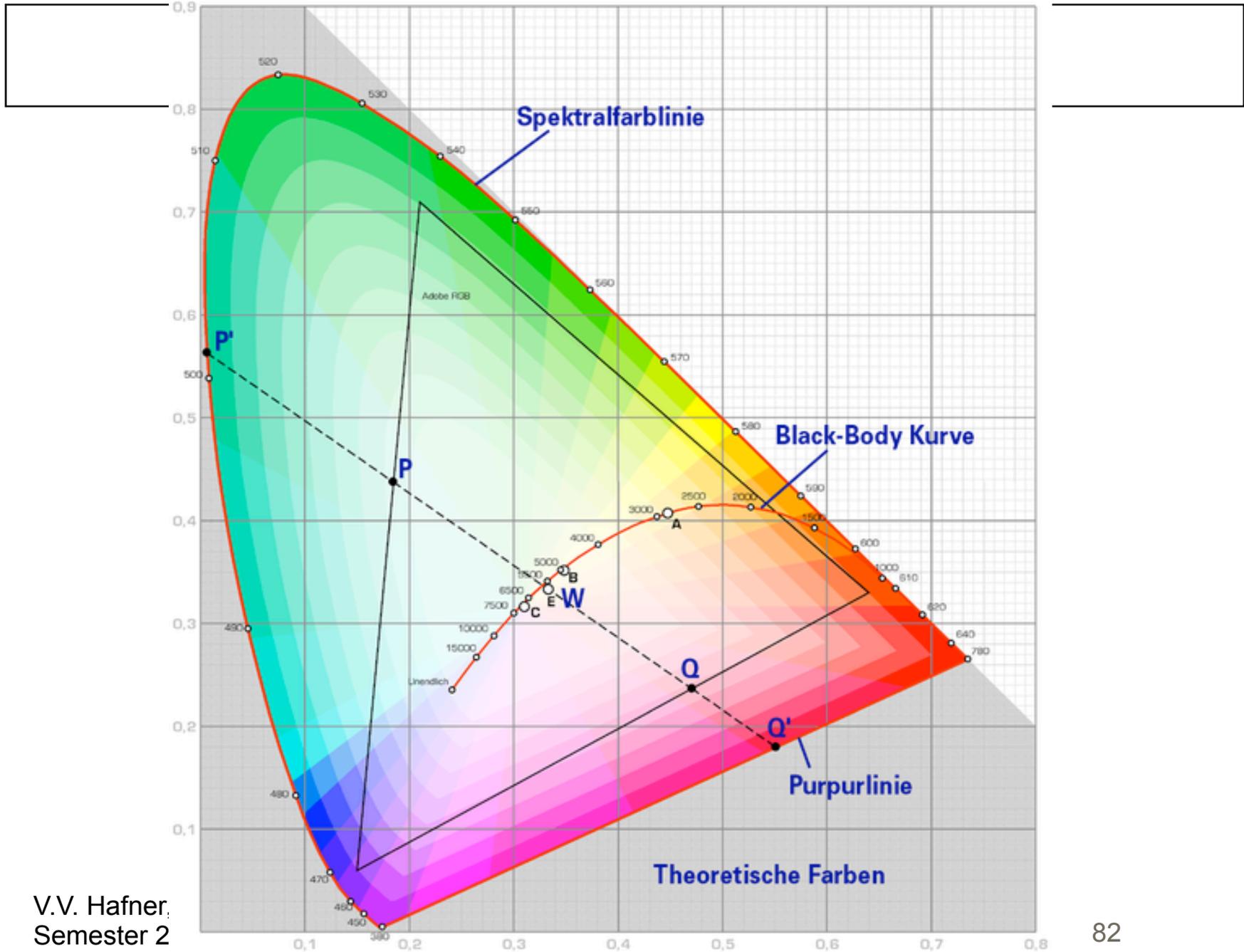


Farbtafel: Additive Farbmischung

Linearer Raum:

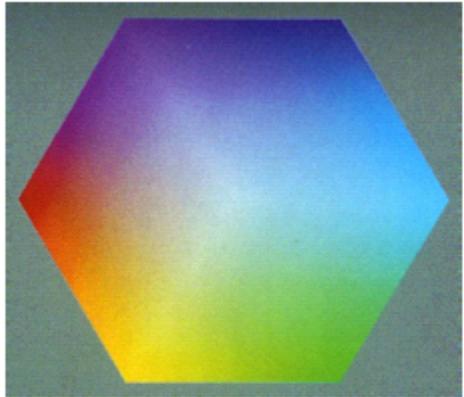
- Jede Farbe auf der Linie IJ kann aus I und J gemischt werden.
- Insbesondere können alle Farben aus zwei Spektralfarben gemischt werden.
- Alle Farben innerhalb des Dreiecks können aus I, J, K gemischt werden.
- Daraus folgt eine Einschränkung: Bei additiver Farbmischung können nicht alle Farben aus 3 Farben erzeugt werden.



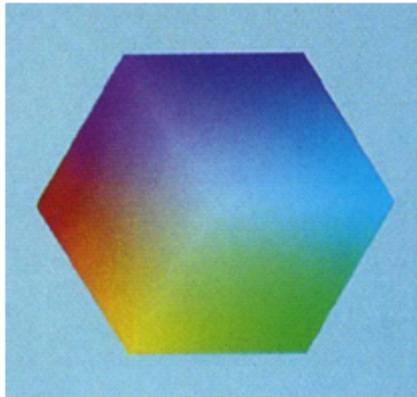


RGB-Modell

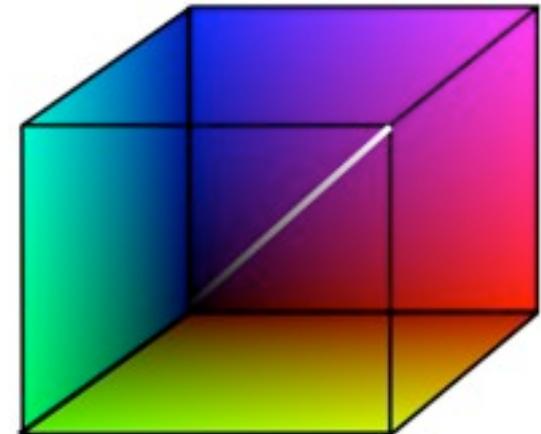
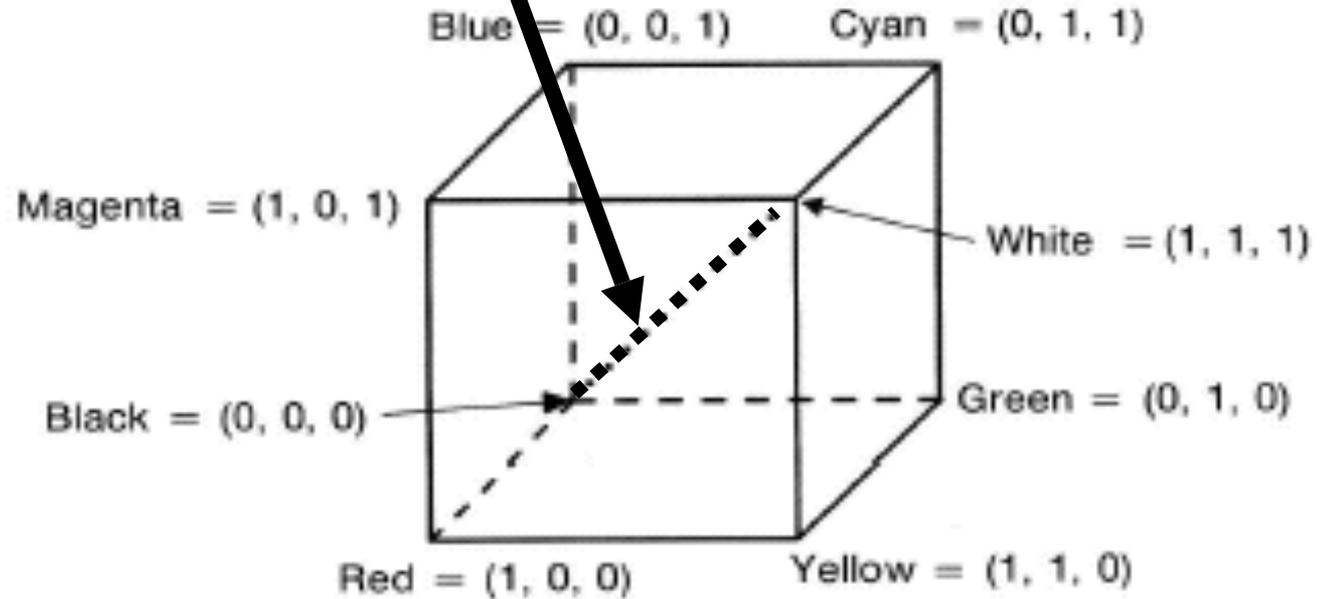
Unbunt-Linie: Schwarz-grau-weiß



Draufsicht

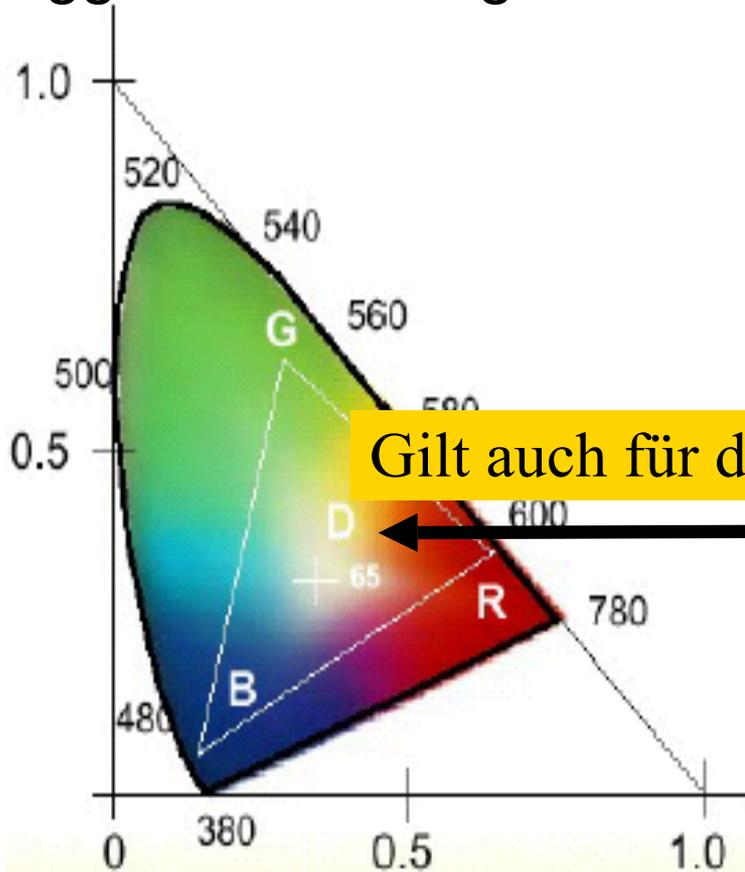


Schnitt



RGB-Modell

RGB-Darstellungen in der Praxis unterschiedlich
ggf. Kalibrierung



Gilt auch für diese Darstellung L

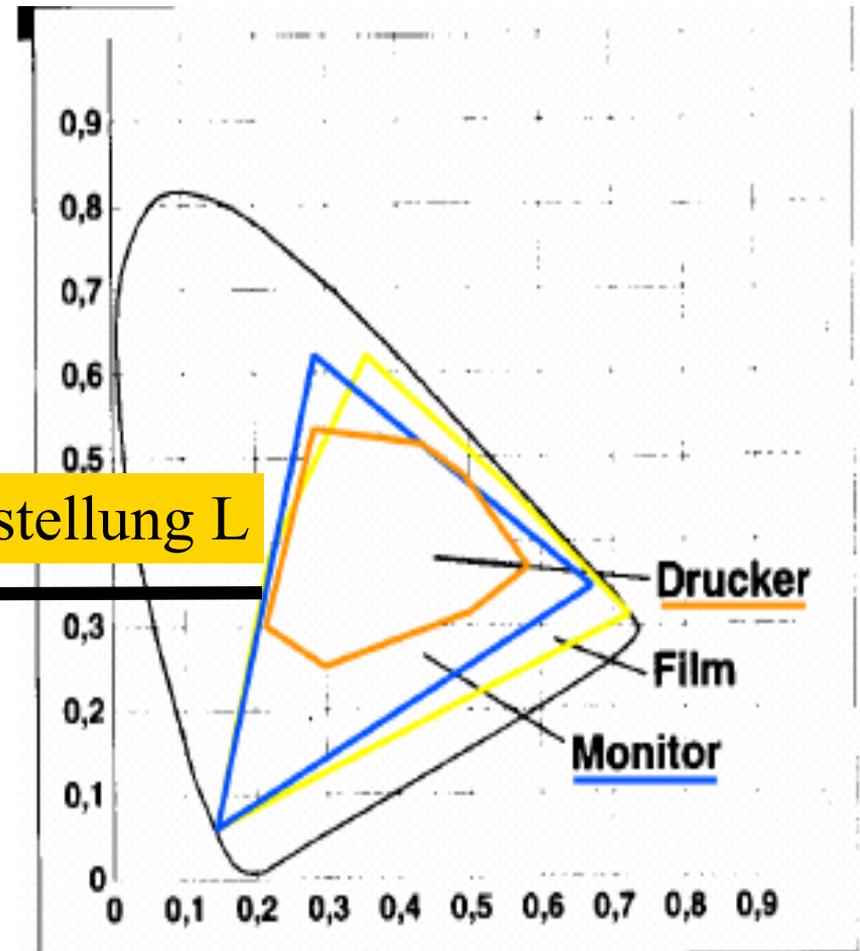


Bild mit RGB-Teilen



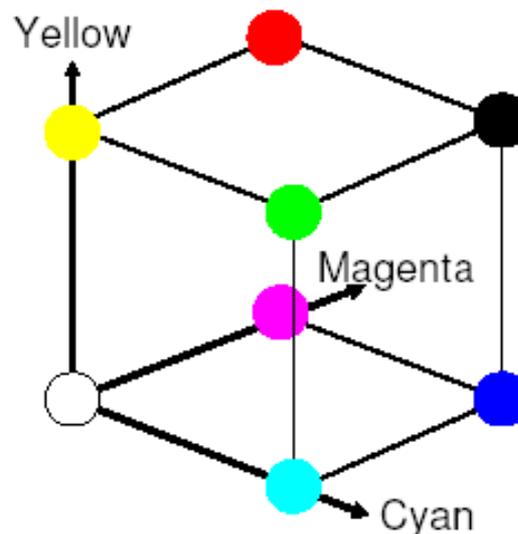
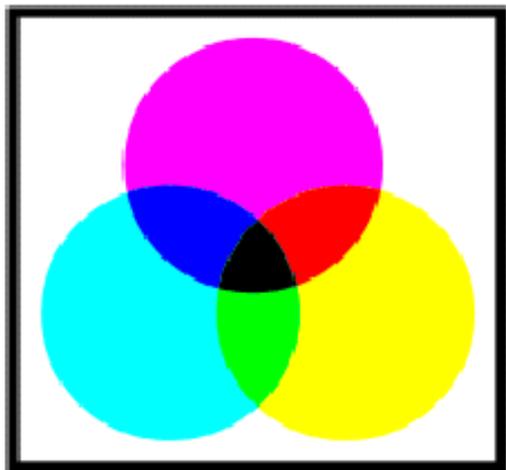
CMY(K)

Subtraktives Modell

Meistverwendetes Modell zur Ausgabe auf reflektierenden Ausgabemedien (z.B. Farbdrucker)

Anschaulich: Farbfilter subtrahieren Farbwerte

Für Drucker oft vierte Komponente "schwarz" (black), deshalb CMYK (schwarz aus CMY nur unvollständig und mit viel Farbauftrag erzeugbar).



CMY(K)

Konvertierung RGB \leftrightarrow CMY

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} C \\ M \\ Y \end{pmatrix}$$

$$\begin{pmatrix} C \\ M \\ Y \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

- gilt nur theoretisch
- kompliziertere Verfahren in der Druckindustrie

HSB-System

An subjektiver Wahrnehmung orientiertes Modell:

H = Hue (Farbton)

S = Saturation (Sättigung)

B = Brightness (Helligkeit)

Im ursprünglichen
drei-dimensionalen XYZ-Modell:

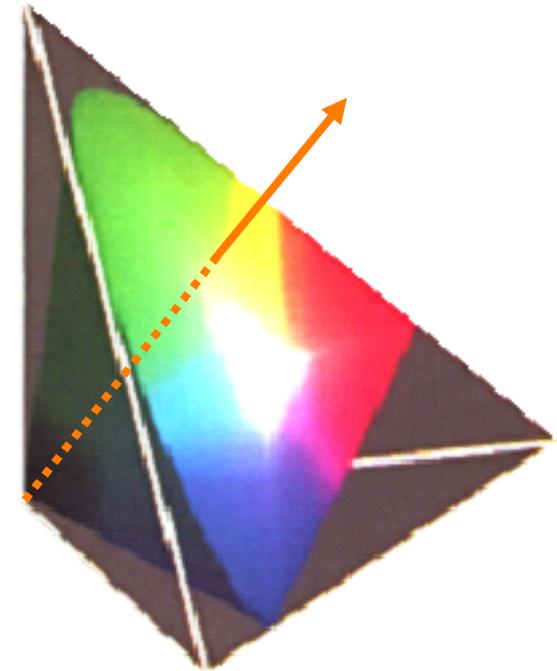
B = Länge des Farbvektors

H = Winkel ausgehend vom Weißpunkt

S = Entfernung vom Weißpunkt

(Vom Prinzip ähnlich:

HSV mit V=„Value“, HSI mit I = Intensität)



YUV Modell

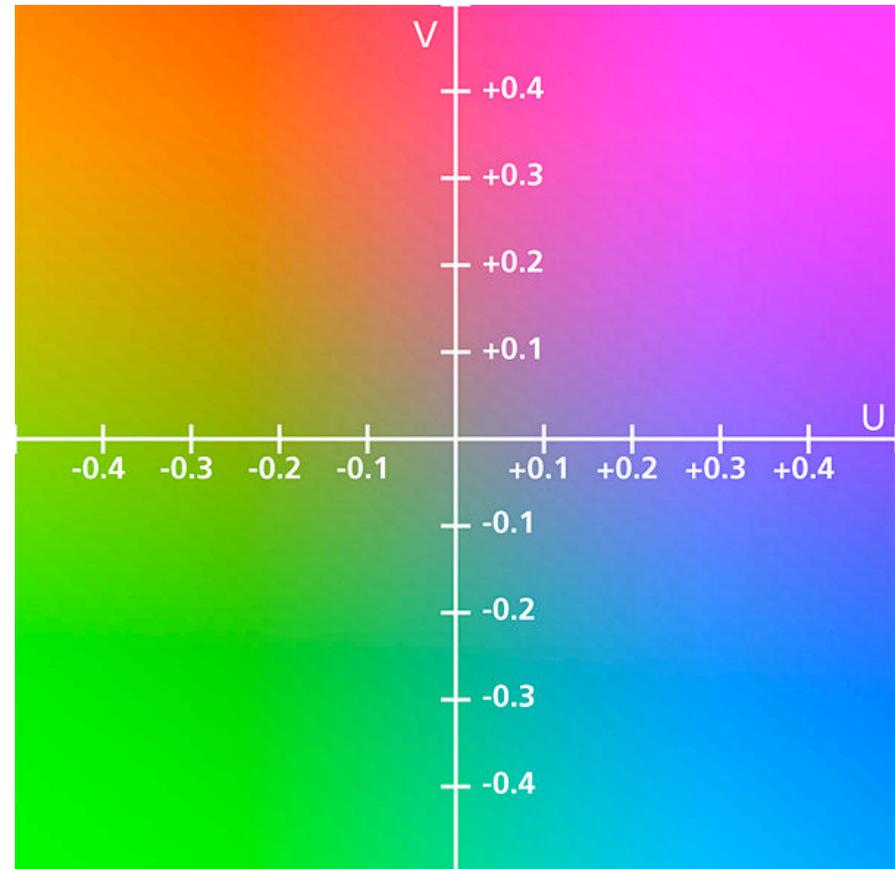
vom Schwarzweiss- zum Farbfernsehen



YUV Modell

Das YUV-Farbmodell verwendet zur Darstellung der Farbinformation zwei Komponenten, die **Luminanz** (Lichtstärke pro Fläche) (luma) und die **Chrominanz** oder Farbanteil (chroma), wobei die Chrominanz wiederum aus zwei Komponenten U und V besteht.

Die Entwicklung des YUV-Farbmodells geht auf die Entwicklung des analogen **Farbfernsehens** in der Norm PAL zurück. Es findet heutzutage auch beim analogen NTSC-Farbfernsehen Verwendung.



YUV

Abgeleitet aus RGB. Bildinformation getrennt in:

Helligkeitssignal

$$Y := 0,59 G + 0,30 R + 0,11 B$$

Wichtung entspricht Helligkeitsempfinden (grün am stärksten)

Zwei Chrominanz-Signale:

$$U := 0,493 (B-Y) \quad (\text{Differenz Blau} - Y)$$

Farbanteil zwischen Gelb und Blau

$$V := 0,877 (R-Y) \quad (\text{Differenz Rot} - Y)$$

Farbanteil zwischen Cyan und Rot

Konvertierungen RGB \Leftrightarrow YUV

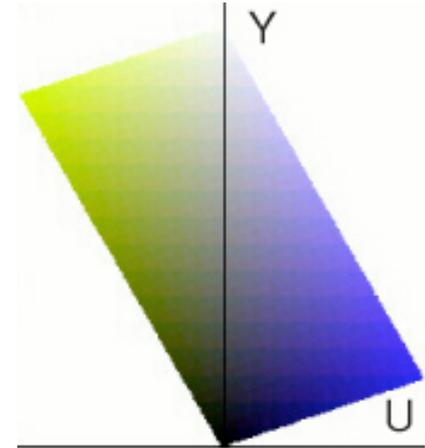
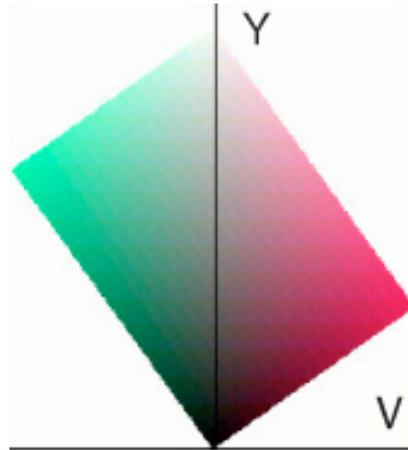
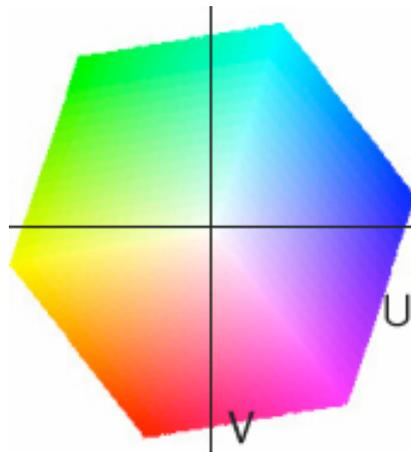
Transformations-Matrizen

(leicht unterschiedliche Werte in der Literatur,
dabei auch U und V manchmal vertauscht)

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.437 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

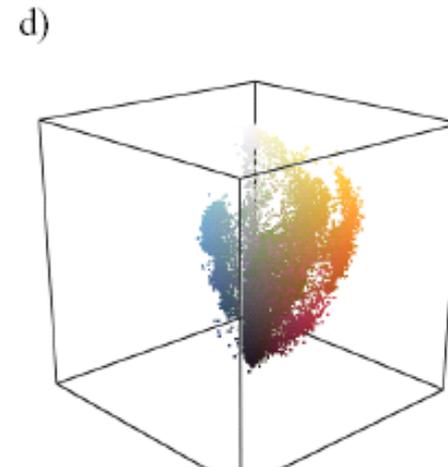
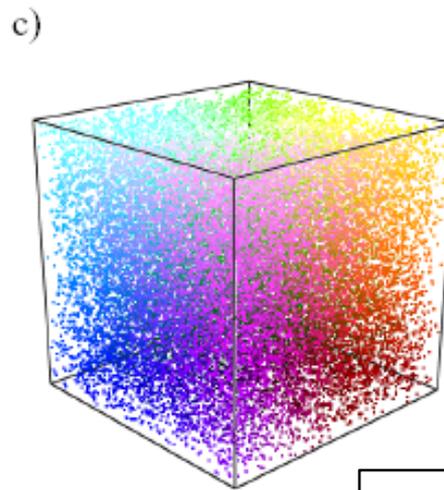
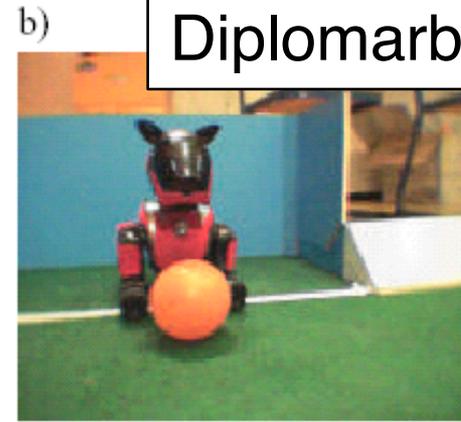
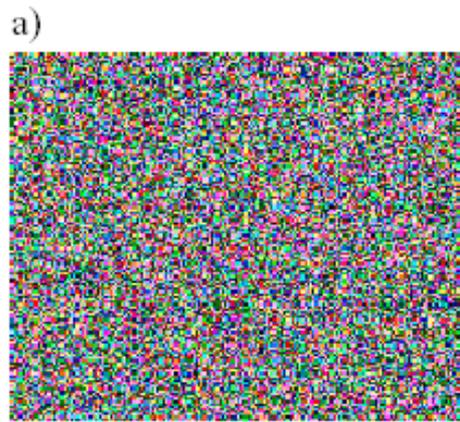
$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1.140 \\ 1 & -0.394 & -0.581 \\ 1 & 2.028 & 0 \end{bmatrix} \begin{bmatrix} Y \\ U \\ V \end{bmatrix}$$

YUV



YUV

AIBO benutzt YUV



Vertikal: Y-Achse (Helligkeit)

(c) YUV-Anteile des Zufallsbildes (a)

(d) YUV-Anteile des Bildes (b) – nur Teil des Raums benutzt

Bilderzeugung

Erfassen räumlicher Gegebenheiten:
Objekte erzeugen/reflektieren Licht.

Bilderzeugung durch systematische Erfassung von
Lichtstrahlung.

Abbildungsvorgang

Lichtstrahlen aus verschiedenen Quellen im 3D-Raum:

Lichtstrahlen treffen auf Abbildungsfläche

- Im Auge: Netzhaut
- In der Kamera: Film, Sensor-Feld
 - *Modellierung als Projektion 3D nach 2D
(abhängig von Kamera-Konstruktion)*

Lichtstrahlen werden umgewandelt:

- Nervensignale im Auge/Gehirn
- Chemische Prozesse auf Film
- Digitales Bild

Elektronische Bildaufnahme, Kamera

Lichtempfindliche Sensoren messen Lichtstärke

z.B. CCD = Charge Coupled Device:

Ladung proportional zur Lichtstärke

Kamera: Rasterförmige (Zeile, Feld) Anordnung von Sensor-Zellen zur Aufnahme von Bildern.

Nach Aufnahme (Verschluss schließen): Übertragung aus Zellen in einen Speicher zur weiteren Verarbeitung.

Das gespeicherte Bild (frame) ist die Matrix der von jeder Zelle empfangenen Werte.

Die Bildpunkte heißen Pixel (= picture element).

Erzeugung Digitaler Bilder

1. Räumliche Abtastung:

- Jedes Sensorelement erhält Licht aus bestimmten Bereich gemäß Kamera-System:
 - Linsen
 - Anordnung der Zellen: Hardwaremäßig vorgegebene Rasterung (rechteckig, quadratisch, sechseckig)

2. Zeitliche Abtastung:

- Sensorelement erhält Licht während einer bestimmten Zeit
 - Framerate: Anzahl der Bilder pro Sekunde
 - Sequentielle Abtastung kann zu Verzerrungen führen

Quantisierung

3. Quantisierung (Digitalisierung):

Analog-Digital-Wandler („Frame Grabber“)
transformiert Intensitätswerte
z.B. in 8-bit oder 16-bit-Werte

- Grauwerte bei „Schwarz-Weiß“-Kanal
- Farbwerte in Farbkanälen

Ergebnis ist eine Matrix von Pixelwerten.

Quantisierung mit

128, 64,

32, 16,

8, 4

Grauwerten

V.V. Hafner, HU Berlin Winter-
Semester 2008/09

Vorlesung Kog
Visuelle Sensoren



Haberäcker, S.13

Weitere Softwaretechnische Aufbereitung

Verbesserung des Bildes

- Interpolation, Rauschunterdrückung usw.

Kompressionsalgorithmen

- Speicherersparnis
- z.T. mit Kompressionsverlusten

Digitale Bilder (1-Kanal-Bilder)

Felder von „Pixeln“ („picture elements“),
jeweils x-y-Position, „Grauwert“

Darstellung als Matrix $[g(x,y)]_{x,y}$ von Grauwerten mit

- Anordnung: (1,1) ist oben links
- x nach unten, y nach rechts
- Wertebereich z.B. 8 bit, 16 bit
- 8-bit-Darstellung:
 - 0 für Schwarz
 - 255 für Weiß
- Binärbild: 1 bit (0=Schwarz)

1,1	1,2	1,3	1,4	1,5
2,1	2,2	2,3	2,4	2,5
3,1	3,2	3,3	3,4	3,5
4,1	4,2	4,3	4,4	4,5

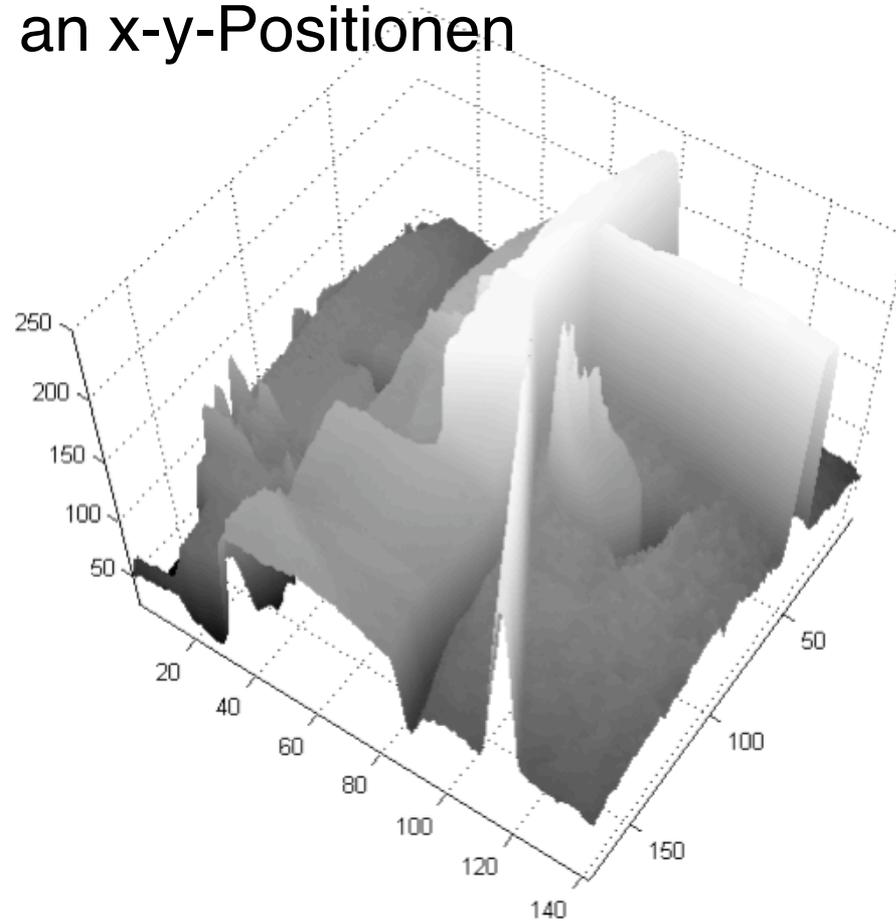
Digitale Bilder (1-Kanal-Bilder)

Felder von „Pixeln“ („picture elements“),
jeweils x-y-Position, „Grauwert“

110	120	124	121	120	122	122	117	122	121	120	121	122	121	123	121	123	123	123	125	121	117	110	100	100	1
118	120	122	122	120	125	122	121	120	121	120	121	122	121	122	121	123	123	123	125	121	115	114	115	114	1
120	121	119	121	119	122	120	120	119	121	121	121	121	121	123	123	123	123	123	125	121	115	114	115	114	1
120	119	121	122	120	119	118	121	121	122	121	120	122	120	122	123	120	123	123	122	117	112	110	115	119	1
117	119	122	121	117	118	119	119	121	121	122	120	120	120	120	122	124	124	123	122	118	111	106	116	121	1
117	121	119	118	119	120	120	120	122	119	120	122	120	122	120	118	121	123	121	122	119	112	104	116	123	1
117	121	119	120	120	120	123	122	122	120	123	122	121	122	121	122	126	121	121	121	121	112	104	111	120	1
118	118	122	120	122	125	124	123	124	125	122	122	121	121	122	122	126	123	123	121	112	104	110	122	1	
120	121	124	124	126	130	131	128	127	129	126	125	124	121	120	123	122	122	121	118	111	108	111	123	1	
122	125	129	129	127	129	131	131	129	135	133	128	128	125	126	123	122	122	119	113	107	108	113	119	1	
125	127	130	130	128	129	129	129	131	136	138	134	132	133	131	126	121	122	117	111	103	100	108	109	1	
129	130	130	132	128	128	129	128	129	131	130	133	134	131	128	124	118	114	108	102	89	88	90	90	1	
124	130	130	130	130	129	128	127	126	122	121	119	119	118	116	112	109	102	96	89	77	75	77	77	1	
120	127	128	125	124	121	119	115	116	111	109	102	98	94	96	92	89	87	84	80	71	62	62	66	1	
117	116	123	117	107	103	102	97	97	93	91	86	80	76	76	75	71	73	72	73	66	53	51	55	1	
94	96	100	100	91	85	83	79	76	71	70	71	66	63	62	62	61	60	61	60	63	56	51	50	1	
74	76	79	78	75	71	68	63	59	62	57	62	57	55	53	55	56	54	54	57	59	57	53	53	1	
66	64	64	65	67	64	57	53	52	52	56	56	53	50	52	49	53	52	52	54	56	51	53	54	1	
65	59	55	55	60	58	51	48	47	51	53	54	48	48	50	52	54	54	53	55	50	51	56	1		
64	58	53	55	56	54	48	48	48	52	50	49	46	48	50	52	55	54	51	54	51	53	54	1		
63	61	56	56	58	53	50	48	46	46	48	47	48	48	51	49	51	53	54	57	55	51	54	58	1	
63	61	59	58	60	58	54	47	47	43	44	46	48	50	51	51	48	49	52	53	53	53	58	1		
63	61	59	56	58	60	52	44	47	45	45	48	53	53	50	52	51	49	49	52	51	52	53	54	1	
62	61	62	60	57	56	53	44	47	46	47	48	50	52	49	53	51	51	51	52	53	54	52	55	1	
61	59	65	62	57	53	53	48	50	49	45	48	47	47	48	52	52	51	51	52	54	53	55	56	1	
57	59	62	60	58	55	54	52	50	53	44	48	45	45	50	51	51	51	51	52	52	55	56	54	1	
57	58	62	60	57	56	56	52	49	52	48	47	47	45	49	49	49	51	50	52	51	55	53	51	1	
60	59	62	58	56	57	56	51	51	52	51	48	49	48	50	48	51	51	49	51	54	53	52	50	1	
59	60	61	60	61	60	60	54	53	55	51	49	49	50	50	49	49	53	52	49	50	52	52	51	1	
56	58	57	61	61	60	58	55	53	53	54	52	48	51	50	48	51	54	51	49	49	51	53	52	1	
59	59	59	56	56	57	56	56	55	56	55	48	47	50	52	49	51	50	48	52	50	51	50	1		
63	60	62	60	58	54	57	58	56	56	58	51	50	48	52	48	50	48	49	48	50	51	51	51	1	
64	59	61	63	60	55	56	57	59	55	59	52	53	49	49	48	48	48	50	51	52	51	52	55	1	
61	62	64	60	59	59	57	57	60	54	54	49	54	51	51	51	49	49	51	53	53	52	52	53	1	
58	61	61	58	56	60	59	57	57	55	51	50	53	53	52	53	52	48	51	51	51	52	54	51	1	
59	61	59	59	56	60	60	60	55	58	54	51	52	54	50	49	52	50	49	54	55	51	48	1		
59	60	60	60	58	60	60	60	54	57	53	53	53	53	51	48	52	51	53	52	55	52	50	49	1	
59	61	61	61	59	57	56	56	54	55	53	55	52	53	56	52	53	53	50	53	54	48	50	51	1	
59	59	59	61	59	58	53	56	55	56	57	56	53	55	55	52	54	53	48	53	53	53	52	52	1	
63	59	60	58	57	57	58	55	53	54	58	58	58	57	53	57	55	52	50	53	54	59	56	52	1	
62	60	56	58	58	56	59	56	57	54	53	53	55	55	55	57	55	53	53	53	54	55	55	54	1	
59	59	57	58	62	58	58	56	58	57	56	55	56	54	55	52	55	57	54	55	57	55	54	53	1	
62	58	57	59	61	59	55	54	54	56	60	59	55	51	52	54	52	56	54	55	55	58	53	51	1	
62	58	58	58	58	58	56	53	56	57	53	52	52	53	53	55	55	55	55	55	55	55	55	53	1	

Digitale Bilder (1-Kanal-Bilder)

Felder von „Pixeln“ („picture elements“)
(Grau-)Werte an x-y-Positionen



Digitale Bilder (1-Kanal-Bilder)

Daten-Komprimierung: Originalbild rekonstruierbar

- RLE (run length encoding, Lauflängenkodierung)
für Binär-Bilder nur Anfangspunkt und Länge von
(horizontalen) schwarzen Linien angeben.
- Weitere Möglichkeiten: Baumstrukturen, Richtungsketten.
Oft angepasst an spezielle Formen der Weiterverarbeitung.

Daten-Reduktion: Originalbild nicht vollständig rekonstruierbar

- Reduktion der Grauwertskala
- Informationstheoretische Verfahren
- Physiologische Unterscheidungsfähigkeit ausnutzen

Histogramme

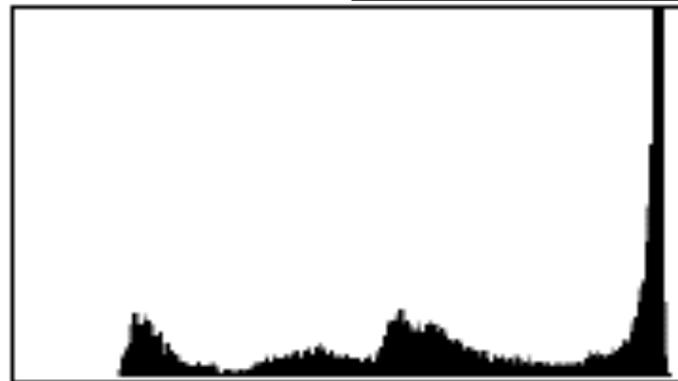
Relative Häufigkeit von Grauwerten $a \in [0 \dots 255]$

$\text{Hist}(a) := A/M$

$A :=$ Anzahl Pixel (x,y) mit $g(x,y)=a$

$M :=$ Anzahl aller Pixel

Diplomarbeit Jünger



Häufigkeiten für Grauwerte als Abszissen.
Maximum bei hohem Weißanteil.

Histogramme für Zeilen/Spalten

Relative Häufigkeit von Grauwerten $a \in [0 \dots 255]$ für Zeilen
bzw. Spalten

Spalten:

$$\text{hist}_{\text{Spalte}_i}(a) := A_{\text{Spalte}_i} / M_{\text{Spalte}_i}$$

$A_{\text{Spalte}_i} :=$ Anzahl Pixel (x,y) mit $g(x,y)=a$ in Spalte i

$M_{\text{Spalte}_i} :=$ Anzahl aller Pixel in Spalte i

Zeilen:

$$\text{hist}_{\text{Zeile}_j}(a) := A_{\text{Zeile}_j} / M_{\text{Zeile}_j}$$

$A_{\text{Zeile}_j} :=$ Anzahl Pixel (x,y) mit $g(x,y)=a$ in Zeile j

$M_{\text{Zeile}_j} :=$ Anzahl aller Pixel in Zeile j

v.v. Se Frage: gilt $\text{Hist}(a) = \sum_i \text{hist}_{\text{Spalte}_i}(a) = \sum_j \text{hist}_{\text{Zeile}_j}(a)$?

Histogramme: Anwendungen

Relative Häufigkeit im Bild z.B. anwendbar für
Quantisierung

Schwelwertverfahren

(Schwellwerte für Klassifizierung bestimmen)

Relative Häufigkeit für Zeilen bzw. Spalten

Position von Bildinhalten

(in Schnittpunkten bzgl. Histogramm)

Ähnlichkeit von Bildern

Analoge Verfahren mit Histogrammen für Farbwerte

Statistische Merkmale von Bildern

Mittlerer Grauwert (Mittelwert):

$$m = (1 / LR) \cdot \sum_{i=1 \dots L} \sum_{j=1 \dots R} g(i,j)$$

Maß für Helligkeit eines Bildes

Mittlere quadratische Ableitung (Varianz):

$$\sigma^2 = (1 / LR) \cdot \sum_{i=1 \dots L} \sum_{j=1 \dots R} (g(i,j) - m)^2$$

Maß für Kontrast eines Bildes

Entropie:

$$H = \sum_{a=0 \dots 255} \text{Hist}(a) \log_2 \text{Hist}(a)$$

Maß für mittlere Unsicherheit des Grauwerts eines Bildpunktes

Maß für mittlere Anzahl von Bits pro Bildpunkt

Mehrkanal-Bilder

Kanäle mit unterschiedlichen Informationen

- Farbbilder (3 Kanäle bei entsprechendem Farbmodell)
- Multispektralbilder (weitere Kanäle, z.B. für erweiterte Farbempfindlichkeit oder andere Informationen)

Bildelement (Pixel):

- x-y- Position, Intensitätswerte je Kanal

Darstellung als Matrix

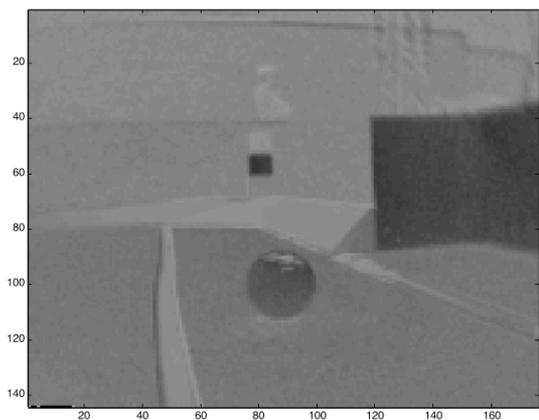
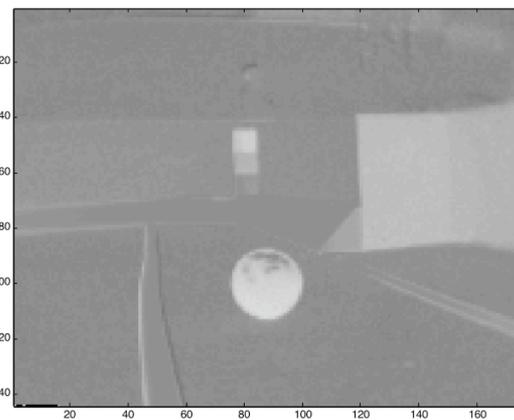
$$[f_1(x,y), f_2(x,y), f_3(x,y)]_{x,y}$$

bzw. je Kanal eine Matrix

$$[f_i(x,y)]_{x,y}$$

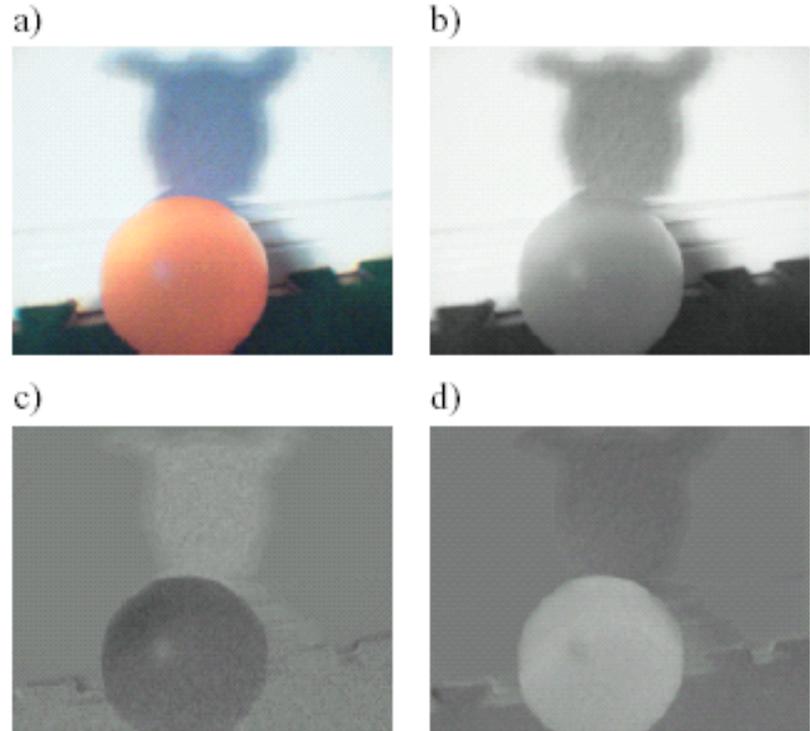
(interpretierbar als je ein „Graubild“ pro Kanal)

Farbkanäle in RGB



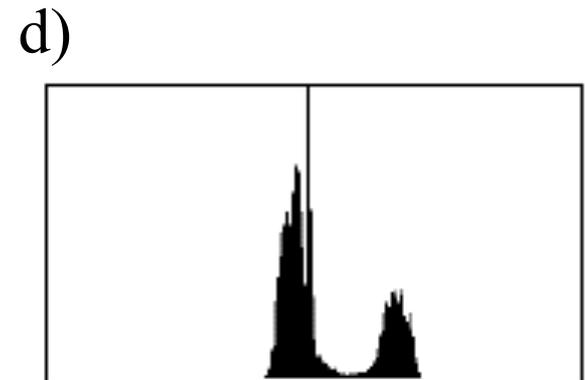
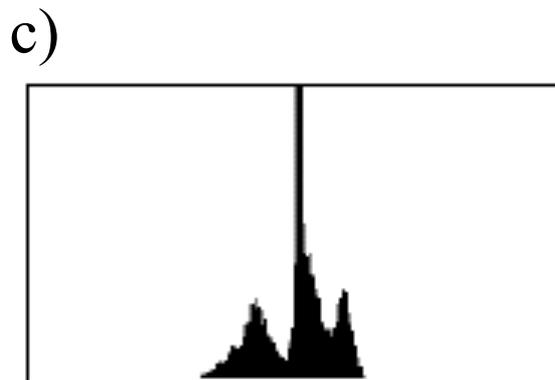
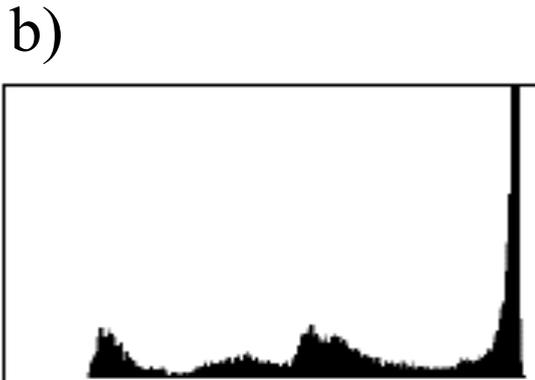
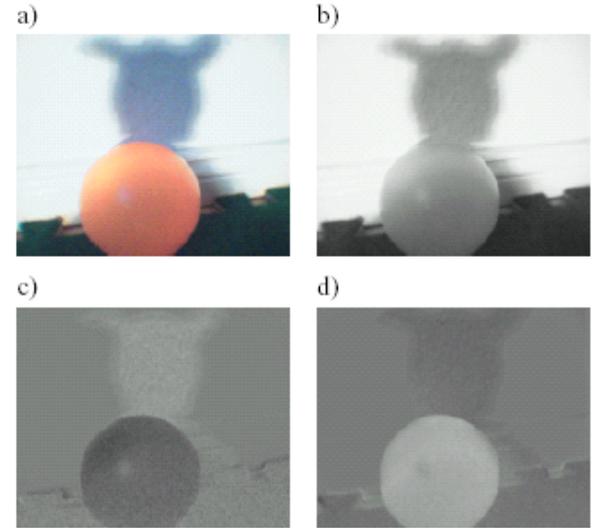
Farbkanäle in YUV

- a) Bild
- b) Y-Kanal (Helligkeit) als Graubild
- c) U-Kanal (Blau) als Graubild
(blauer Schatten: hell,
orange Ball: dunkel)
- d) V-Kanal (Rot) als Graubild
(orange Ball
hat höchste Intensität)



Histogramme der YUV-Kanäle

Maxima entsprechend den hohen Weißanteilen



Häufigkeits-Bild

Original-Pixel ersetzt durch Grauwert-Pixel entsprechend der Häufigkeit seiner Farbe:

Pixel mit selten benutzter Farbe sind dunkel.

Pixel mit oft benutzter Farbe sind hell.

a)



b)

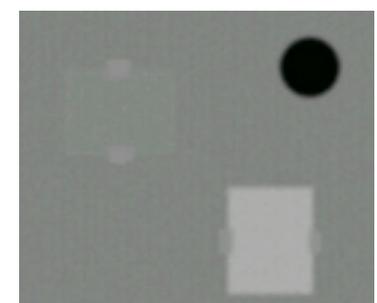
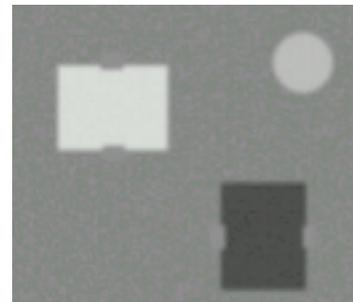
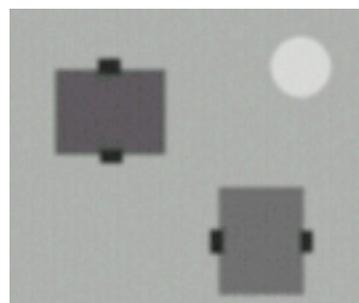
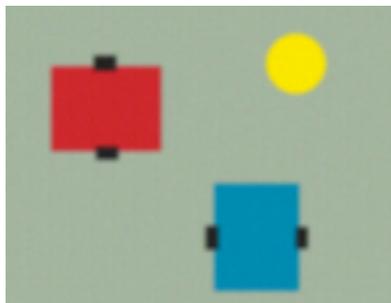


Farbkanäle in YUV

Starke Helligkeit (Y) bzw. starkes Rot bzw. starkes Blau erscheinen hell in der SW-Wiedergabe der Kanäle.

Die einzelnen Kanäle können wie Graubilder behandelt werden (Kompression/Reduktion, Verarbeitung)

Angaben über Nutzen der Kanäle aus Histogramm, Mittelwert, Varianz, Entropie ableitbar



Y (hell)

V (rot)

U (blau)

Bildqualität

Qualität bestimmt durch

- Wertebereiche der Intensitätswerte
(z.B. Anzahl der Grauwert-Töne: 8 bit, 16 bit)

- Rastergröße

bzgl. korrekter Wiedergabe

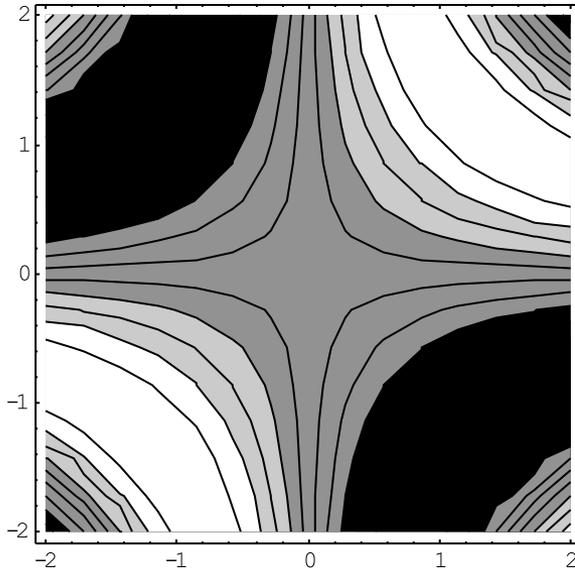
Bei ungenügender Rastergröße entstehen neue Muster: „Alias“

- Die Rasterpunkte wirken als diskrete Abtastpunkte:
Nur an bestimmten Stellen wird gemessen.
- Die Rastergröße muss ausreichen, um insbesondere kleine Strukturen korrekt wiederzugeben.

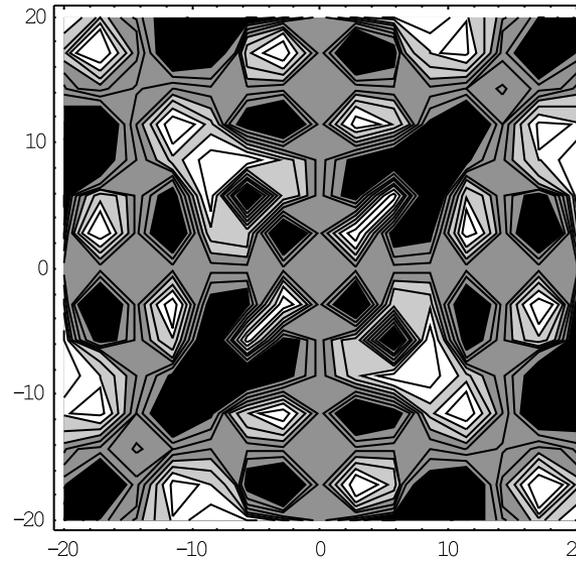
bzgl. menschlicher Wahrnehmung

- Rasterpunkte kleiner als Unterscheidungsfähigkeit
(abhängig von Entfernung des Betrachters)

Darstellung von $\sin(xy)$ in ungenügender Rasterung

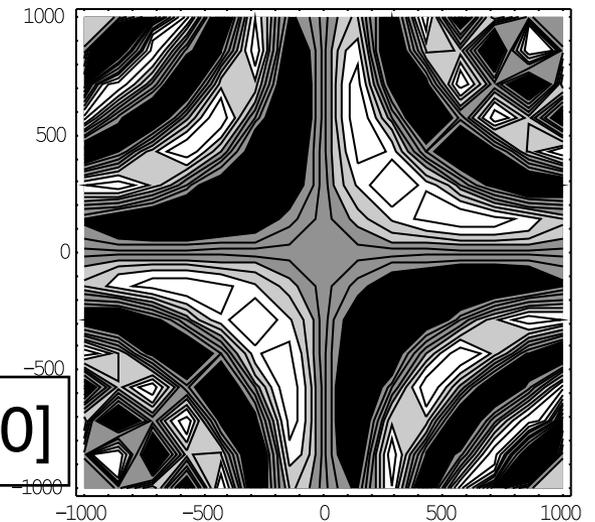


$x, y \in [-2, 2]$



$x, y \in [-20, 20]$

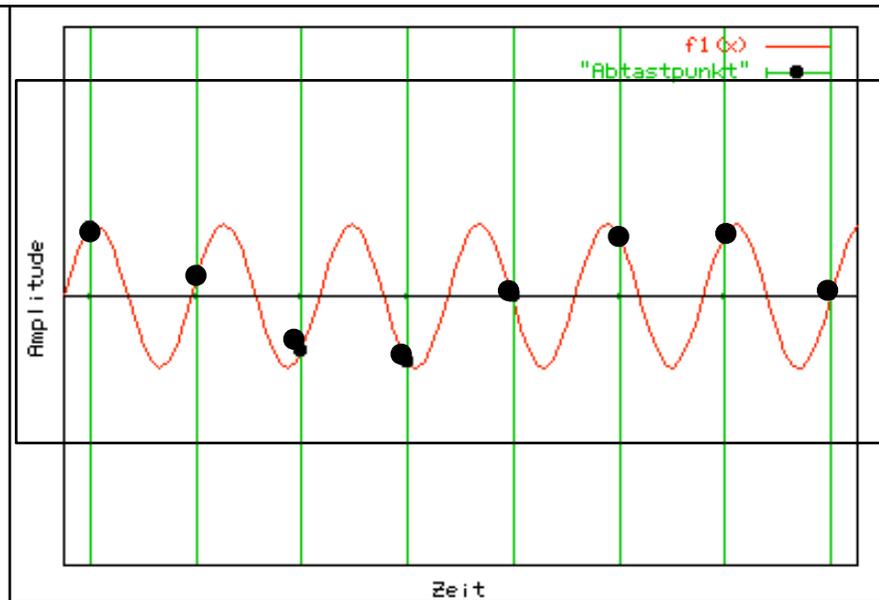
$x, y \in [-1000, 1000]$



Abtasttheorem

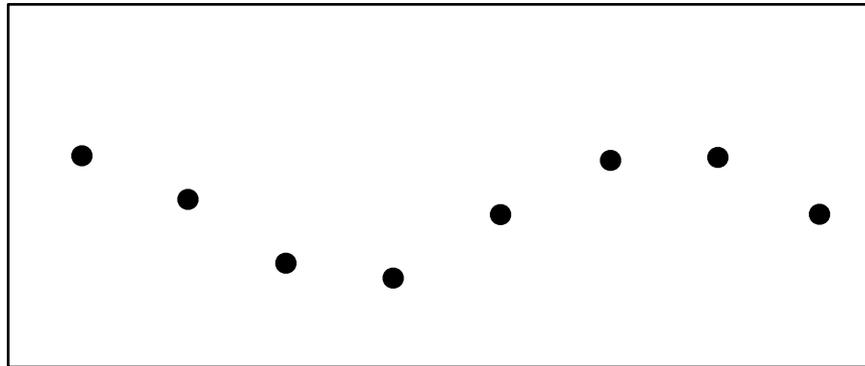
Problem:

Die rote Kurve wird nur an wenigen Punkten gemessen:
Es werden nur die schwarzen Punkte ermittelt.



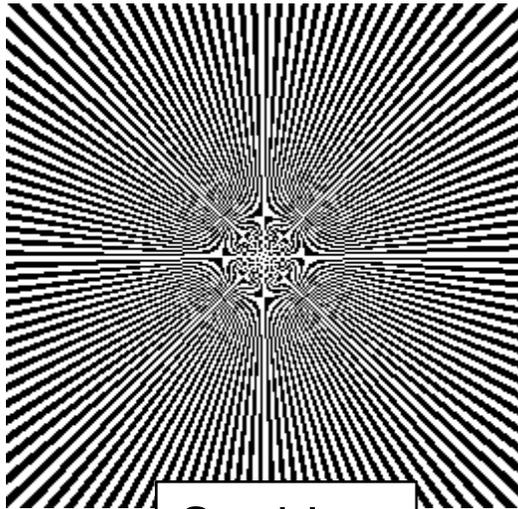
Abtasttheorem

Der Verlauf der schwarzen Punkte wird als eine Kurve niedrigerer Frequenz interpretiert: Alias

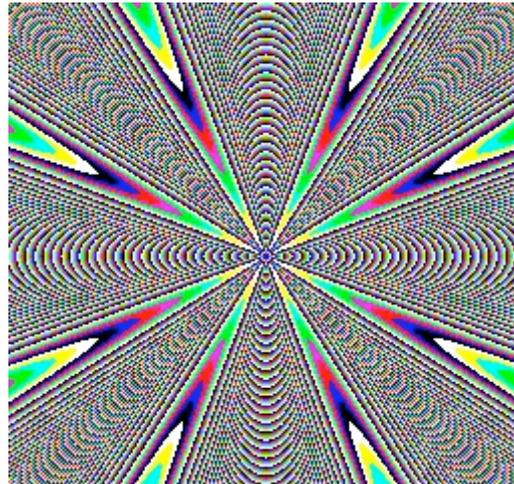


Moiré Effekt

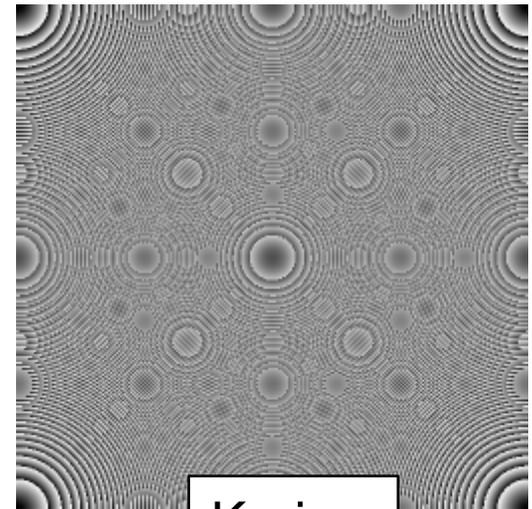
www.jbum.com/jbum/pixmagic/galmoire.html



Strahlen



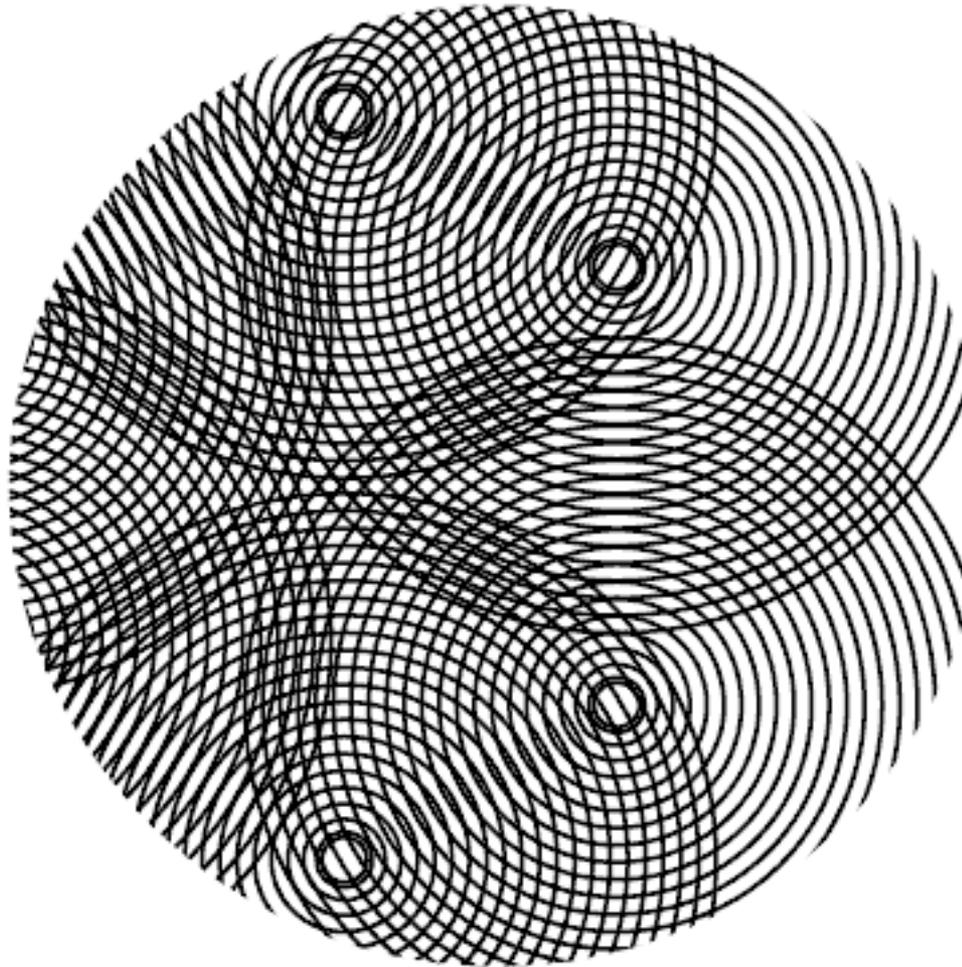
Konzentrische Kreise



Kreise

Pixelgröße jeweils nicht ausreichend für exakte Wiedergabe.
Es entstehen tief-frequente Muster.

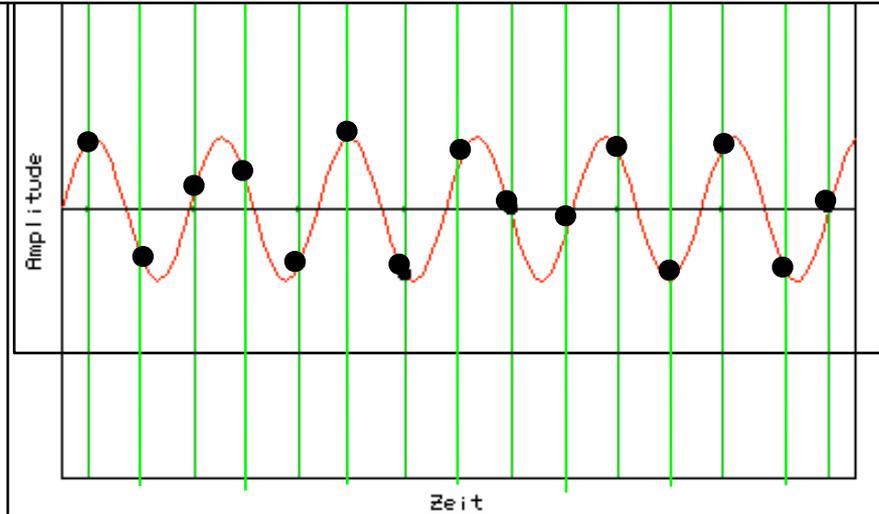
Moiré Effekt



Abtasttheorem

Wie oft müsste man messen?

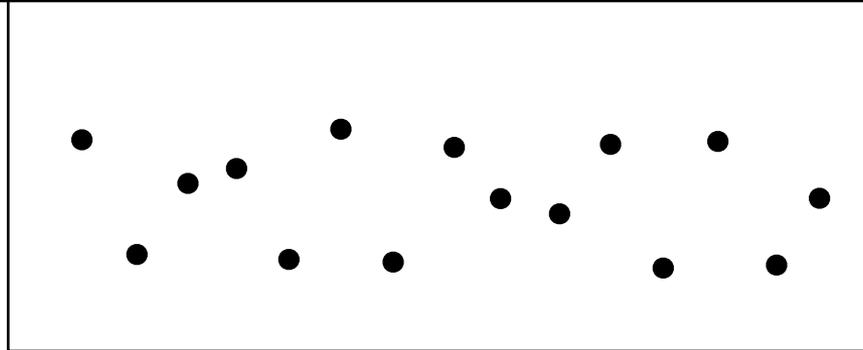
Beispiel: Verkleinerung der Messintervalle
-- Mehr der Messpunkte



Abtasttheorem

Wie oft müsste man messen?

Beispiel: Verkleinerung der Messintervalle
-- Mehr der Messpunkte



Abtasttheorem

Abtasttheorem

Für korrekte Wiedergabe muss gelten:

Pro Wellenlänge T mindestens 2 Abtastpunkte,

d.h. Abtastrate $\Delta x < T/2$. (Nyquist-Bedingung)

oder:

Abtast-Frequenz mindestens doppelt so groß

wie höchste vorkommende Frequenz.

Beispiel:

Sinusförmige Intensitätsfunktion $\sin(\omega x)$ mit Frequenz $\omega = 2\pi / T$ und Wellenlänge T .

Bei Abtastrate Δx werden Frequenzen $\omega > \pi / \Delta x$ nicht mehr korrekt wiedergegeben.

Abtasttheorem

Die Bedingung muss in jeder Richtung des Bildes gelten.
Wenn die Bedingung nicht erfüllt ist, passiert zweierlei:

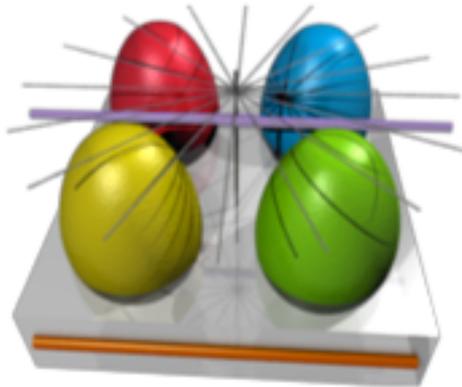
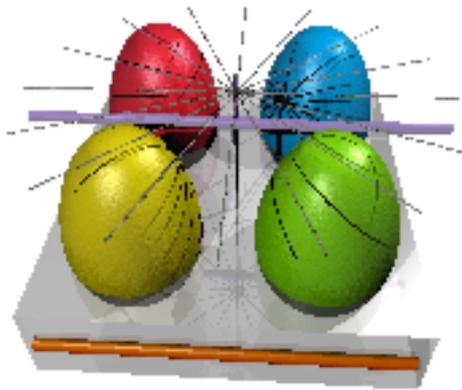
Korrekte Bildinhalte werden nicht übermittelt.
Es entstehen zusätzliche störende Muster.

Die störenden Muster können durch vorherige
Tiefpassfilter verhindert werden:

Nicht korrekt abtastbare hochfrequente Bildinhalte
werden eliminiert.

Anti-Aliasing

Bei Antialiasing (auch Treppenglättung oder Kantenglättung genannt) werden Signalanteile oberhalb der Nyquist-Frequenz durch Tiefpass-Filterung gedämpft.



Kamera, Abbildung, Projektion

Voraussetzung für scharfe Darstellung:

Jeder Bildpunkt wird nur von einem Licht-Strahl aus der Umgebung getroffen (näherungsweise: Licht aus einem engen Raumwinkel)

Das kann erreicht werden durch

- Lochkamera
- Kollimator (Bündel von Röhren)
- Linsensysteme (Bündelung von Lichtstrahlen: Aus dem gleichen Ursprung stammende Strahlen werden auf den gleichen Bildpunkt gelenkt.)

Bei großer Linse (großer Öffnung der Blende):
höhere Lichtausbeute, aber geringere Tiefenschärfe.

Kalibrierung + Korrekturen

Die Bilder (bzw. allgemein: Sensordaten) sind nicht immer so, wie man glauben möchte ...



Fig. 2. Barrel distortion of the field with a 4.2 mm lens



Kameramodellierung

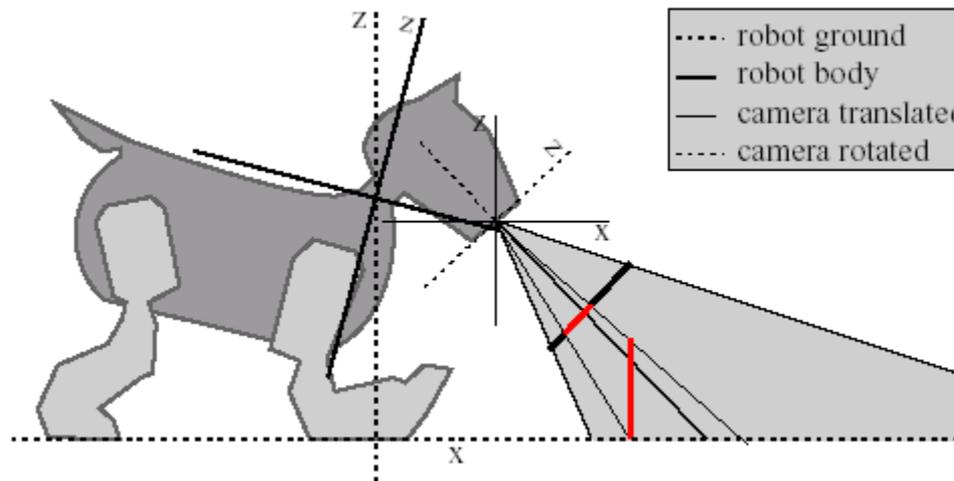
Unterschiedliche Koordinatensysteme.

Grauer Kegel:

Blickrichtung, optische Achse, virtuelle Bildebene

Rote Linien:

Ein Objekt und seine Projektion auf die Bildebene



Vorwärtsmodell

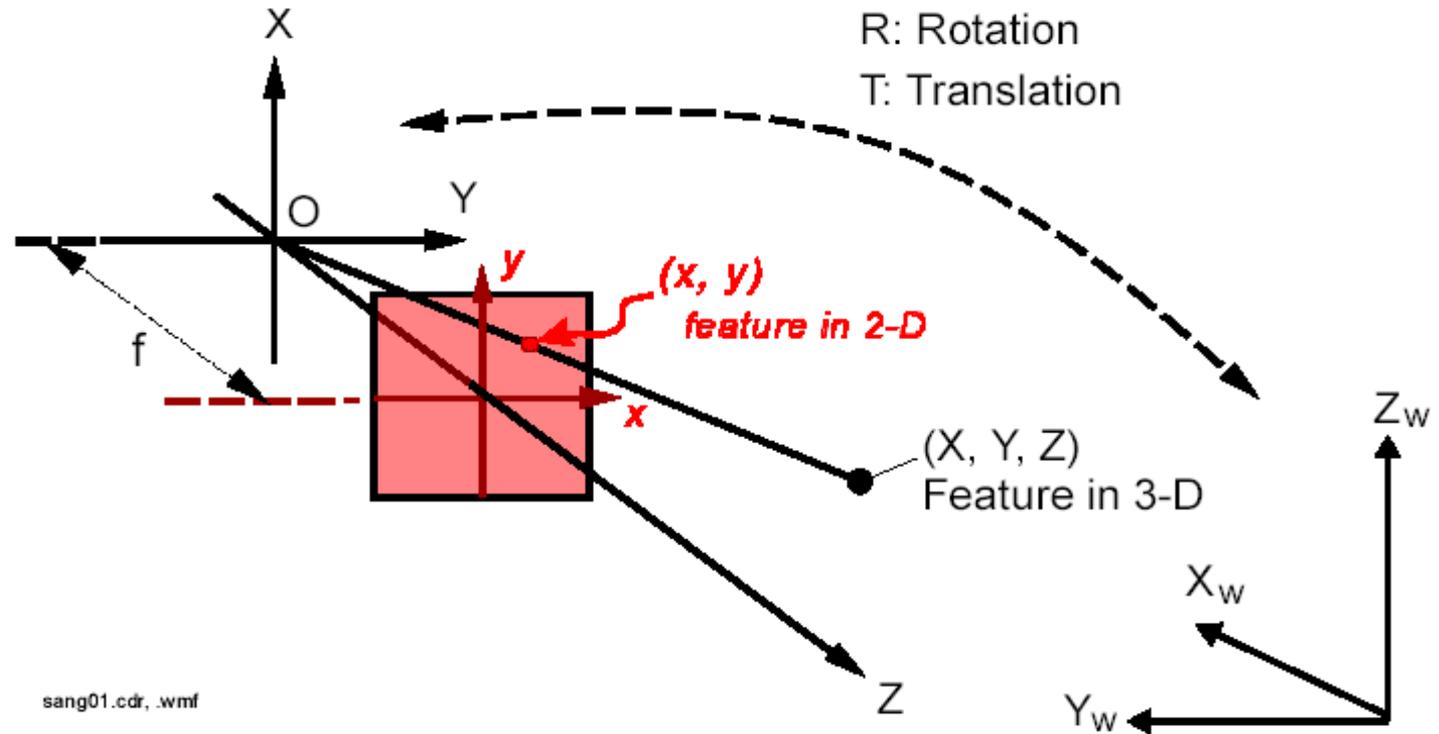
Projektion Raum \rightarrow Bildebene

Reduktion der 3D-Welt auf 2D-Abbildung
mit räumlicher Information bzgl.

- Orientierung der Objekte
(Lage der Bildebene relativ zu Objekten)
- Lichtverhältnisse (Beleuchtung, Schatten)
- Relative Größe der Objekte
- Beziehungen zwischen Objekten

Interpretation ist mehr als nur Umkehrung

Kameramodel (,,Vorwärtsmodell“)



X, Y, Z : Kamera-Koordinatensystem

Z -Achse = optische Achse, Bildebene senkrecht dazu.

f : Brennweite

X_w, Y_w, Z_w : Welt-Koordinatensystem

Kameraparameter („Kameramatrix“)

Extrinsische Parameter:

Pose im Raum (bzgl. Weltkoordinaten X_w, Y_w, Z_w)

- Lage des Brennpunkts (3 DOF)
- Orientierung (3 DOF):

Kamerakoordinaten X, Y, Z mit Ursprung im Brennpunkt und Z in Richtung optische Achse

Intrinsische Parameter:

Lage der Bildebene (bzgl. Kamerakoordinaten)

- Koordinate f (=Brennweite) auf Z -Achse (1 DOF)
- Durchstoßpunkt der optischen Achse (2 DOF):

Bildkoordinaten x, y mit Ursprung auf Z -Achse und Orientierung parallel zu X - Y -Ebene

Vorwärtsmodell

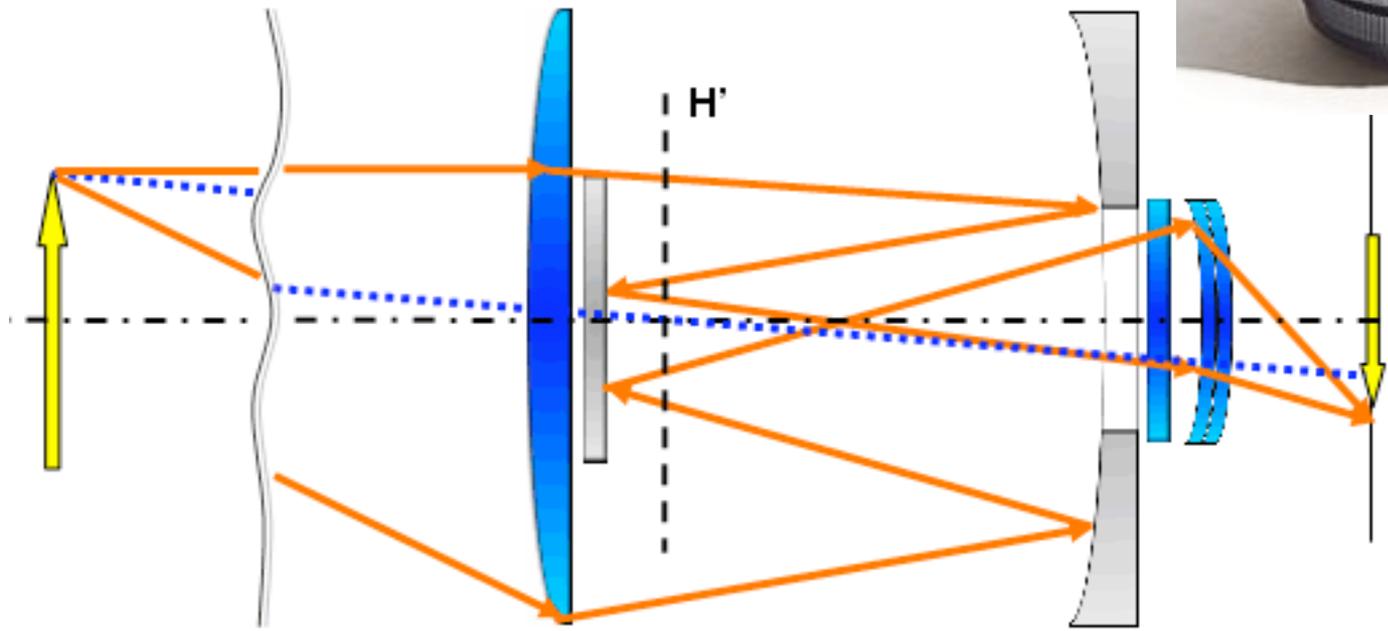
Zu betrachten:

- Szene: Weltobjekte in Weltkoordinaten X_w, Y_w, Z_w
- Kamera mit Kameramatrix
- Bild: Bildobjekte in Bildkoordinaten x, y

Vorwärtsmodell:

Berechnung der Bildobjekte zu den Weltobjekten bei gegebener Kameramatrix

Spiegelteleobjektiv



Perspektivische Projektion (Zentralperspektive)

Die Bildkoordinaten (x,y) sind durch Kamerakoordinaten (X,Y,Z) eindeutig bestimmt.

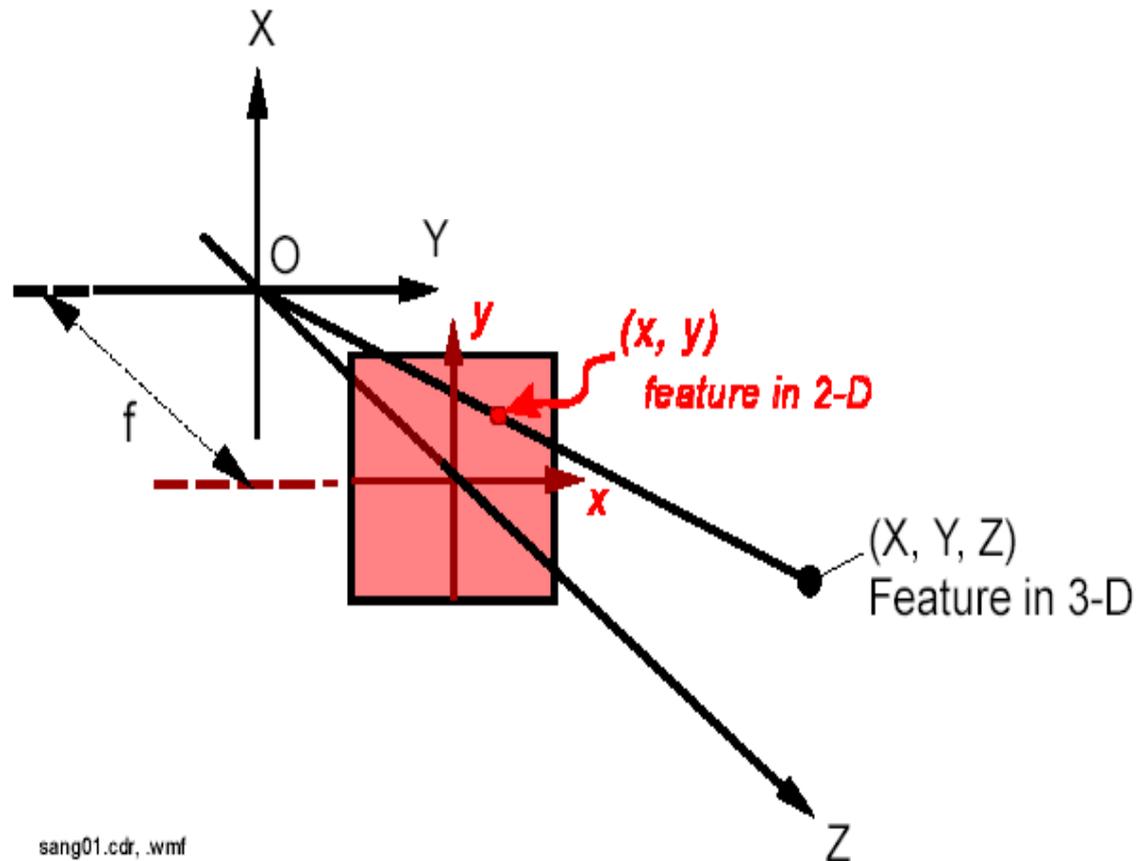
Strahlensatz:

$$Z : f = X : x = Y : y$$

$$x = f/Z \cdot X$$

$$y = f/Z \cdot Y$$

Exakt gültig nur unter idealen Bedingungen (Lochkamera)



Perspektivische Projektion

Zentralprojektion durch Transformationsmatrix Z im Raum der Kamerakoordinaten, dargestellt in homogenen Koordinaten

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix} \begin{pmatrix} \lambda X \\ \lambda Y \\ \lambda Z \\ \lambda \end{pmatrix} = \begin{pmatrix} \lambda X \\ \lambda Y \\ \lambda Z \\ \lambda Z/f \end{pmatrix} \Rightarrow f/Z \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}$$

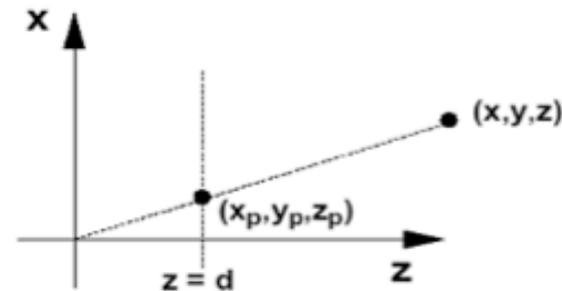
Homogene Transformationsmatrizen

Translation:	\underline{T}	$= \begin{pmatrix} 1 & 0 & 0 & t_x \\ 0 & 1 & 0 & t_y \\ 0 & 0 & 1 & t_z \\ 0 & 0 & 0 & 1 \end{pmatrix}$	Skalierung:	\underline{S}	$= \begin{pmatrix} s_x & 0 & 0 & 0 \\ 0 & s_y & 0 & 0 \\ 0 & 0 & s_z & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$
Rotation um x-Achse:	\underline{R}_x	$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha & 0 \\ 0 & \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$	Perspektive:	\underline{P}	$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & \frac{1}{d} & 1 \end{pmatrix}$
Rotation um y-Achse:	\underline{R}_y	$= \begin{pmatrix} \cos \beta & 0 & \sin \beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin \beta & 0 & \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$	Orthogonale Projektion:	$\underline{P}_{orth,z=0}$	$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$
Rotation um z-Achse:	\underline{R}_z	$= \begin{pmatrix} \cos \gamma & -\sin \gamma & 0 & 0 \\ \sin \gamma & \cos \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$			

Perspektivische Projektion

Projiziere alle Punkte entlang der z -Achse auf die $z = d$ Ebene, Augpunkt im Ursprung

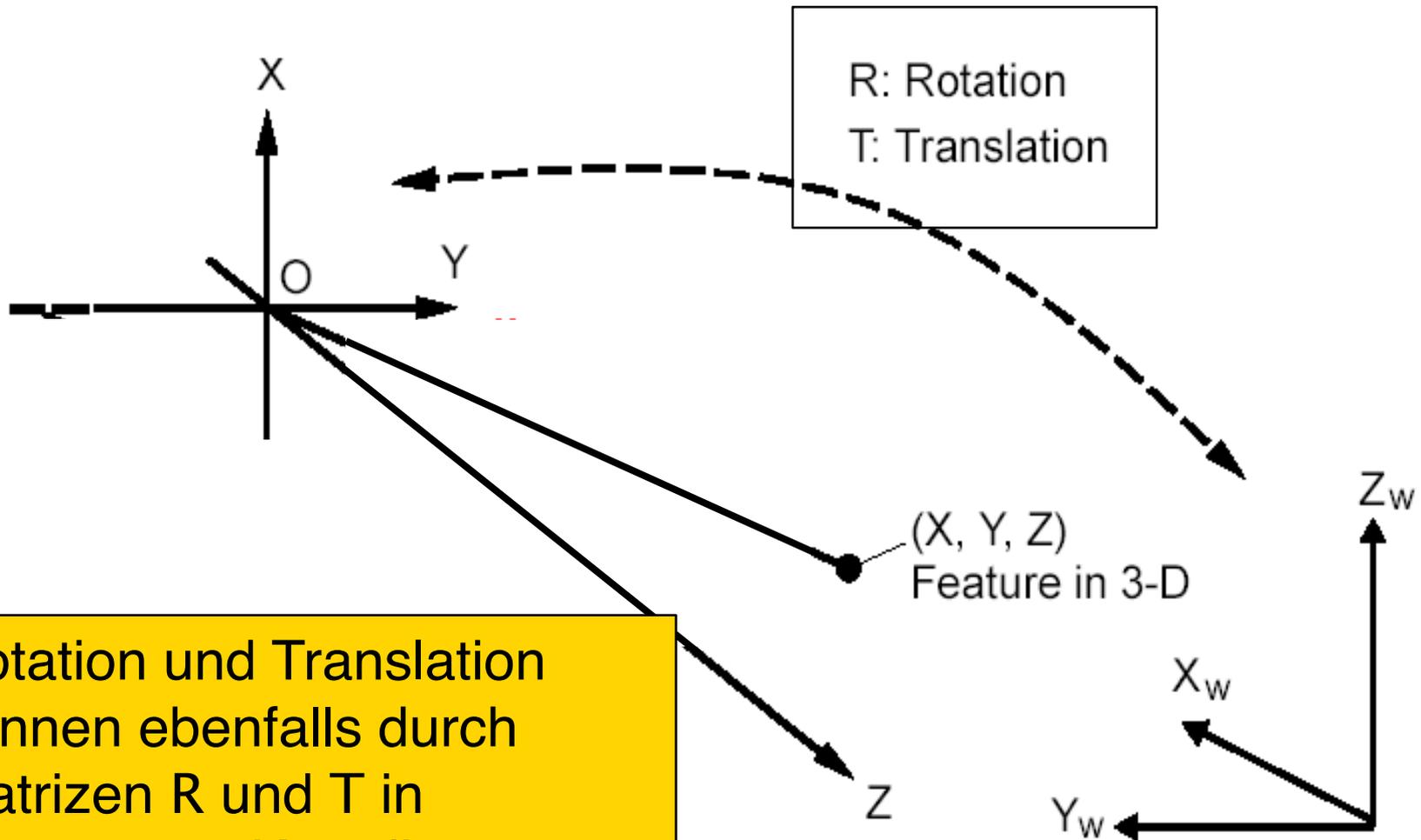
$$\begin{aligned}x_p &= \frac{d \cdot x}{z} = \frac{x}{z/d} \\y_p &= \frac{d \cdot y}{z} = \frac{y}{z/d} \\z_p &= d\end{aligned}$$



homogenize

$$\begin{pmatrix} x * d / z \\ y * d / z \\ d \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \\ z / d \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/d & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

Übergang zu Kamerakoordinaten



Rotation und Translation können ebenfalls durch Matrizen R und T in homogenen Koordinaten dargestellt werden.

Projektionsmatrix

Es ergibt sich eine 4x4-Matrix P' (in homogenen Koordinaten)

$$P' = Z \cdot R \cdot T$$

für den Übergang von der Objektdarstellung in Weltkoordinaten zur Bilddarstellung in Kamerakoordinaten:

$$(X, Y, Z, 1)^T = P' \cdot (X_w, Y_w, Z_w, 1)^T$$

Daraus kann eine 3x4-Matrix P für den Übergang von Weltkoordinaten zu Bildkoordinaten erzeugt werden (Z ist konstant f):

$$(x, y, 1)^T = P \cdot (X_w, Y_w, Z_w, 1)^T$$

Perspektivische Abbildung von Geraden

Objekt-Gerade in Kamerakoordinaten:

$$(A, B, C)^T + \alpha \cdot (X, Y, Z)^T$$

wird abgebildet auf eine Bild-Gerade in Bildkoordinaten:

$$\left(f / (C + \alpha Z) \right) \cdot (A + \alpha X, B + \alpha Y)^T$$

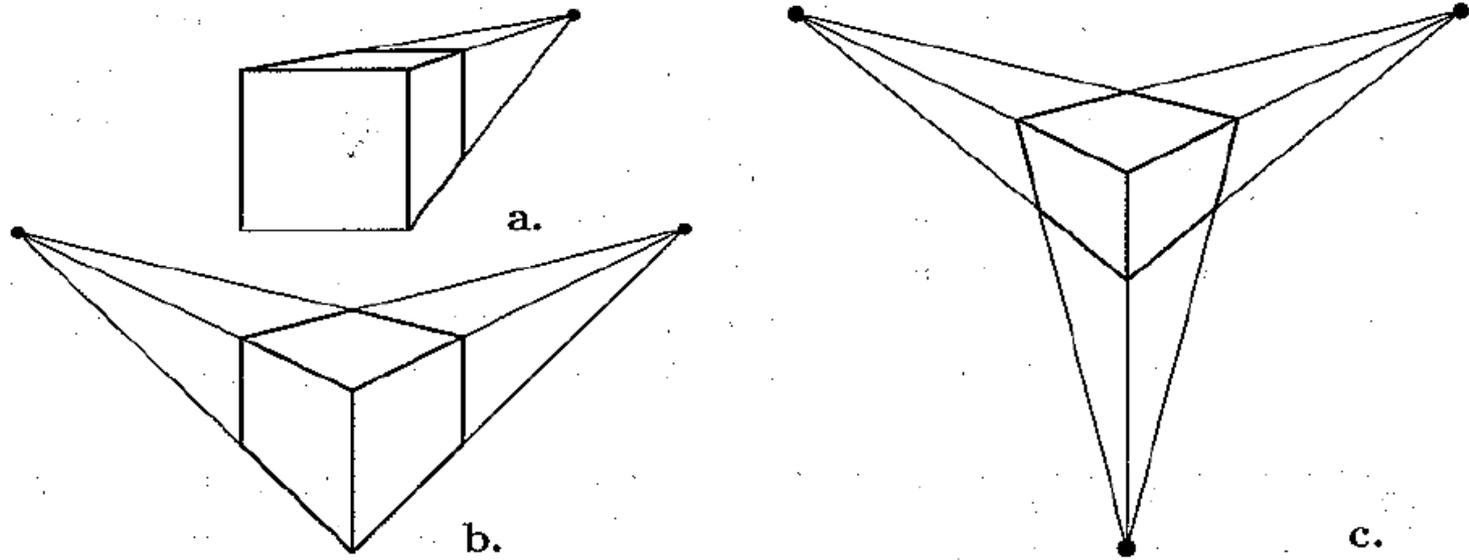
Für $Z \neq 0$ gilt:

$$\lim_{\alpha \rightarrow \infty} \left(f / (C + \alpha \cdot Z) \right) \cdot (A + \alpha X, B + \alpha Y)^T = f / Z \cdot (X, Y)^T,$$

d.h. jeweils für alle Objekt-Geraden mit gleicher Richtung enden die Bild-Geraden in einem Fluchtpunkt.

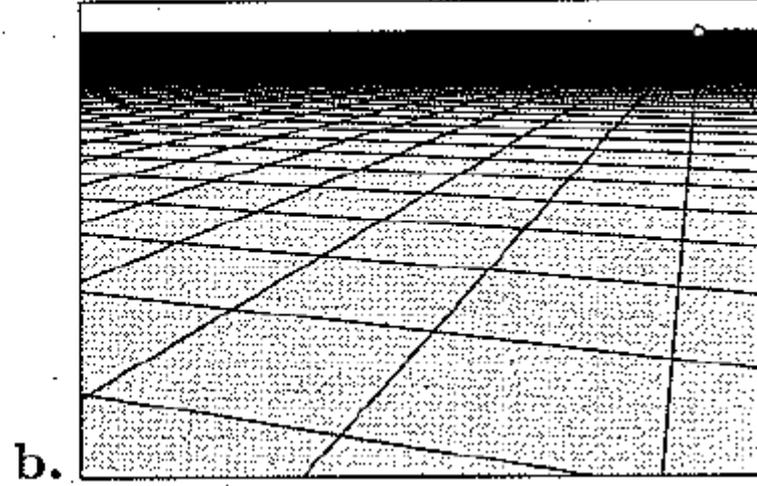
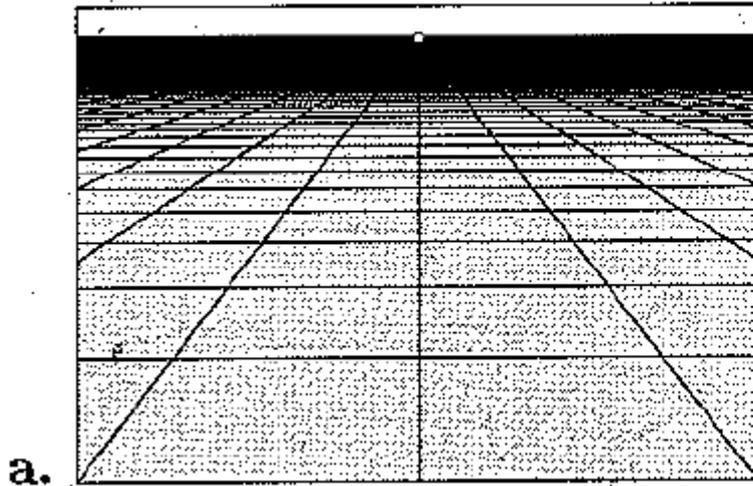
Für $Z=0$ und $C \neq 0$ (Objekt-Gerade verläuft parallel zur Bildebene) ergeben sich vollständige Bild-Geraden im Bild.

Perspektivische Abbildung von Geraden



Je nach Orientierung des Objekts (Kanten parallel zur Bildebene) gibt es 1, 2 oder 3 Fluchtpunkte (Ein-/Zwei-/Dreipunktperspektive)
[aus Mallot]

Perspektivische Abbildung von Geraden



Je nach Orientierung der Ebene (Kanten parallel zur Bildebene) gibt es 1 oder 2 Fluchtpunkte (Ein-/Zweipunktperspektive).

[aus Mallot]

Auch die Winkel zwischen den Geraden variieren.

Rückwärtsmodelle

1. Als Teil der Bildinterpretation:

Berechnung der Weltobjekte zu Bildobjekten bei gegebener Kameramatrix

(Geometrisch schlecht gestelltes Problem wegen Übergang 2D nach 3D)

2. Kamerakalibrierung:

Berechnung der Kameramatrix bei gegebenen Weltobjekten und Bildobjekten

3. Selbstlokalisierung

Berechnung der extrinsischen Kameraparameter bei gegebenen Weltobjekten und Bildobjekten und gegebenen intrinsischen Kameraparametern

Rückwärtsmodell (Interpretation)

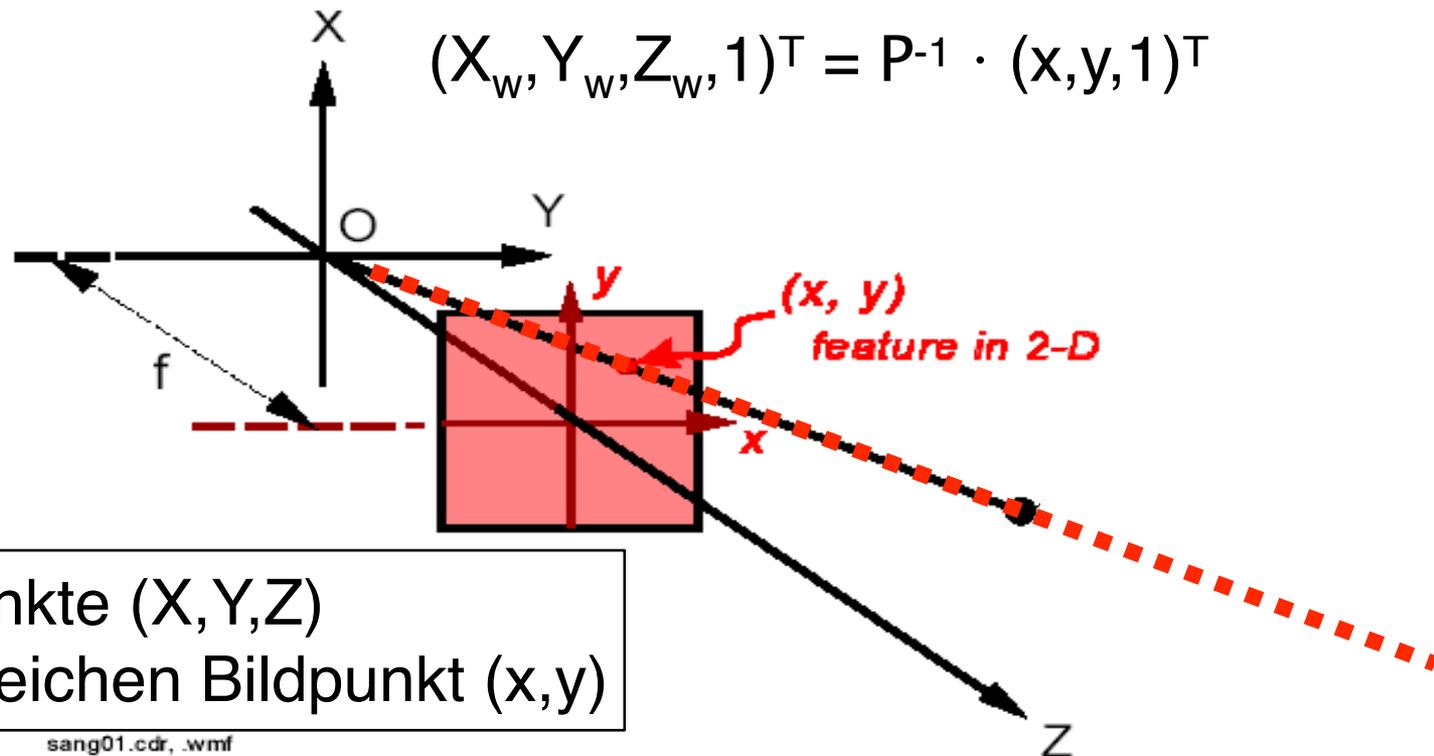
Rekonstruktion des 3D-Raumes aus 2D-Bild.

Mehrdeutigkeiten („schlecht gestelltes Problem“):

Urbilder eines Bildpunktes auf einer Geraden im Raum

Geometrie allein nicht ausreichend für Rekonstruktion

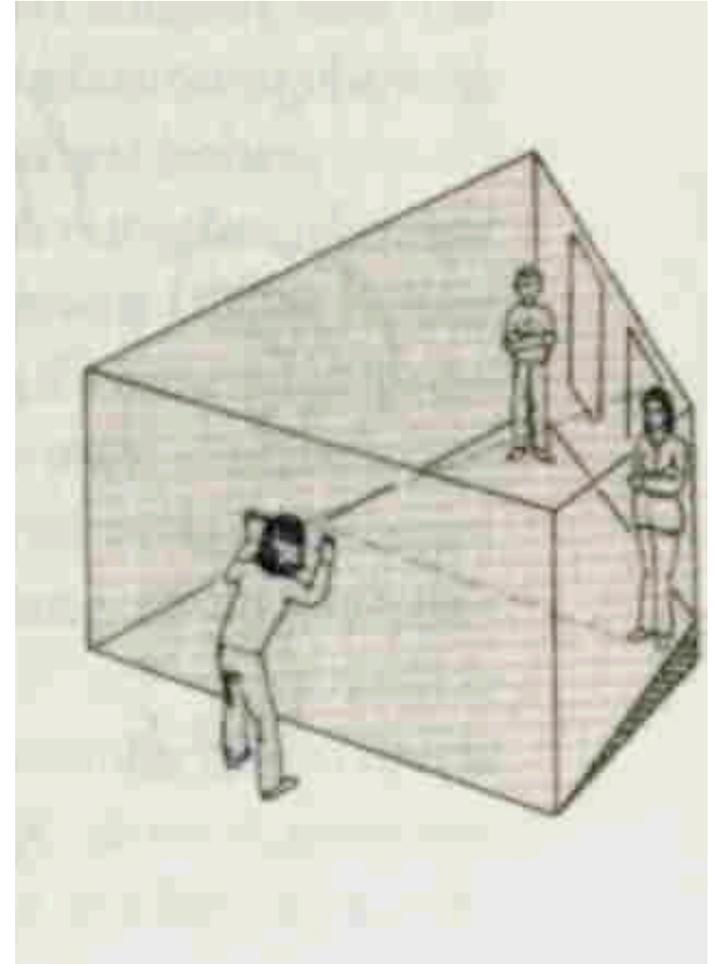
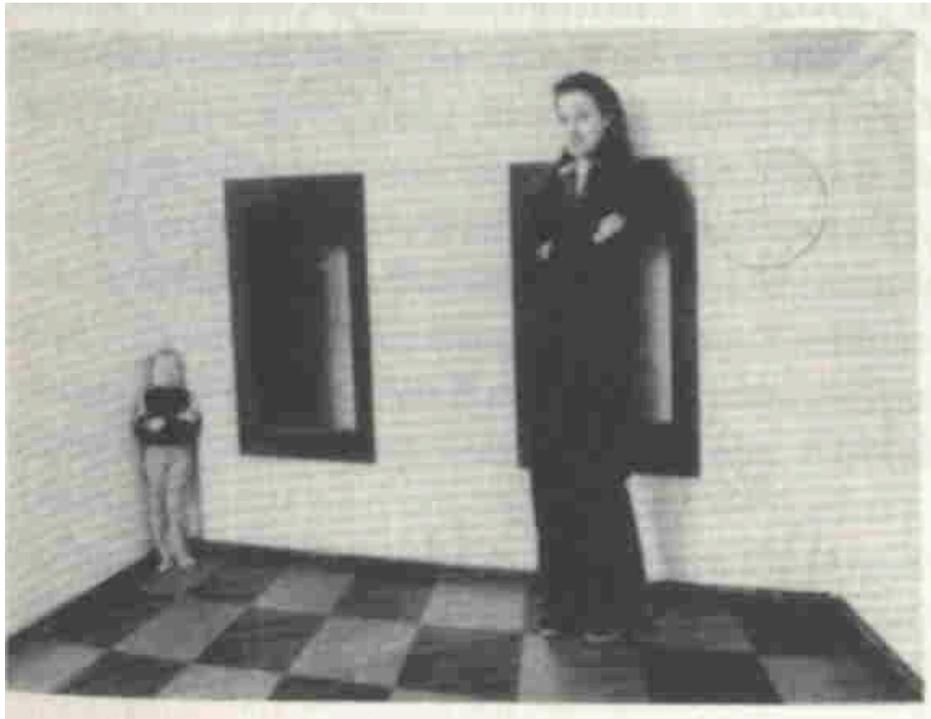
$$(X_w, Y_w, Z_w, 1)^T = P^{-1} \cdot (x, y, 1)^T$$



Alle roten Punkte (X, Y, Z)
haben den gleichen Bildpunkt (x, y)

Rückwärtsmodell (Interpretation)

Größen-Illusion: Ames-Zimer



Rückwärtsmodell (Interpretation)

Stereo-Sehen:

2 Kamerabilder aus unterschiedlichen Positionen
(2 Kameras oder Kamera bewegen)

Urbild von korrespondierenden Bildpunkten im
Schnittpunkt der jeweiligen Geraden

– ermöglicht Rückrechnung auf Tiefe

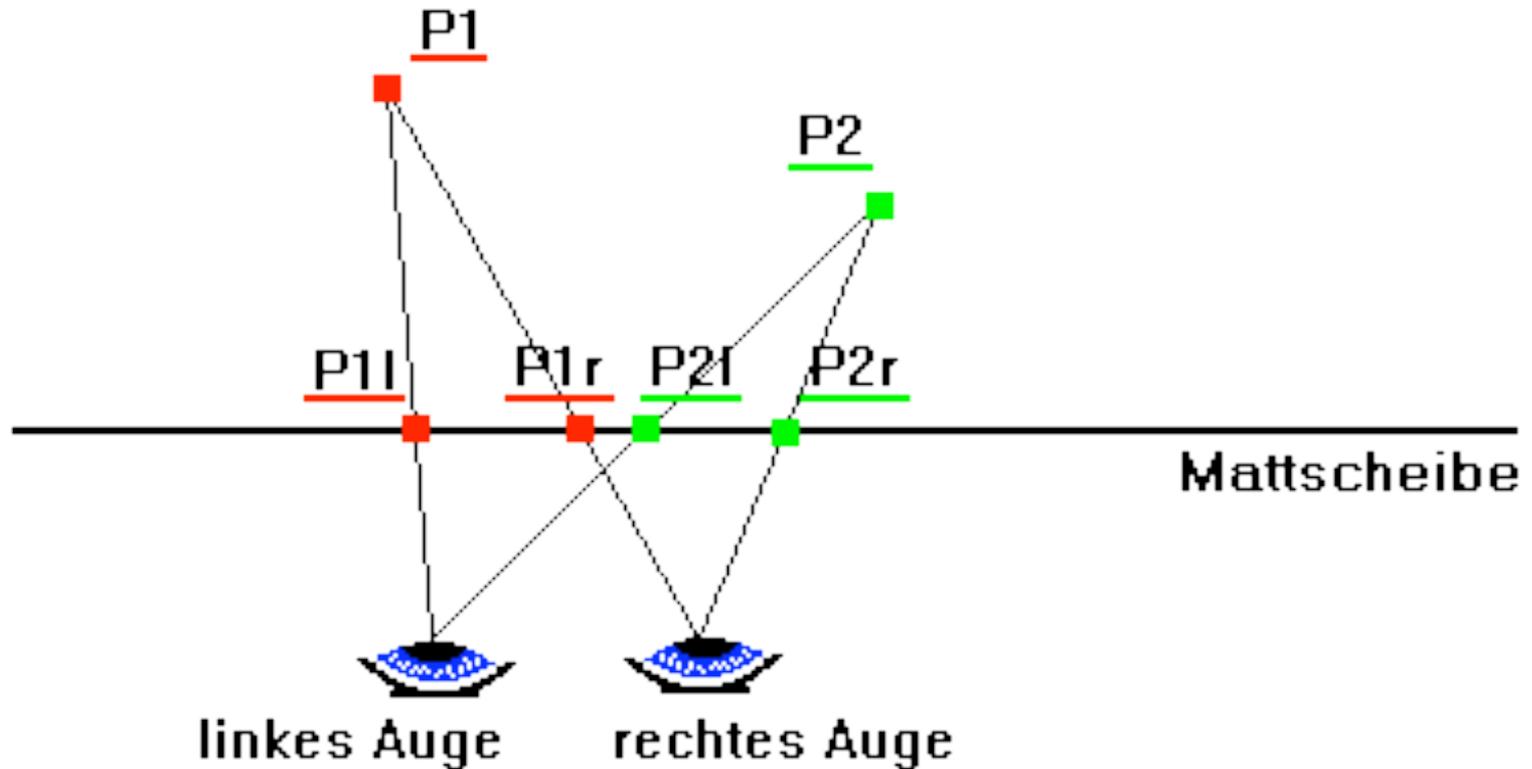
Korrespondenz-Problem:

Welche Bildpunkte gehören zusammen?

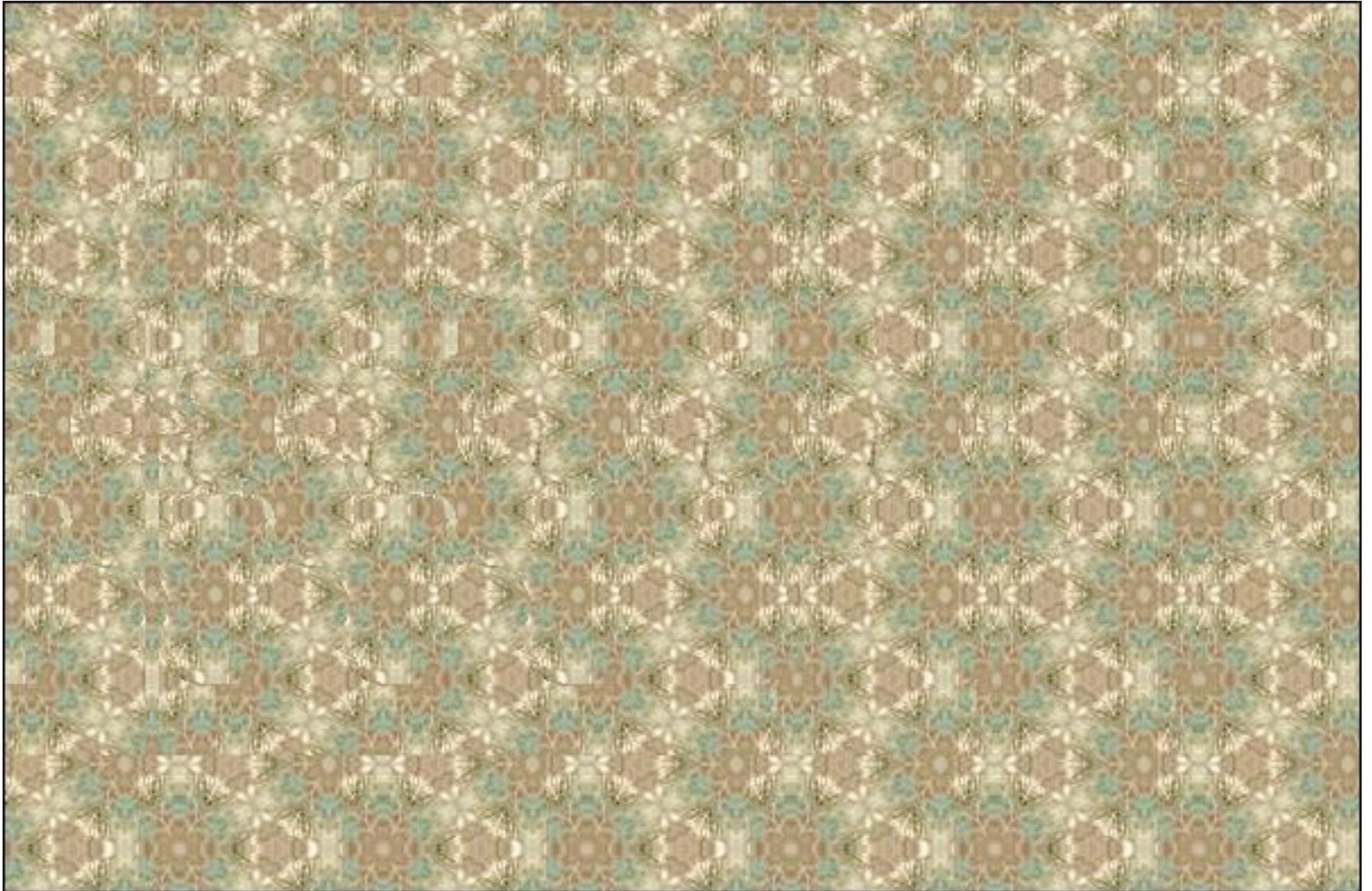
- Vergleich von Bildmerkmalen
- Korrelationsverfahren

Autostereogramm

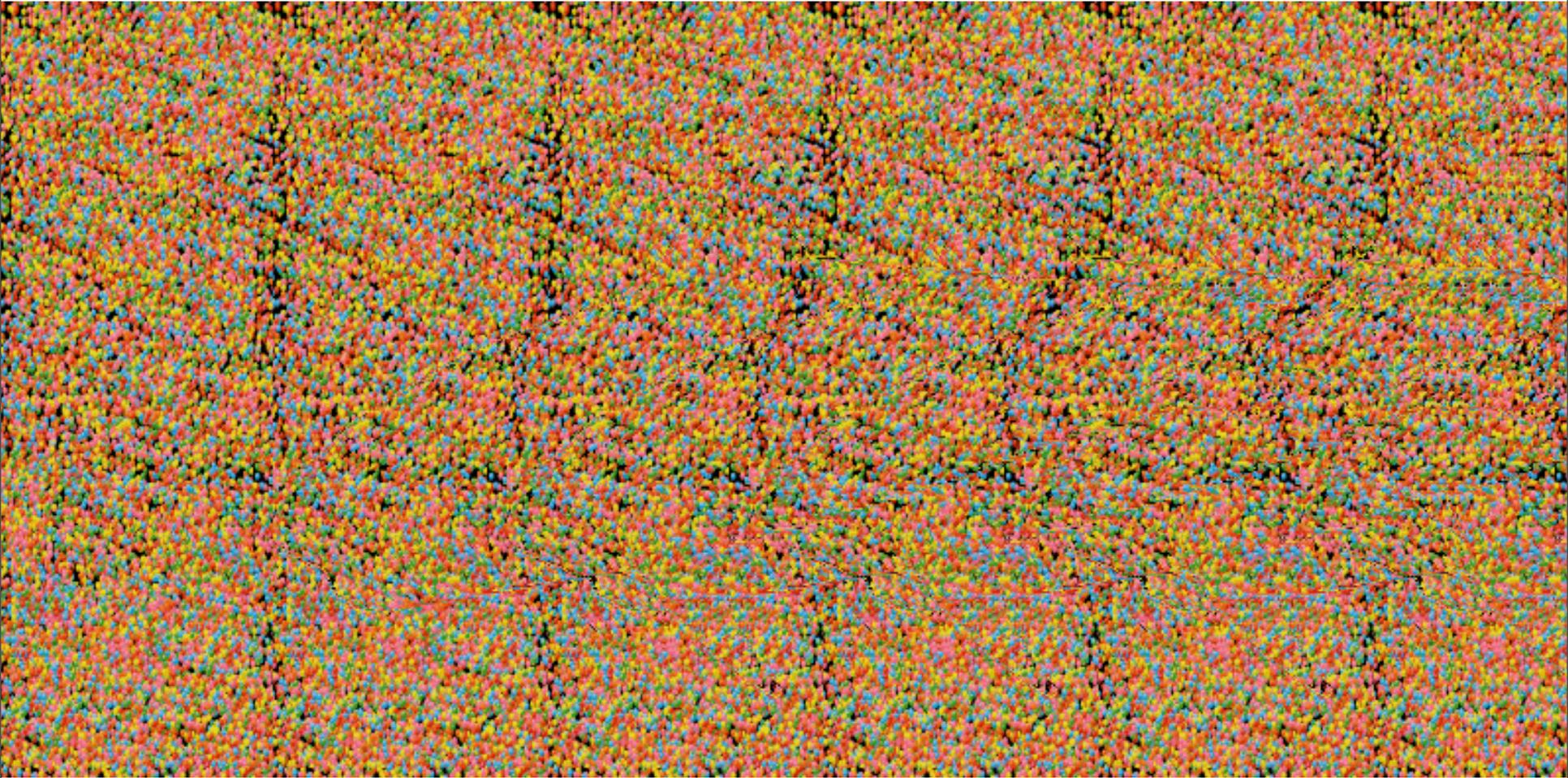
Korrespondenz durch korrespondierende Muster



Autostereogramm

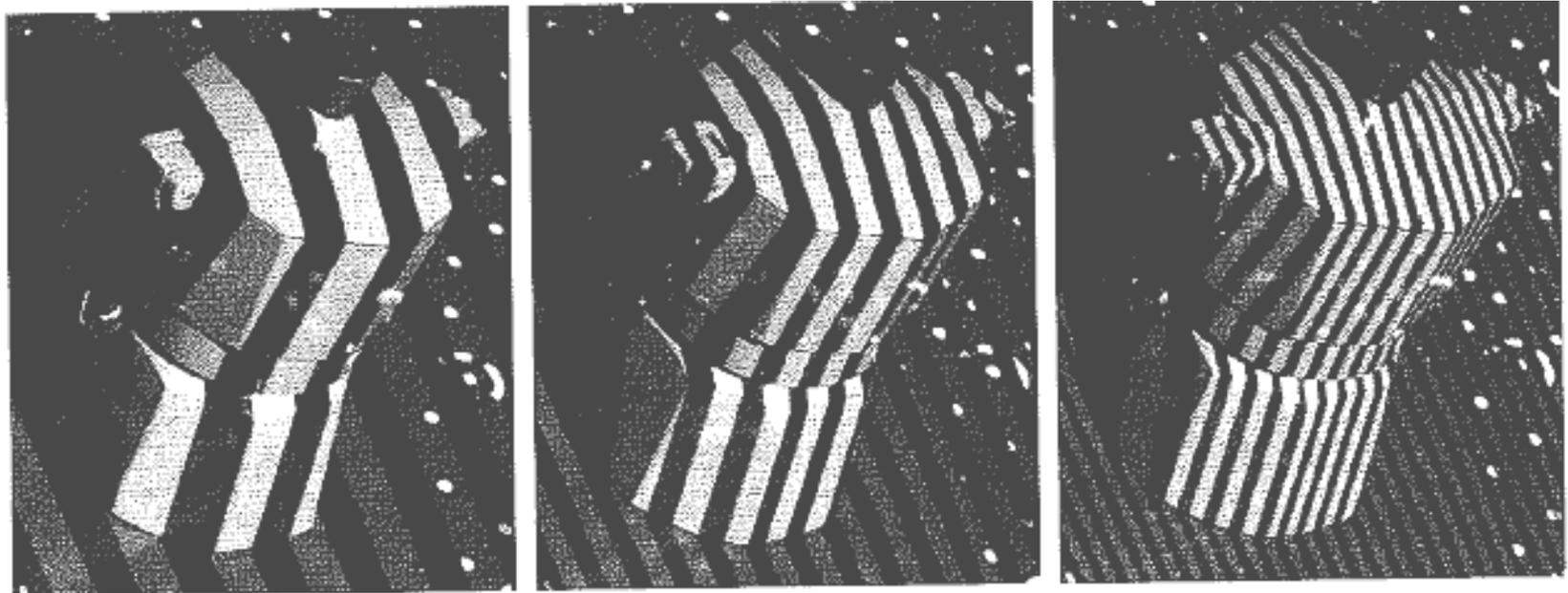


Autostereogramm 4D



Rückwärtsmodell (Interpretation)

Aktive Triangulation: Prinzip Stereosehen
Statt 2. Kamera wird ein Streifenbild projiziert



Jähne, S. 220

Rückwärtsmodell (Interpretation)

Weitere technische Möglichkeiten:

- Kombination mit Entfernungsmessung
(z.B. Laser: misst exakte Entfernung,
liefert aber keine Helligkeits- oder Farbinformation)
- Bildfolgen, Bewegung

Weitere Möglichkeiten bei der Interpretation

Zusatzwissen ausnutzen

(Größe, Umfeld, räumliche Beziehungen, ...)

**Voraussetzung für maschinelle Interpretation:
Bildobjekte (Kanten, Flächen usw.) identifizieren**

Rückwärtsmodell (Kalibrierung)

Übergangsmatrix von Weltkoordinaten zu Bildkoordinaten:

$$(x, y, 1)^T = P \cdot (X_w, Y_w, Z_w, 1)^T$$

Aus korrespondierenden Punkten (Objekt/Bild) können die Einträge in der Matrix berechnet werden:

- Homogenes Gleichungssystem für 11 Parameter lösen (6 Punkte reichen).
- Bzw. Fehlerminimierung bzgl. berechnetem und tatsächlichem Bildpunkt (auch mit weiteren Punkten)

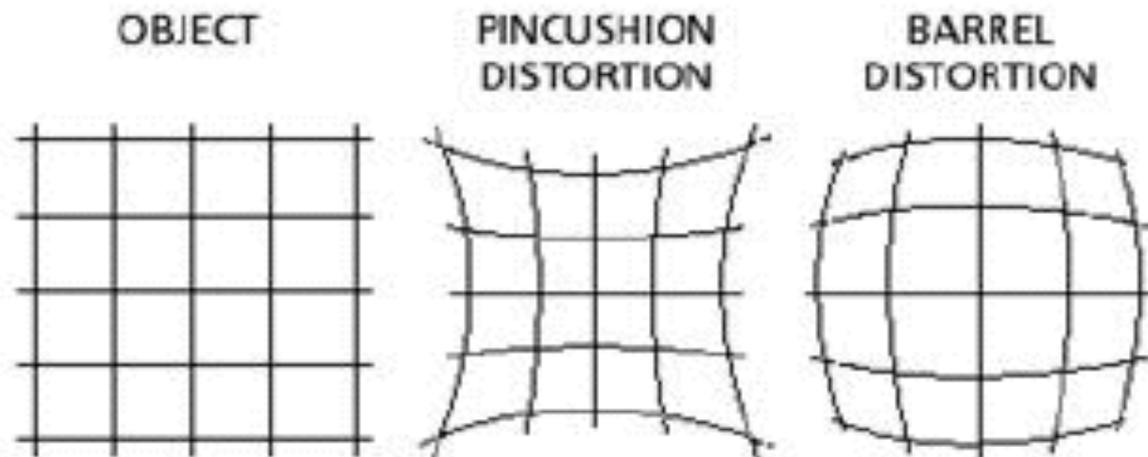
Evtl. Vereinfachungen, z.B. bei bekannten extrinsischen Parametern in definiertem Versuchsaufbau einfachere Bestimmung der intrinsischen Parameter).

Kamerakalibrierung

Für reale Kameras komplexer (keine Zentralprojektion):

- Brechung des Lichts bei Eintritt und Austritt aus Medien
- Linsenmodelle (Farbe, Verzerrung)
- Sensor-Eigenschaften (Norm- Abweichungen einzelner Sensoren)

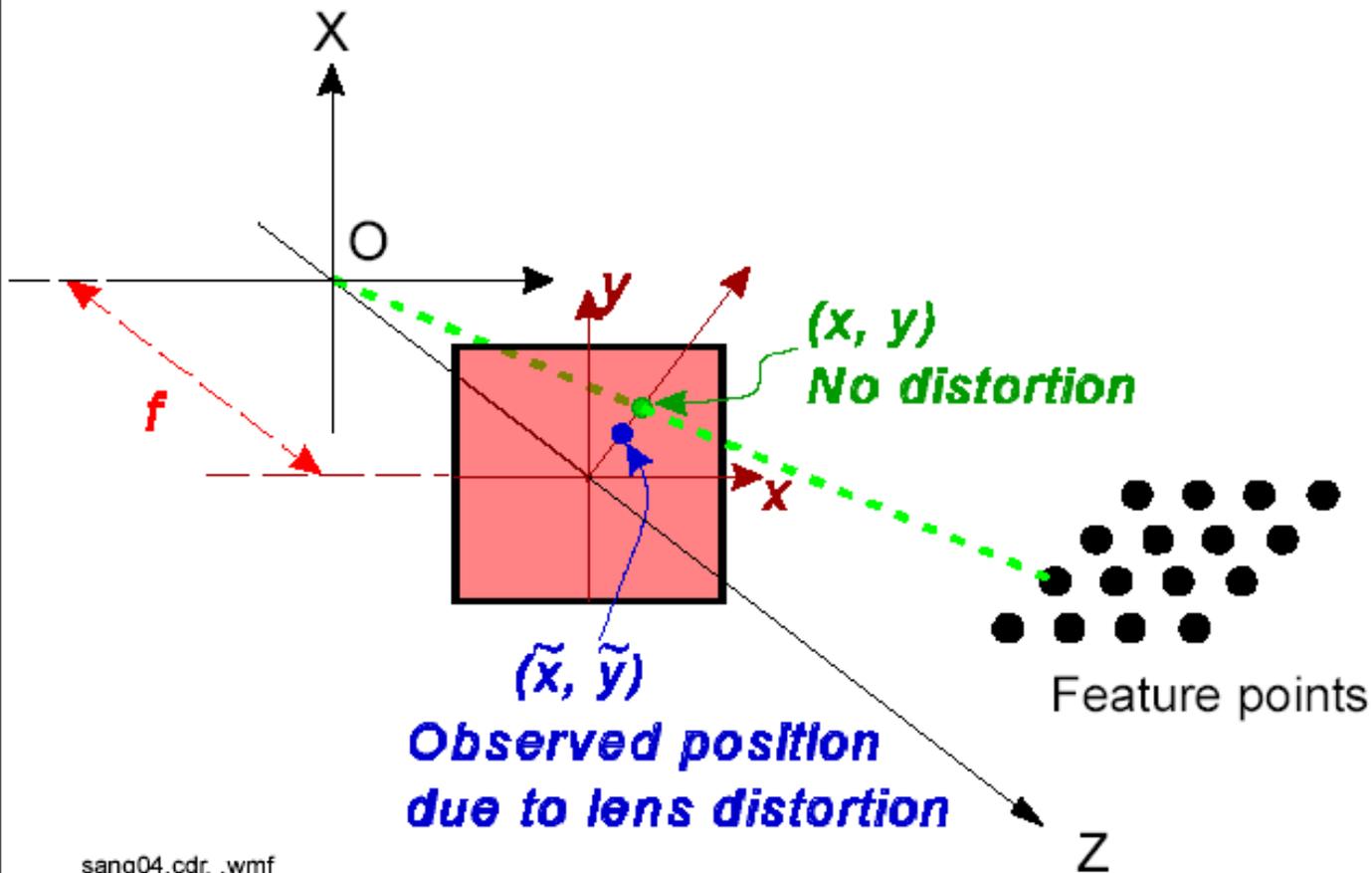
Weitere intrinsische Parameter, aber gleiches Verfahren.



Kamera-Kalibrierung bzgl. Projektion

Ziel: Verzerrungen ausgleichen

Methode: Abbildungsparameter bestimmen mittels zugeordneter vermessener Punkte im Raum und im Bild.



Kamera-Kalibrierung

Beispiel Deckenkamera Small-Size-League
(FU-Fighters Berlin)



Fig. 2. Barrel distortion of the field with a 4.2 mm lens

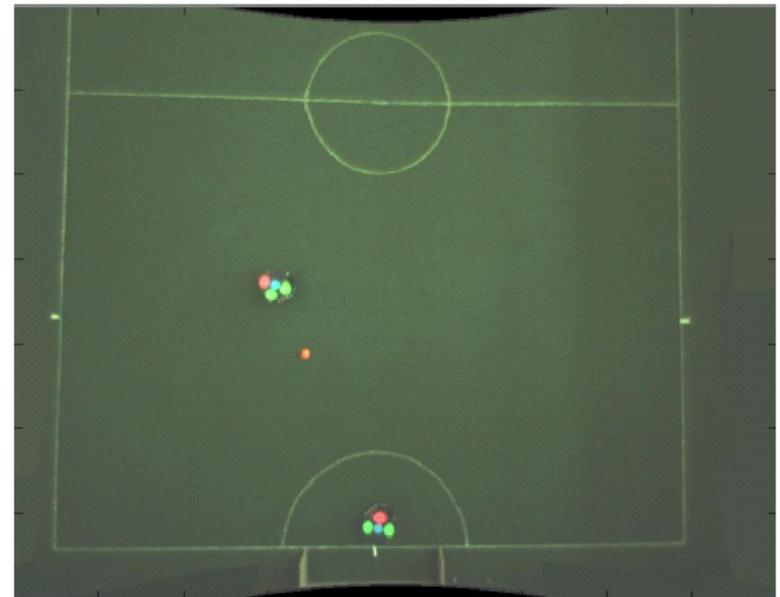


Fig. 6. Corrected field image

Kamera-Kalibrierung bzgl. Farbe

Farb-Korrekturen

- Optik-bedingt insbesondere am Bildrand
 - (AIBO: Am Rand blau verstärkt)
- Sensorbedingt: Unterschiedliche Empfindlichkeit

Für Kalibrierung: Farbwerte Soll/Ist vermessen

Varianten für Auswertung (Projektion und Farbe):

- Neues Bild berechnen
- Bei Bild-Verarbeitung Korrekturen berücksichtigen
(z.B. Scan-Linie entsprechend modifizieren)

Kalibrierung bzgl. Farbe der Umgebung

Notwendig für Auswertung von Farbinformationen
(z.B. für farbkodierte Objekte)

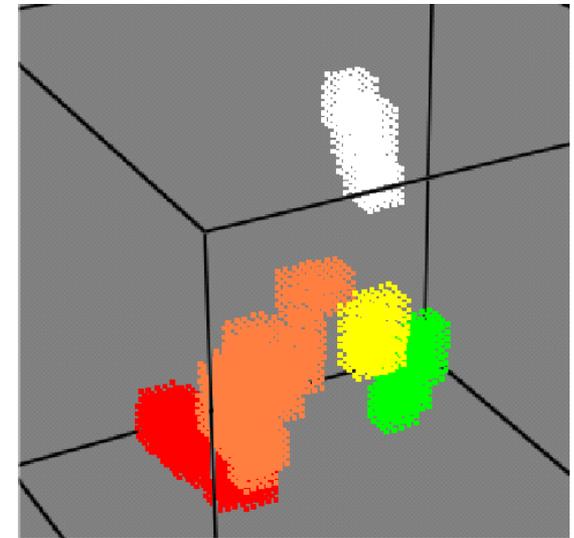
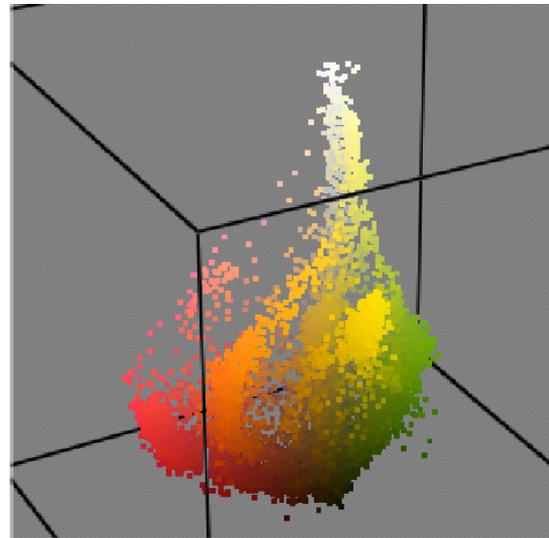
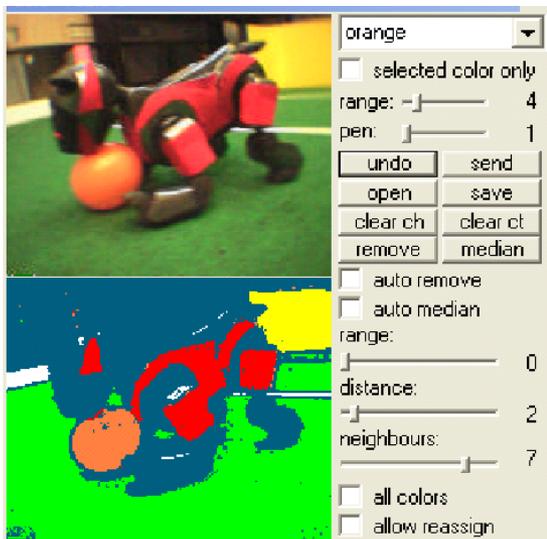


Kalibrierung bzgl. Farbe der Umgebung

Farbtabelle enthält Zuordnung von Farbwerten zu Farbklassen (z.B. Look-up-table).

Oft ist der Bereich einer Farbklasse ein komplexer Ausschnitt des entsprechenden Farbraums.

Der konkrete Bereich ist abhängig von Beleuchtung, Reflexionen, Schatten, ...

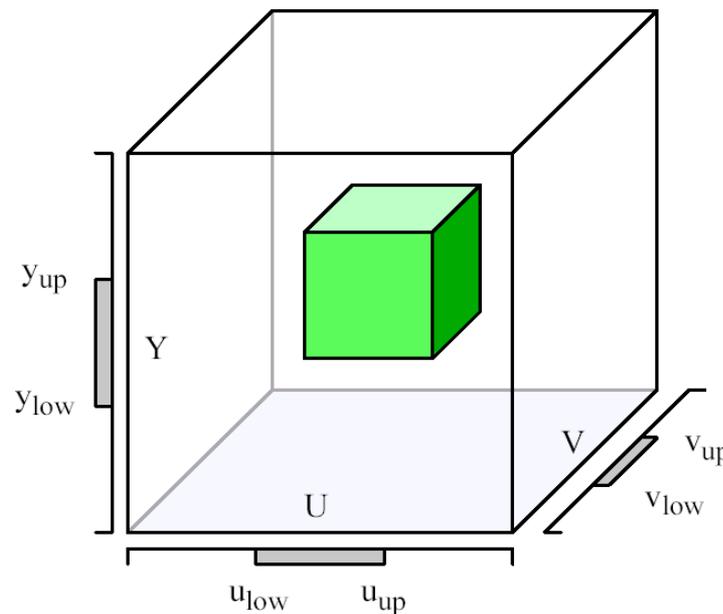


Farbtabelle

Color-look-up-table:

Pixelwerte für „grün“ im Farbraum festlegen

Ziel: als „grün“ klassifizierte Objekte erkennen



Diplomarbeit Jüngerl

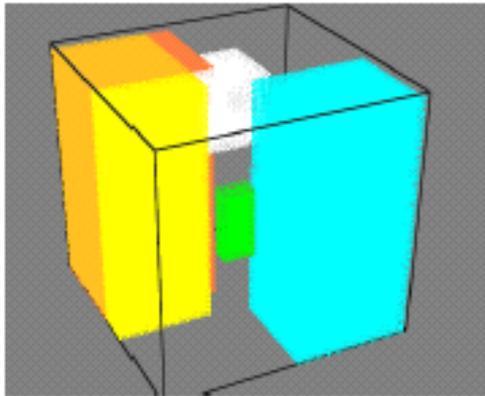
Farbtabelle

a) Farben im Farbraum

b) Bild

c) Klassifiziertes Bild

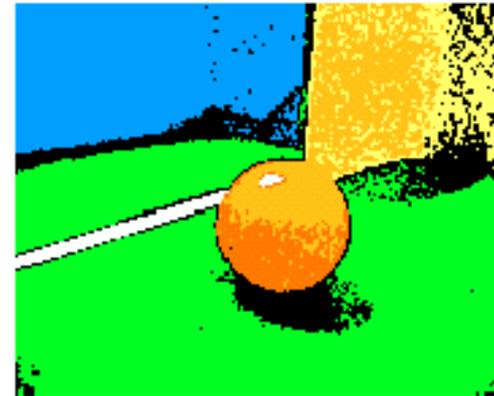
a)



b)



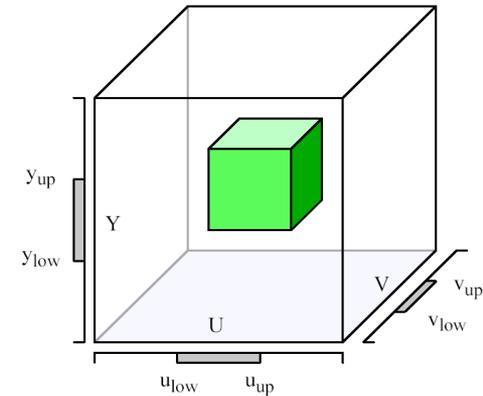
c)



Unterschiedliche Farbräume mit unterschiedlicher Eignung.

Farbtabelle

Beschreibung der Farben:
oft durch Intervalle
bzw. Mengen von Intervallen



Ziel: Gute Klassifizierung,
d.h. möglichst großer Abstand zwischen Bereichen
unterschiedlicher Farben (keine Überschneidungen).

Weiterführende Themen:

- Unterschiedliche Tabellen für unterschiedliche Aufgaben.
- Bessere Lernverfahren (VSM).
- Texturen statt Farben.

Kalibrierung bzgl. Farbe der Umgebung

Kalibrierung:

„wie sieht orange heute aus“

„wie sieht das Objekt heute aus“

Beispielbasiert:

- Farben im Raum (am Roboter) mit Kamera erfassen.
- Im Bild auswählen für Eintrag in Farbtabelle:
 - Manuell („anklicken“), offline
 - Automatisiert (Lernverfahren), auch online möglich

Problem bei Manueller Auswahl:

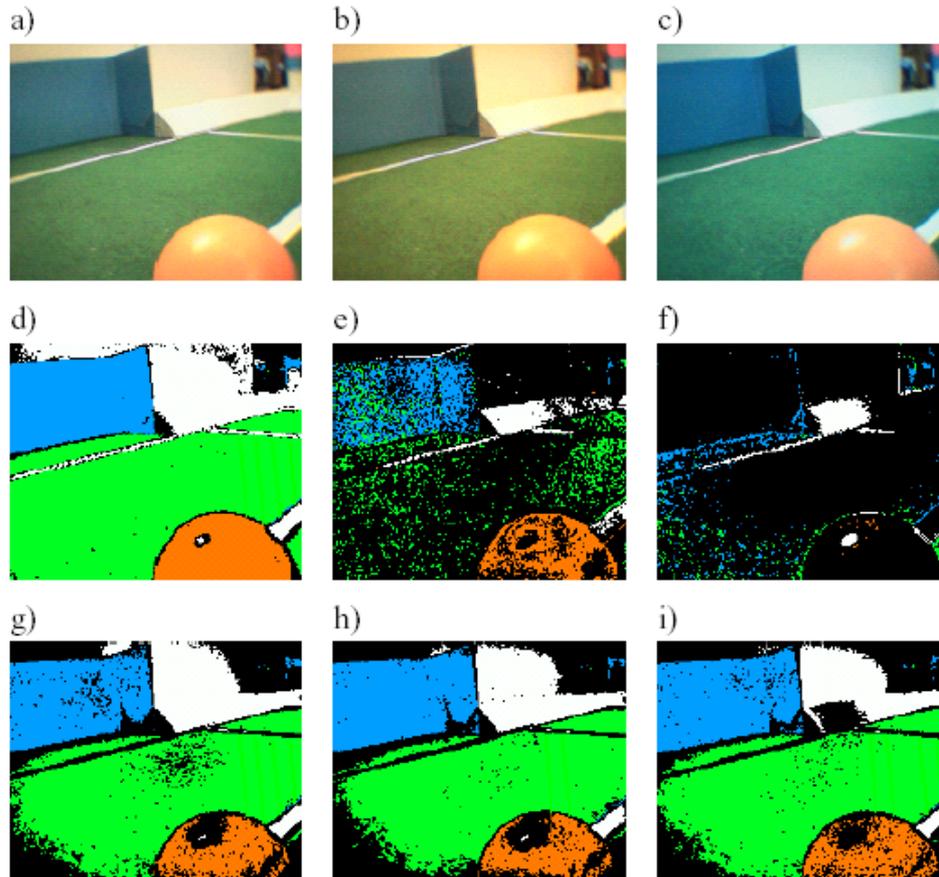
Metamere Farben (Kamera kann anders reagieren als Auge).

Anpassung an Lichtverhältnisse

(a,b,c) unter verschiedenen Bedingungen aufgenommene Bilder

(d,e,f) Farbklassifizierungen mit für (a) erzeugter Farbtabelle

(g,h,i) Farbklassifizierungen mit spezifischen Farbtabelle



Farbkalibrierung

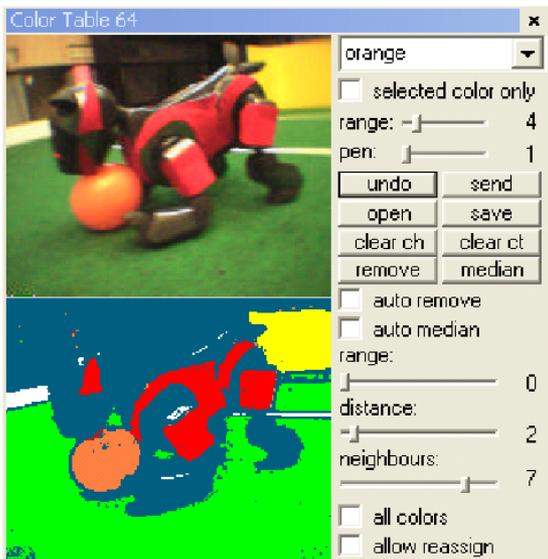
Manuelle Farbkalibrierung (Aibo Team Humboldt)

a) Interface: Bild und Klassifikationsresultat

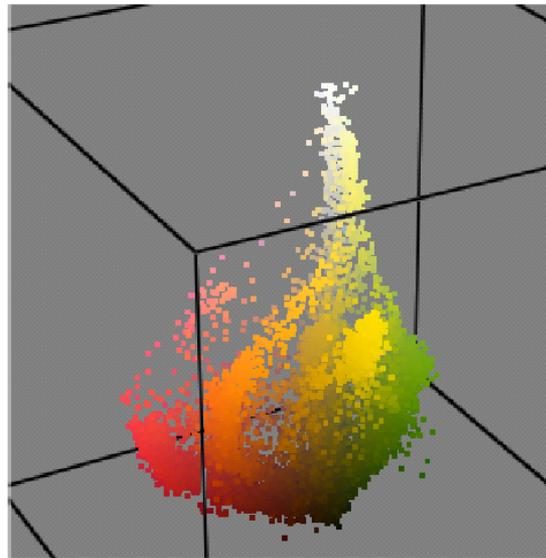
b) Bildraumbelegung des Bildes

c) Bildraumbelegung der Farbtabelle

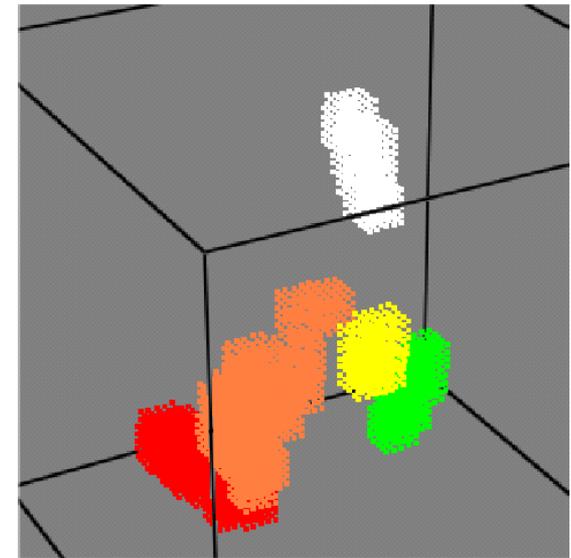
a)



b)



c)



Farbkalibrierung

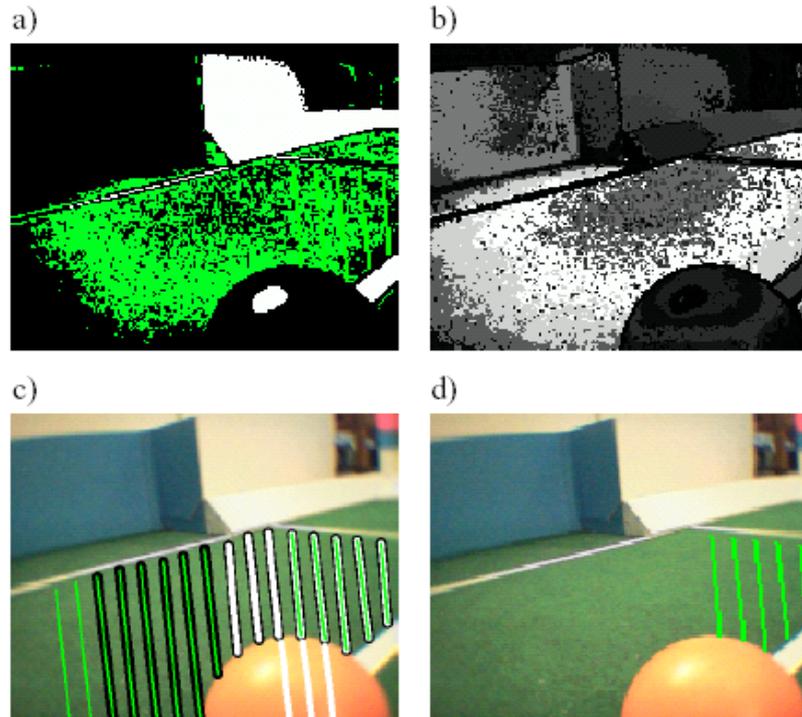
Automatische Farbkalibrierung (Aibo Team Humboldt)

(a) Identifizierte Referenzwerte für „grün“ (vgl. Bild (d))

(b) Häufigkeitsverteilung der Farben.

(c) Scanlinien mit unterschiedlicher Bewertung bzgl. „grün“

(d) Für Kalibrierung benutzte Scanlinien



Farbkalibrierung

Allgemeiner modellbasierter Ansatz:

Modell für Reflexion:

Für jede Wellenlänge λ ist die Reflexion $I(p, \lambda)$ eines Oberflächenpunktes p proportional zu

- Eingestrahlttem Licht auf Oberfläche $E(p, \lambda)$
(Lichtstärke, Orientierung zum Licht)
- Reflektivität der Oberfläche $S(p, \lambda)$
(Anteil, Farbe, Orientierung zum Beobachter)

$$I(p, \lambda) = E(p, \lambda) \cdot S(p, \lambda)$$

Farbkalibrierung

Einfacher Zusammenhang für einzelne Wellenlänge

$$I(p, \lambda) = E(p, \lambda) \cdot S(p, \lambda)$$

Aber: Komplexer Zusammenhang für die Sensoren wegen Integration über alle Wellenlängen

$$e_x(p) = \int \phi_x(\lambda) I(p, \lambda) d\lambda = \int \phi_x(\lambda) E(p, \lambda) S(p, \lambda) d\lambda$$

$$e_y(p) = \int \phi_y(\lambda) I(p, \lambda) d\lambda = \int \phi_y(\lambda) E(p, \lambda) S(p, \lambda) d\lambda$$

$$e_z(p) = \int \phi_z(\lambda) I(p, \lambda) d\lambda = \int \phi_z(\lambda) E(p, \lambda) S(p, \lambda) d\lambda$$

Farbkalibrierung

$$e_x(p) = \int \phi_x(\lambda) I(p, \lambda) d\lambda = \int \phi_x(\lambda) E(p, \lambda) S(p, \lambda) d\lambda$$

$$e_y(p) = \int \phi_y(\lambda) I(p, \lambda) d\lambda = \int \phi_y(\lambda) E(p, \lambda) S(p, \lambda) d\lambda$$

$$e_z(p) = \int \phi_z(\lambda) I(p, \lambda) d\lambda = \int \phi_z(\lambda) E(p, \lambda) S(p, \lambda) d\lambda$$

Für Farbkalibrierung am Farbpunkt p wird benötigt:

Abweichung von $[e_x(p), e_y(p), e_z(p)]$

in Abhängigkeit von (unbekannten) Einstrahlungswerten $E(p, \lambda)$ bei fixierter (aber unbekannter) Reflektivität $S(p, \lambda)$.

Annahmen zur Vereinfachung z.B.:

- auch $S(p, \lambda)$ als Linearkombination von 3 Basisvektoren

Weiteres: Siehe Literatur.

Rückwärtsmodelle (Selbstlokalisierung)

Gesucht:

Extrinsische Parameter:

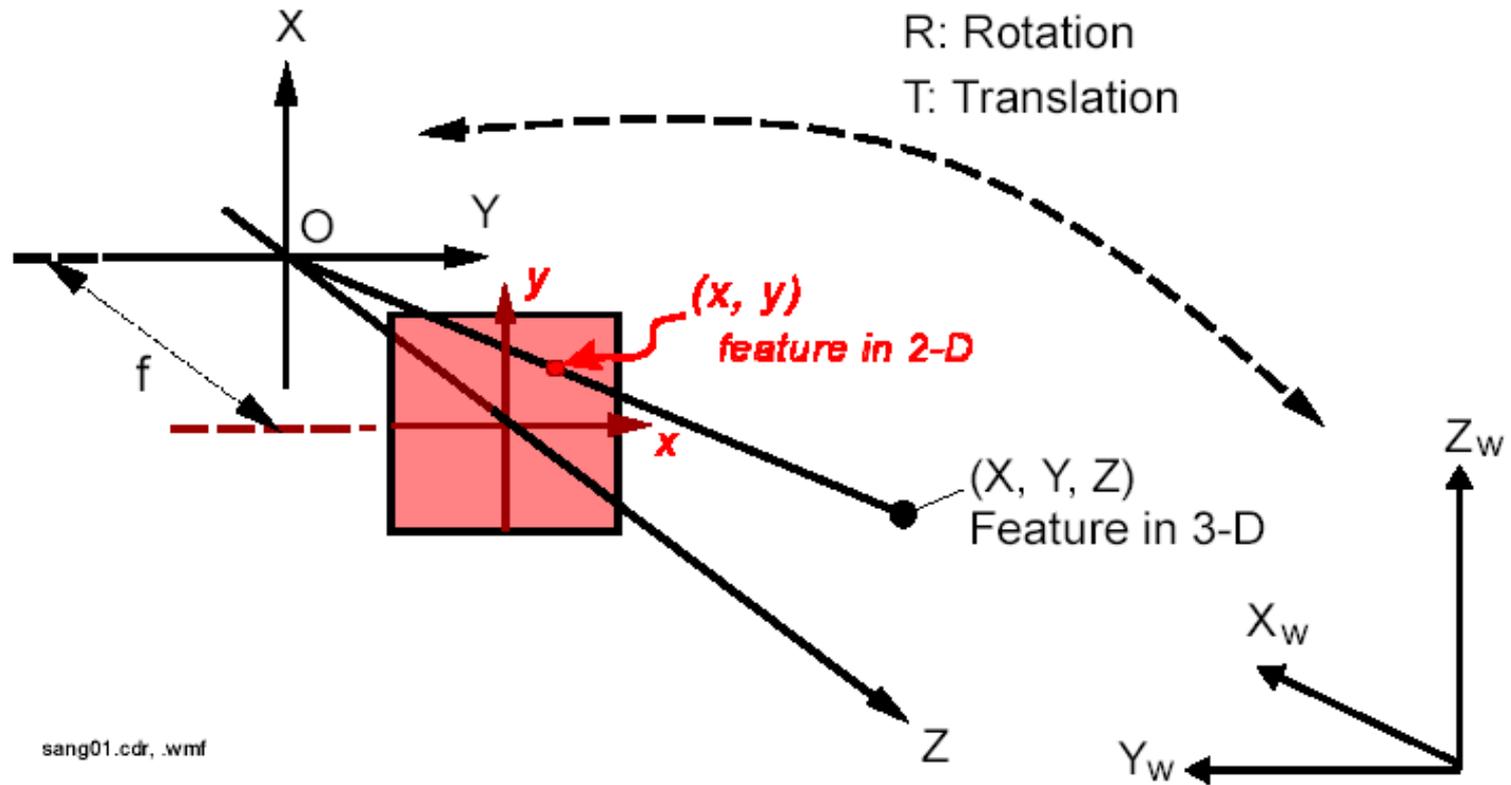
Pose im Raum (bzgl. Weltkoordinaten X_w, Y_w, Z_w)

- Lage des Brennpunkts (3 DOF)
- Orientierung (3 DOF):

Kamerakoordinaten X, Y, Z mit Ursprung im Brennpunkt und Z in Richtung optische Achse

bei bekannten intrinsischen Parametern.

Rückwärtsmodelle (Selbstlokalisierung)



Projektionsabbildung:

$$(x, y, 1)^T = P \cdot (X_w, Y_w, Z_w, 1)^T$$

$$P' = Z \cdot R \cdot T$$

Rückwärtsmodelle (Selbstlokalisierung)

Constraint für mögliche Posen:

Objektpunkt („feature“) mit bekannten Weltkoordinaten und bekannten Bildkoordinaten des zugeordneten Bildpunktes.

Im allgemeinen reichen drei solche Constraints aus für Bestimmung der Pose (evtl. mehrdeutige Lösung).

Bessere Genauigkeit bei mehr Constraints.

Rückwärtsmodelle (Selbstlokalisierung)

Constraint durch

1 Objekt-Punkt X_w mit zugehörigem Bildpunkt:

Legt für jede Pose (6 DOF) im Raum fest,

ob/wie X_w von dort aus in der durch den Bildpunkt vorgegebenen Weise sichtbar ist (für jeden möglichen Kamera-Ort maximal eine Orientierung).

Rückwärtsmodelle (Selbstlokalisierung)

Constraint durch

2 unterscheidbare Objekt-Punkte X_w mit zugehörigen

Bildpunkten:

Der Winkel, in dem die beiden Objekte gesehen werden

(„Peilung“, d.h. Abstand im Bild)

bestimmt eine Oberfläche im Raum (Torus) als mögliche Aufenthaltsorte der Kamera mit entsprechenden Orientierungen.

Rückwärtsmodelle (Selbstlokalisierung)

Weitere Constraints:

- Winkel zwischen Geraden
- Peilung (Winkel, in dem zwei Objekte gesehen werden, entspricht Abstand im Bild)
- Größenverhältnisse,
- ...

Auswertung z.B.

- Constraint-basierte Verfahren
- Minimisierung des quadratischen Fehlers.

Spezielle Bilderzeugung

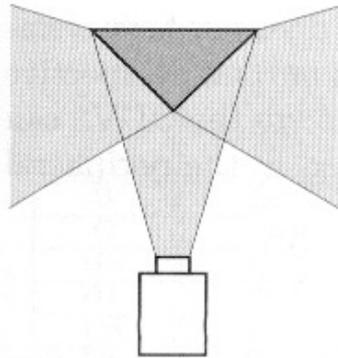
Omnidirektionale Kamera
Spiegel



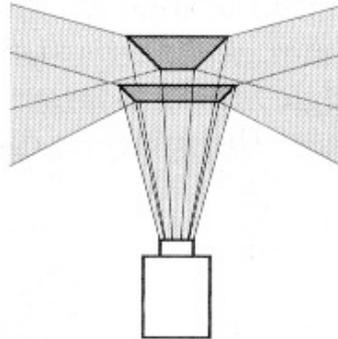
Unterschiedliche Oberflächenkrümmung erlaubt z.B.
zusätzlich vergrößerte Auflösung im Nahbereich

Anpassung der Algorithmen zur Auswertung

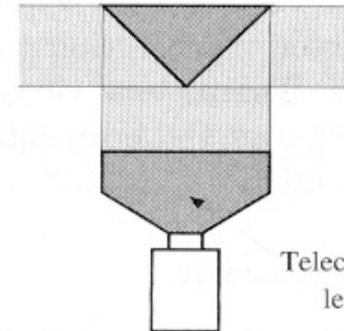
verschiedene Spiegelarten



(a) Conic mirror

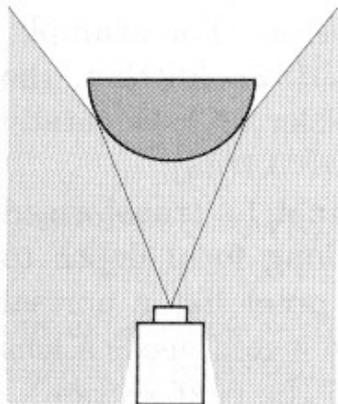


(a2)

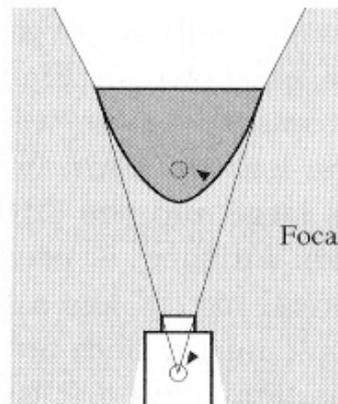


(a3)

Telecentric lens

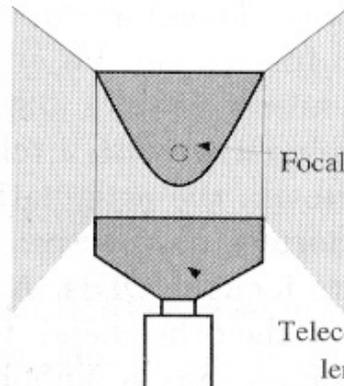


(b) Spherical mirror



(c) Hyperboloidal mirror

Focal point



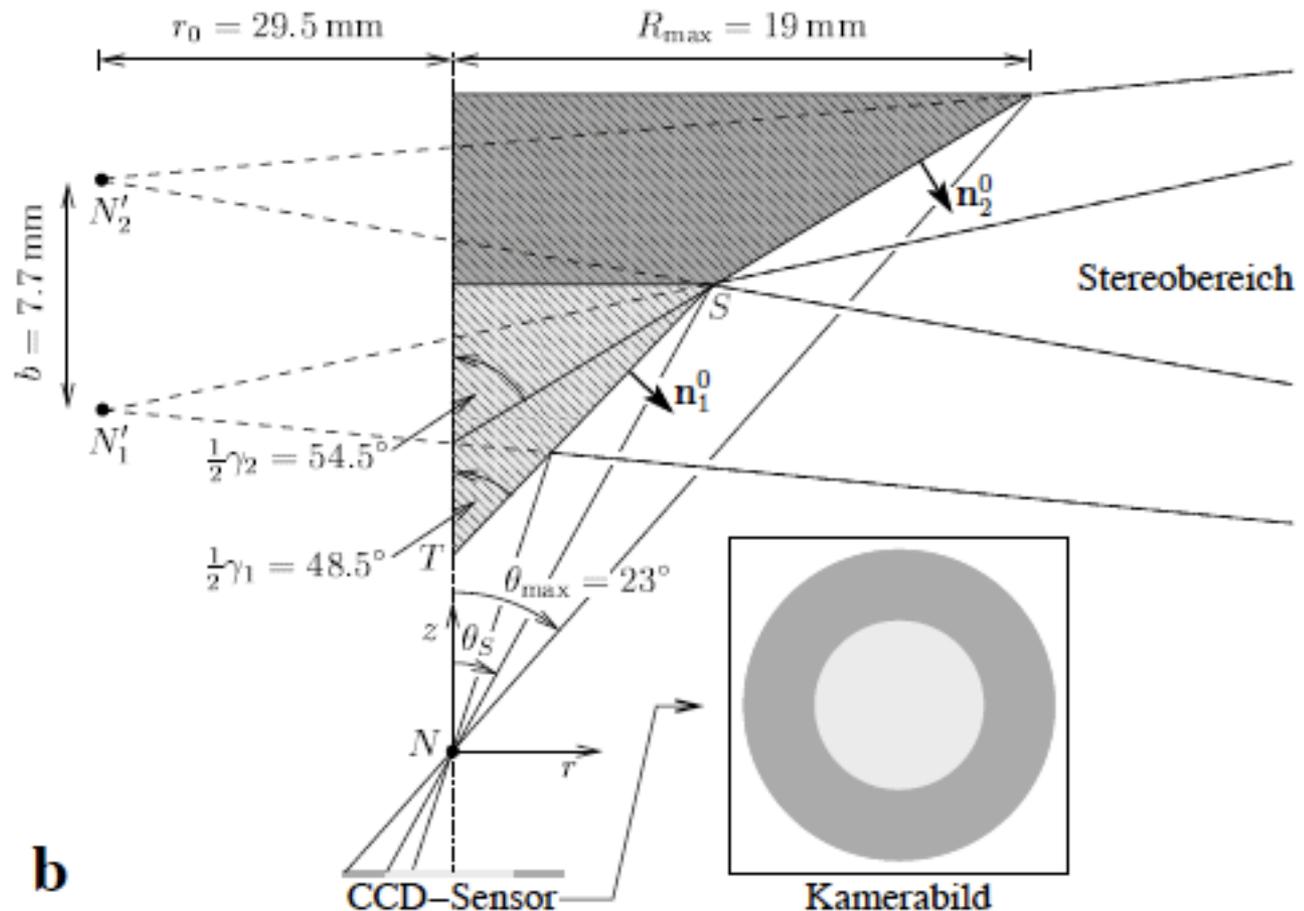
(d) Paraboloidal mirror

Telecentric lens

Omnicam mit 2 Spiegeln



Stereo Omnicam Uni Tübingen



Literatur

- Jähne, Bernd: Digitale Bildverarbeitung. Springer, 2002
- Nehmzow, Ulrich, Peter Haberäcker: Scientific Methods in Mobile Robotics. Springer, 2006
- Burger, Wilhelm, Mark James Burge: Digitale Bildverarbeitung. Springer, 2005
- Haberäcker, Peter: Digitale Bildverarbeitung. Hande 1985.
- Klette, Reinhard, P. Zamperoni: Handbuch der Operatoren für die Bildbearbeitung. Vieweg 1995.
- Mallot, Hans-Peter A.: Sehen und die Verarbeitung visueller Informationen. Vieweg 1998
- Watkins, Christopher, Alberto Sadun, Stephen Marenka: Modern Image Processing. Academic Press 1993.
- Borenstein, Johann (et al.): Where am I . University of