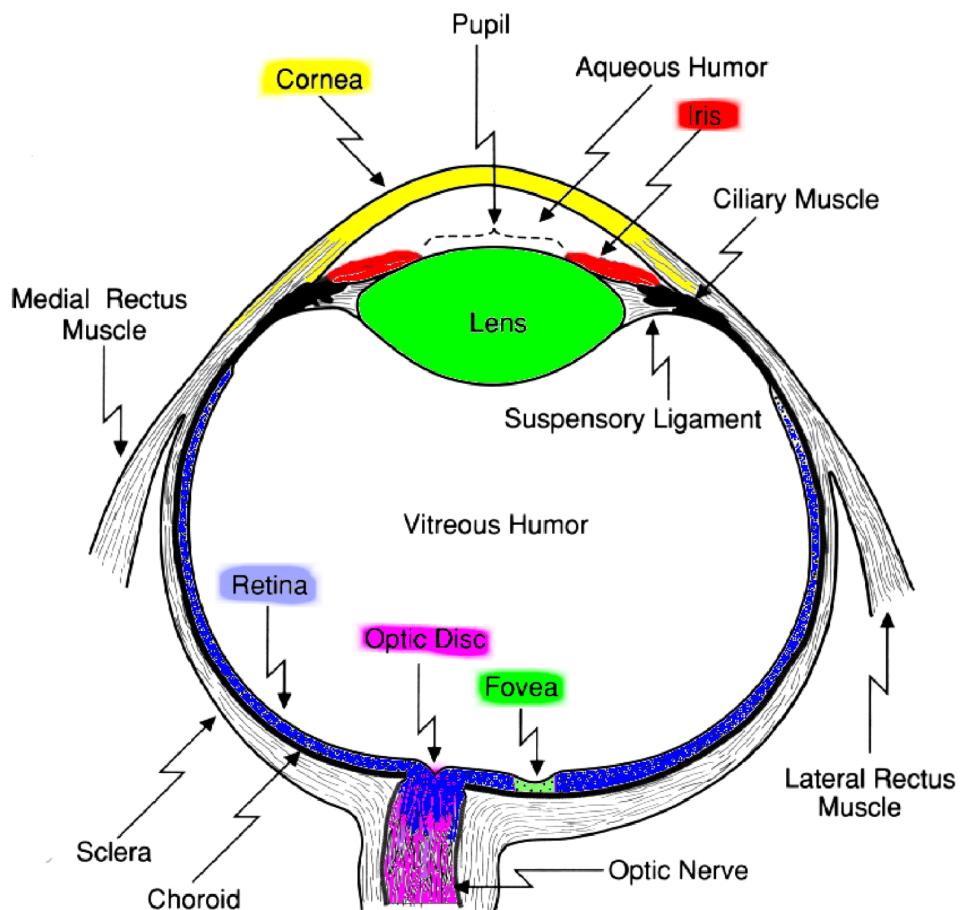


## 2. Bildaufnahme

Quellen:

- [EVC\\_Skriptum\\_CV.p.8](#) bis [EVC\\_Skriptum\\_CV.p.14](#)

## Das Menschliche Auge:



- **Linse** des Auges fokussiert Licht aus der Umgebung
- Bild wird **seitenverkehrt & auf dem Kopf stehend** auf die Retina projiziert
- **Retina (Netzhaut)** = lichtempfindliche Membran auf der Rückseite des Auges

## Fotorezeptoren

- Spezialzellen, die Lichtreize in **neuronale Impulse** umwandeln
- Zwei Typen:
  - **Stäbchen**
    - Verantwortlich für **Schwarz-Weiß-Sehen**
    - Hohe Lichtempfindlichkeit → wichtig für das **Dämmerungssehen**

- **Zäpfchen**
  - Zuständig für **Farbsehen**
  - Funktionieren bei **Tageslicht**
  - Drei Typen für unterschiedliche **Wellenlängenbereiche** (rot, grün, blau)

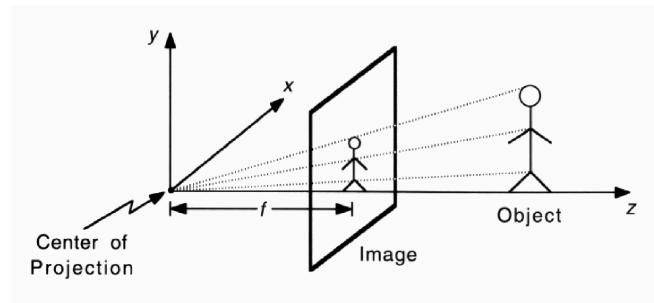
## Signalverarbeitung

- Reaktion der Fotorezeptoren: **chemisch-elektrisch**
- Entstehung eines **neuronalen Impulses**
- Weiterleitung über mehrere **Neuronenschichten**
- Schichten sind durch **Synapsen** miteinander verbunden

## Besondere Strukturen

- **Fovea centralis** (im Zentrum des gelben Flecks / *Macula lutea*):
  - Ort des **schärfsten Sehens**
  - Dichteste Anordnung von Zäpfchen
- **Sehnerv (Nervus opticus)**:
  - Leitet visuelle Impulse ans Gehirn
  - **50 % der Fasern**: Informationen aus der **Fovea**
  - **50 % der Fasern**: Informationen aus der **restlichen Retina**

## Perspektivische Projektion



## Grundprinzip

- Einfachstes Modell der Bildaufnahme ohne Linse, nur mit punktförmiger Blende
- Beschreibt perspektivische Projektion eines 3D-Punktes auf eine 2D-Bildecke
- Idealisiert: keine geometrischen Verzerrungen, keine Unschärfe durch Linsen

## Annahmen (zur Vereinfachung der Herleitung)

- Zentrum der Projektion (Punkt O) = Ursprung des Kamerakoordinatensystems

- Optische Achse (Kamaraachse) ist entlang der z-Achse der Kamera ausgerichtet
- Bildebene liegt vor dem Zentrum der Projektion ( $\rightarrow$  keine Bildumkehrung)

## Aufbau des Modells

- $O$  = Zentrum der Projektion (Punkt, durch den alle Lichtstrahlen verlaufen)
- Bildebene = Fläche, auf die das 3D-Bild projiziert wird
- $f$  = Brennweite (*Abstand von O zur Bildebene*) (*Bei Kameras: "Abstand zwischen Linse und Chip/Film"*)
- Optische Achse = Gerade durch  $O$ , senkrecht zur Bildebene

## Mathematische Beziehung (perspektivische Projektion)

- Gegeben: 3D-Punkt  $P = (X, Y, Z)$
- Projektionsgleichungen:
  - $x = \frac{f}{Z} \cdot X$
  - $y = \frac{f}{Z} \cdot Y$
- $x$  und  $y$ : Position des projizierten Punktes auf der Bildebene

## Eigenschaften der Projektion

- Nicht-linear: Bildgröße  $\propto 1/Z \rightarrow$  je näher ein Objekt, desto größer erscheint es
- 1:n-Abbildung: mehrere 3D-Punkte können auf denselben 2D-Punkt projiziert werden
- 3D-Linien  $\rightarrow$  2D-Linien (außer wenn parallel zur optischen Achse)
- Winkel und Abstände bleiben nicht erhalten
- Parallelle Linien bleiben nur erhalten, wenn sie parallel zur Bildebene liegen
- Informationsverlust: 3D  $\rightarrow$  2D nicht umkehrbar; Tiefeninformation geht verloren

## Praktische Anwendung

- Vereinfachtes Modell zur Beschreibung von Bildaufnahmeprozessen
  - Grundlage in Computervision, Grafik und Kamerakalibrierung
  - Idealisierung realer Kameras zur mathematischen Beschreibung der Abbildungsgeometrie
- 

## Linsen

### Motivation

- Die einfache Lochkamera hat **praktisch keine Bedeutung** in der realen Bildaufnahme

- Für eine **scharfe Projektion** benötigt man ein sehr kleines Loch → lässt kaum Licht durch
- Folge: **lange Belichtungszeiten**, ungeeignet für bewegte Szenen oder praktische Anwendungen

## Lösung: optische Linsen

- Statt Lochblende werden **Linsen oder Linsensysteme** verwendet
- Diese bieten bessere Lichtbündelung und schärfere Abbildung
- Abbildungsverhalten ist komplexer, aber wesentlich **praxisnäher**

## Modell der dünnen Linse

- Vereinfachtes, idealisiertes Modell
- Ersetzt die Lochblende durch eine **symmetrische, unendlich dünne Linse**
- **Lichtstrahlen werden an einer virtuellen Ebene** (in der Linsenmitte) gebrochen
- Abbildungsgeometrie bleibt gleich wie bei der Lochkamera → perspektivische Projektion

## Brennpunktverhalten

- **Parallele Strahlen** zur optischen Achse werden im **Brennpunkt** (Abstand  $f$  zur Linse) gebündelt
- Die **Brennweite  $f$**  ist ein zentraler Parameter der Abbildung
- Fokuspunkt entsteht auf der Bildebene, wenn das Objekt korrekt positioniert ist

## Linsengleichung

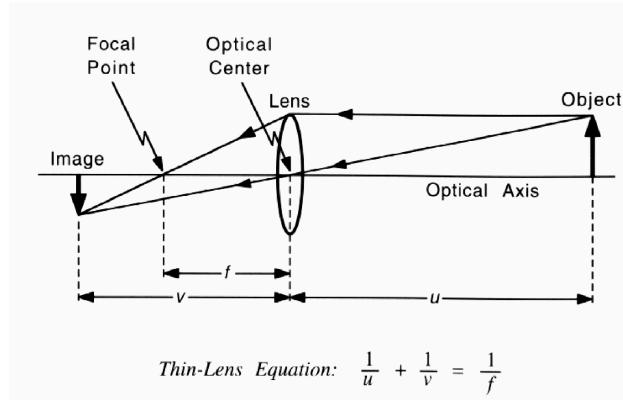


Abbildung 8: Linsengleichung

Liegt ein Objekt in einer endlichen Entfernung  $u$ , werden dessen **Lichtstrahlen** zu einem Bild **hinter dem Brennpunkt** in der Entfernung  $v$  fokussiert. Die zur Linsenachse senkrecht stehende Ebene wird **Bildebene** genannt. Die Beziehung zwischen der **Entfernung  $u$  vom Objekt zur Linse** und der **Entfernung  $v$  von der Linse zur Bildebene** wird durch die **einfache Linsengleichung** beschrieben, wie in Abbildung 8 dargestellt:

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

Diese besagt, dass **Objekte, die sehr weit weg sind ( $u = \infty, v = f$ )**, im Brennpunkt scharf abgebildet sind, alle Objekte, die näher sind, dahinter.

# Tiefenschärfebereich (DOF (Depth of Field))

- Wegen einfacher Linsengleichung immer nur Objekte mit bestimmter Entfernung zur Kamera scharf
- alles andere unscharf
- In Praxis Objekte mit gewissen DOF scharf abgebildet.
- DOF = Entfernung zwischen dem nächsten und weitesten entfernten Objekt einer Szene, welches scharf dargestellt wird

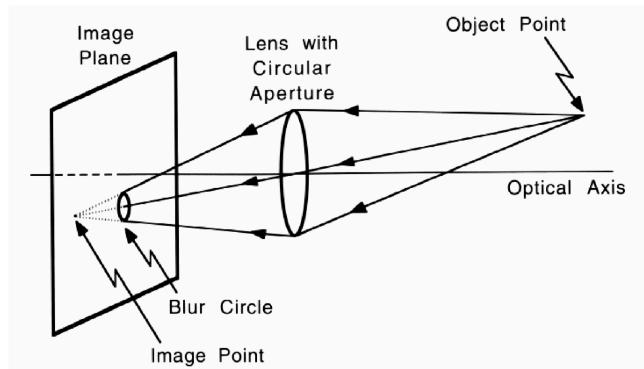
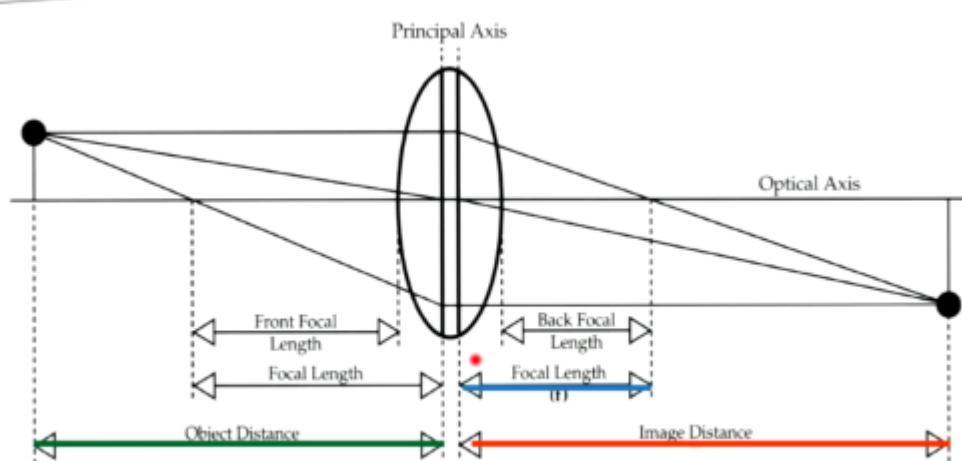


Abbildung 9: Tiefenschärfenbereich



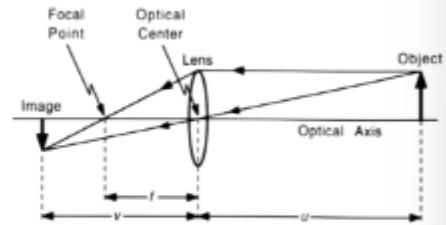
FOCAL LENGTH = Distance from focus point to principal axis.

Test-ähnliches Beispiel:

## Beispiel



- Eine Person steht 5m vor einer Kamera. Die fokale Länge der Linse beträgt 200mm. Wie muss der Bildabstand gewählt werden, damit die Person scharf auf der Bildebene erscheint?



$$\begin{aligned}
 & \bullet u = 5000 \text{ mm} \quad \frac{1}{u} + \frac{1}{v} = \frac{1}{f} \Rightarrow \frac{1}{5000} + \frac{1}{v} = \frac{1}{200} \Rightarrow \frac{1}{v} = \frac{1}{200} - \frac{1}{5000} \Rightarrow \\
 & \bullet f = 200 \text{ mm} \\
 & \bullet v = ? \quad \frac{1}{v} = \frac{24}{5000} \Rightarrow v = \frac{5000}{24} \Rightarrow v = 208,333
 \end{aligned}$$

## Einflussfaktoren

- Sensorauflösung**
- Größe der Linse / Blendenöffnung**
- Brennweite der Linse**

## Blende und Lichtbündel

- Blende = ringförmige Öffnung vor der Linse
- Bestimmt den **Kegelwinkel** des Strahlenbündels, das auf die Bildebene trifft
- Kleine Blende**:
  - Weniger Strahlen erreichen die Bildebene
  - Geringere Lichtmenge → dunkleres Bild
  - Größerer DOF** (mehr vom Bild erscheint scharf)
- Große Blende**:
  - Mehr Strahlen pro Objektpunkt → höhere Lichtausbeute
  - Kleinerer DOF** (weniger Tiefenbereich scharf)

## Unschärfekreise (Blur Circles)

- Wenn Objekt nicht im richtigen Abstand  $v$  zur Linse steht → Strahlen bündeln sich nicht exakt
- Statt eines Punktes entsteht ein **Unschärfekreis** auf der Bildebene
- Benachbarte Objektpunkte ergeben überlappende Kreise → **Bild wird unscharf**

## Zusammenhang: Blende – Blur – DOF

- Größere Linse** = größere Unschärfekreise = **kleinerer DOF**

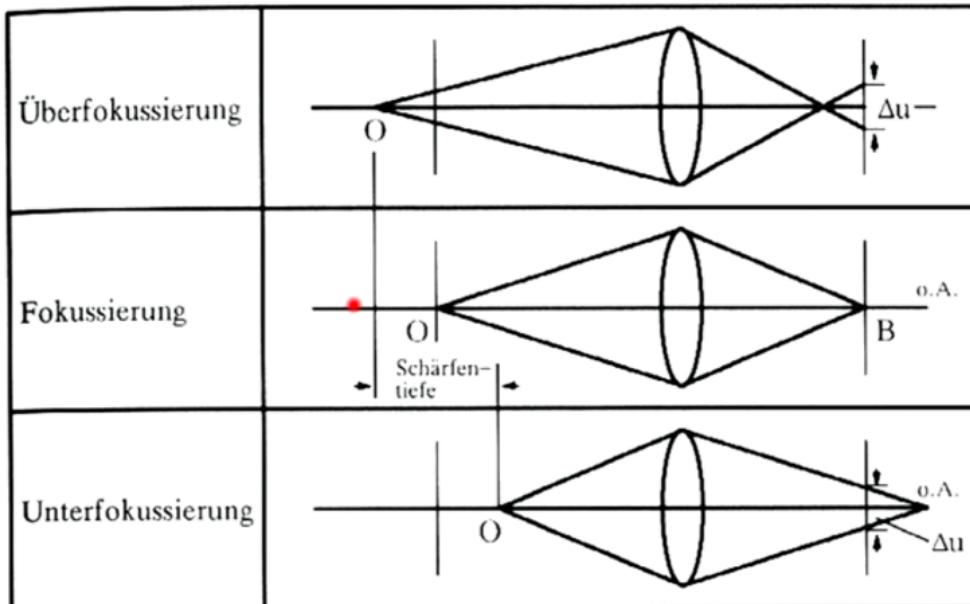
- Kleinere Linse = kleinere Unschärfekreise = **größerer DOF**

## Einfluss der Sensorsauflösung

- **Höhere Auflösung:** Unschärfekreise können besser abgebildet werden → **kleinerer DOF**
- Geringere Auflösung: Unschärfen weniger deutlich → scheinbar größerer DOF

## Einfluss der Brennweite

- Längere Brennweite (z. B. Teleobjektiv) → **kleinerer DOF**
- Kürzere Brennweite (z. B. Weitwinkelobjektiv) → **größerer DOF**



## Radiometrie

### Grundlagen

- Radiometrie = **Messung elektromagnetischer Strahlung**, inkl. sichtbarem Licht
- Ziel: Beschreibung der Lichtmenge, die von der realen Welt auf ein Bild projiziert wird

### Strahldichte (Radiance)

- Beschreibt die **Menge des reflektierten Lichtes** eines Oberflächenpunktes
- Richtungsabhängige Größe
- Gibt an, wie viel Licht in eine bestimmte Richtung von einem Punkt abgestrahlt wird
- Entspricht der **Lichtmenge**, die durch ein **kleines Raumwinkel-Element** ausgesendet wird

# Bestrahlungsstärke (Irradiance)

- Gibt an, wie viel **Licht** von einer Oberfläche **empfangen** wird
- Entspricht der Helligkeit eines Bildpunktes
- Wird aus der Strahldichte berechnet, indem über alle Richtungen integriert wird

## Spektrum

- Radiometrische Größen beziehen sich auf **Strahlung aller Frequenzen**
- Betrachtet man **einzelne Frequenzbereiche**, spricht man vom **Spektrum**
- Wichtige spektrale Größe:
  - **Spektrale Bestrahlungsstärke (spectral irradiance)**
  - Gibt an, wie viel **Strahlungsenergie pro Wellenlänge** auf eine Fläche trifft

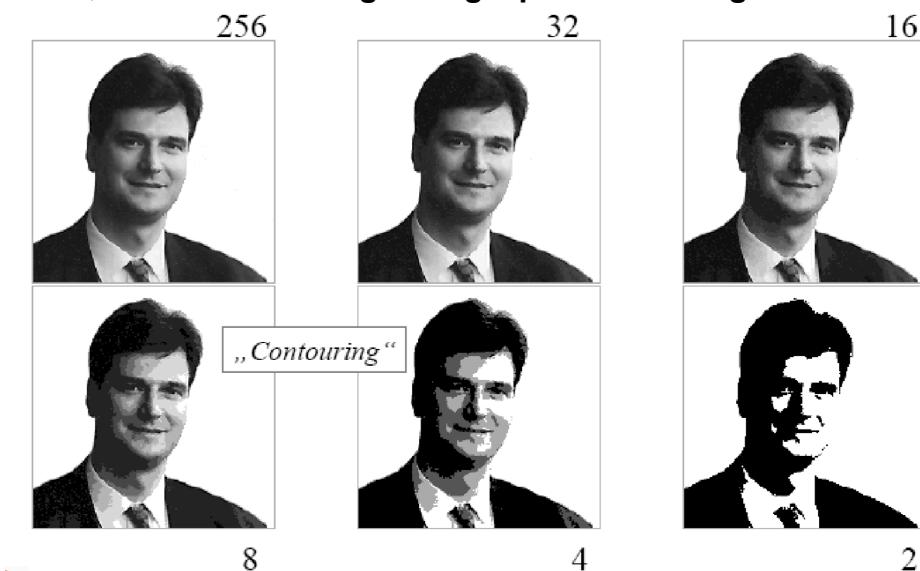
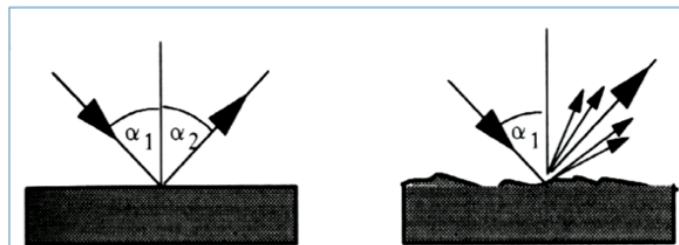


Abbildung 10: Einfluss der Bitanzahl auf die Bildqualität

The radiometric relation between the world and its projection is formed by:

- Amount of light that is reflected by a surface point = **Radiance**
- Amount of light that is projected from this point onto the image = **Irradiance**
- measured in watts per square meter ( $W/m^2$ ),



## Sampling

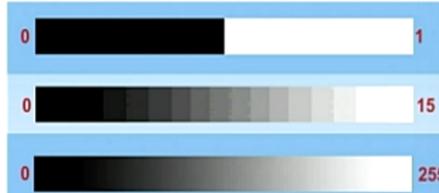
### Bilddauflösung – Überblick

- **Bildauflösung** = Maß für den **Detailreichtum** eines Bildes
- Abhängig von:
  - **Radiometrischer Auflösung**
  - **Sensorauflösung**
  - **Räumlicher (geometrischer) Auflösung**
  - **Zeitlicher Auflösung**

## Radiometrische Auflösung

- Gibt an, wie fein ein System **Helligkeitsunterschiede unterscheiden** kann
- Beschrieben durch **Anzahl der Bits** (z. B. 8 Bit = 256 Graustufen)
- Mensch kann ~120 Grauwerte unterscheiden → 256 Graustufen genügen

$$2^1 = 2 \text{ levels (0,1)}$$



$$2^2 = 4 \text{ levels (0,1,2,3)}$$

$$2^8 = 256 \text{ levels (0-255)}$$

$$2^{12} = 4096 \text{ levels (0-4095)}$$

- Weniger Graustufen → geringerer Speicherbedarf, aber:
  - **Qualitätsverlust**
  - **Contouring-Effekte** (sichtbare Tonwertsprünge)  
(siehe hier:

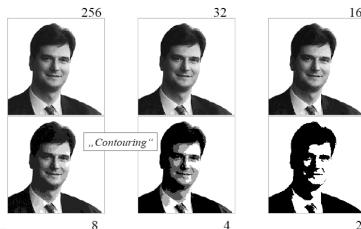


Abbildung 10: Einfluss der Bitanzahl auf die Bildqualität

)

## Sensorauflösung

- Anzahl der **Pixel (Bildpunkte)** eines digitalen Bildes
  - z. B.  $640 \times 480 = 307.200$  Pixel = ~0.3 MP
- Häufige Angabe: **Megapixel**
- Alternativ: **dpi / ppi** (dots/pixels per inch) – sinnvoll nur bei Scannern
- Für Kameras ungeeignet, da Objekt-Sensor-Abstand variabel

## Räumliche (geometrische) Auflösung

- Fähigkeit, **benachbarte Objektstrukturen getrennt** zu erfassen
- Maßeinheiten:
  - **Linien pro Millimeter (L/mm)** in Optik/Photographie
  - Definiert durch:

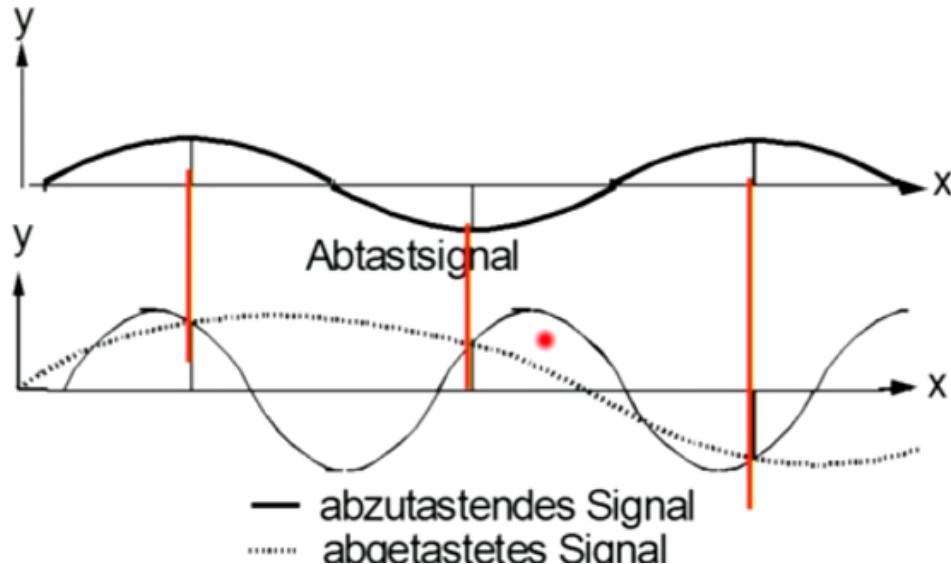
- **minimale Entfernung** zwischen zwei erkennbaren Merkmalen
- **minimale Größe** eines erkennbaren Merkmals

## Zeitliche Auflösung

- Anzahl der aufgenommenen Bilder **pro Sekunde**
  - z. B. 25 fps (frames per second)
- Wichtig bei Videoaufnahmen oder Bewegungsanalysen

## Sampling (Abtastung)

- **Umwandlung** eines kontinuierlichen Signals (z. B. in Raum/Zeit) in eine **diskrete Folge**
- Zentrale Grundlage: **Nyquist–Shannon Sampling Theorem**
  - Ein bandbegrenztes Signal kann exakt rekonstruiert werden, wenn:
    - **Samplingrate > 2 · B** ( $B$  = höchste Frequenz im Signal)
    - d. h.  $f_{abtast} > 2 \cdot f_{maxSignal}$



- Anwendung bei **analoger/digitaler Wandlung** in digitalen Kameras:
    - Raum (Pixelraster)
    - Zeit (fps)
    - Spektrum (z. B. Farbbandsampling)
- mehr Beispiele zu Nyquist: [7. Clipping und Antialiasing](#)

## Bildsensoren

### Allgemeines

- Ein **Bildsensor** wandelt ein **optisches Bild** in ein **elektronisches Signal** um
- Es gibt zwei Arten digitaler Pixelsensoren:

- **CCD (Charge-Coupled Device)**
- **CMOS (Complementary Metal-Oxide-Semiconductor)**

## CCD-Sensoren

- Bestehen aus einer **Matrix lichtempfindlicher Fotodioden**
  - Zusätzlich enthalten: **MOS-Kondensatoren, Feldeffekttransistoren, Steuerleitungen, Leitungspfade**
- **Belichtungszeit:** Licht erzeugt elektrische Ladung in der Fotodiode
- Ladung wird im Kondensator gespeichert (parallel zur Fotodiode geschaltet)
- Auslesung erfolgt **zeilenweise**
- Spannung wird in der Kamera oder extern in **digitale Information** umgewandelt
- Arbeitet **linear**, wie analoge Kameras → basiert auf **Flächenintegral-Prinzip**

## CMOS-Sensoren

- Besitzen **aktive Pixelsensoren**, jedes Pixel ist individuell auslesbar
- Spannung, die der **Helligkeit** entspricht, liegt **ständig an**
- Output ist **nicht-linear (logarithmisch)**
  - Ähnelt dem Prinzip der **menschlichen Bildwahrnehmung**
- Vorteile gegenüber CCD:
  - **Niedriger Energieverbrauch**
  - **Schnellere Lesegeschwindigkeit**
  - **Günstigere Herstellung**

## Farbe auf CCD-Sensoren

### Allgemeines

- **Bildsensoren** erfassen **nur Helligkeit**, nicht direkt Farbe
- Um **Farbinformation** zu erhalten, werden **Farbfilter** eingesetzt
- Drei gängige Methoden zur Farberfassung:
  - **Field Sequential Technik**
  - **3-Chip-Technik**
  - **Color Filter Array (CFA)**

### Field Sequential Technik

- Ursprünglich 1900er Jahre, von **Produkin Gorski** entwickelt

- Drei **Schwarz-Weiß-Bilder** nacheinander aufgenommen – je mit **Rot-, Grün- oder Blaufilter**
- Darstellung durch **Lichtprojektion** (Laterna Magica)
- Nachteil: **Geisterbilder** bei Bewegung, da jede Farbe zu einem **anderen Zeitpunkt** aufgenommen wird
- Heute noch bei **Mikroskopen** in Verwendung

### Sequential Color Three-Pass CCD Imaging System

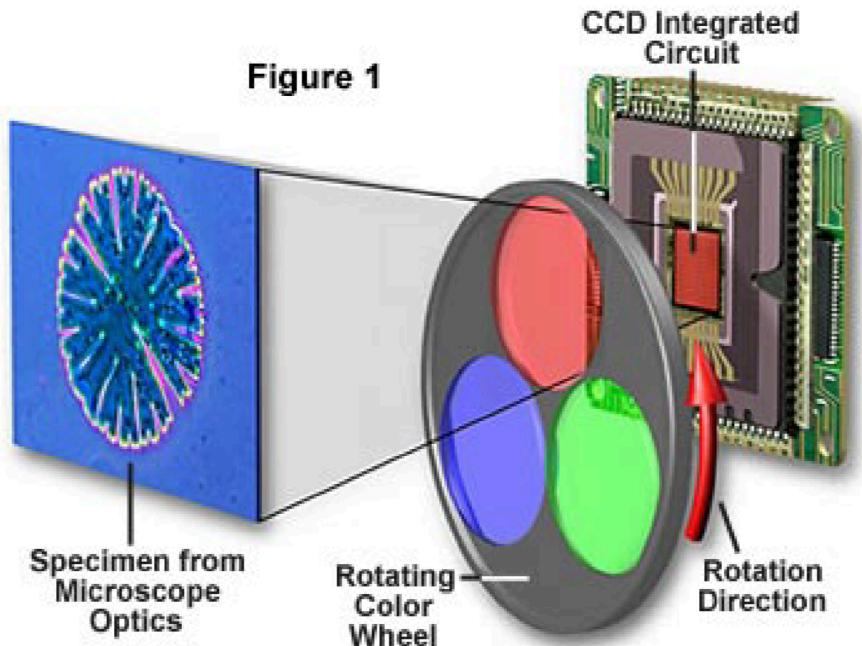


Abbildung 11: Farbaufnahme mittels Color Wheel

## 3-Chip-Technik

- **Prisma** trennt Licht in **Rot, Grün, Blau** (RGB)
- Jeder Farbkanal wird von **eigenem CCD-Sensor** aufgenommen
- Exakte Ausrichtung nötig (Der Lichtstrahl muss bei allen 3 Filtern den selben Pixel treffen), um **Farbtreue** zu garantieren
- nur eingesetzt bei **Videokameras** und **Camcordern**

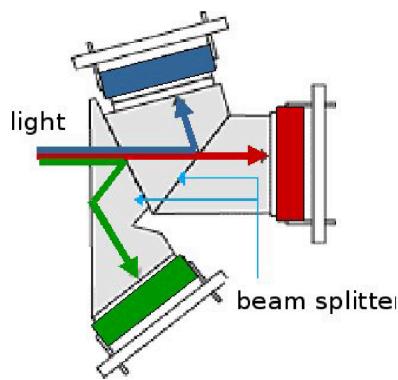


Abbildung 12: Farbtrennung durch Prisma

## Color Filter Array (CFA)

- Häufigste Methode bei **einzelnen CCD- oder CMOS-Sensoren**
- Über jedem Pixel liegt ein **Farbfilter** (Mosaikstruktur)
- **Bayer-Filter** ist Standard:
  - 50 % Grün
  - 25 % Rot
  - 25 % Blau
  - Anpassung an **menschliche Farbwahrnehmung** (grünempfindlichster Bereich)

## Bayer-Muster und Demosaicing

- Output eines Bayer-Sensors = **Rohdatenbild** (Bayer-Muster)
- Jedes Pixel enthält **nur eine Farbkomponente**
- **Demosaicing:** Interpolation der fehlenden Farben
  - Nutzt Annahme: **Geringe Unterschiede** zwischen benachbarten Farbwerten
  - Einfache Verfahren: **Durchschnittswerte** benachbarter Pixel gleicher Farbe
  - Erweiterte Verfahren: Interpolation **entlang von Kanten**
  - Farbverhältnisse (z.B. Rot-Grün) bleiben konstant → erst Grün interpolieren, dann Rot und Blau

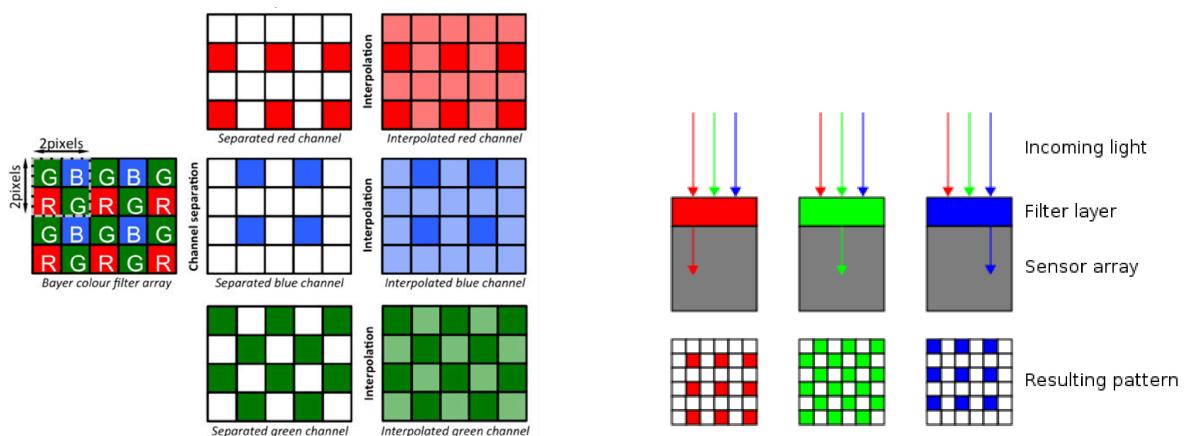


Abbildung 13: Prinzip des Bayer-Filters (links) und Ausführung mittels Filter pro Pixelsensor (rechts).

## Auflösung und Verarbeitung

- Effektive **geometrische Auflösung** des Sensors beträgt **nur ein Drittel**
- Demosaicing kann erfolgen:
  - **In der Kamera** (z.B. JPEG, TIFF-Export)
  - **Extern** durch Programme wie **nomacs** ([www.nomacs.org](http://www.nomacs.org))

## Foveon X3 Sensor

- Der **Foveon X3 Sensor** kombiniert die **3CCD-** und **CFA-Technik**
- Er nutzt ein **Array** von Fotodioden, wobei jede **vertikal gestapelt** ist und in einem **zweidimensionalen Raster** angeordnet wird
- Jede Fotodiode ist auf **verschiedene Wellenlängen** des Lichts empfindlich, was bedeutet, dass jede Diode eine **eigene spektrale Sensitivitätskurve** hat
- Das Signal der drei Fotodioden wird verarbeitet, um **drei additive Primärfarben** (Rot, Grün, Blau) zu erzeugen und so ein **fehlerfreies Farbbild** darzustellen

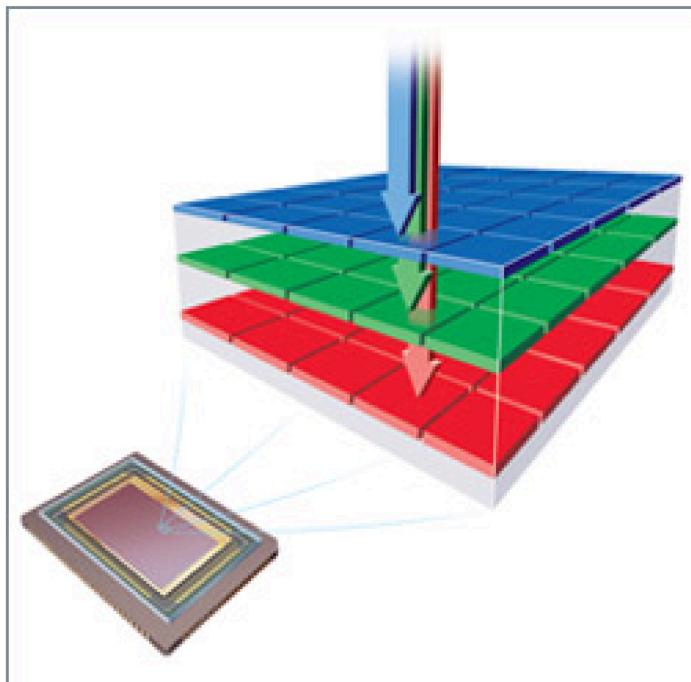


Abbildung 14: Foveon X3 Sensor

## Multi-Shot-Technologie

- Eine **weitere Möglichkeit** zur Erzeugung von **Echtfarbbildern** ist der Einsatz von **beweglichen Sensoren**
- Bei der **Multi-Shot-Technologie** wird der Sensor mithilfe eines **Piezomotors** präzise in **Einpixel-Schritten** bewegt
- Diese Technik ermöglicht es der Kamera, **mehr Farbinformationen** und **Daten** zu erfassen, als bei einer Einzelbildaufnahme
- Im **4-Shot-Modus** werden **vier Aufnahmen** gemacht:

- **Sensorbewegung** erfolgt in vier Schritten (1 Pixel horizontal, 1 Pixel vertikal, erneut 1 Pixel horizontal und 1 Pixel vertikal)
    - Dieser Zyklus sorgt dafür, dass für jedes Pixel **echte RGB-Daten** erfasst werden
  - Nach Beendigung des Zyklus kehrt der Sensor in seine **Ausgangsstellung** zurück

# Ideale vs. Reale Kamera

Reale Kameras weichen von idealen Kameramodellen ab, da sie Linsen mit optischen Phänomenen und nicht-ideale Bildsensoren mit Chipartefakten verwenden. Einige dieser Störungen sind:

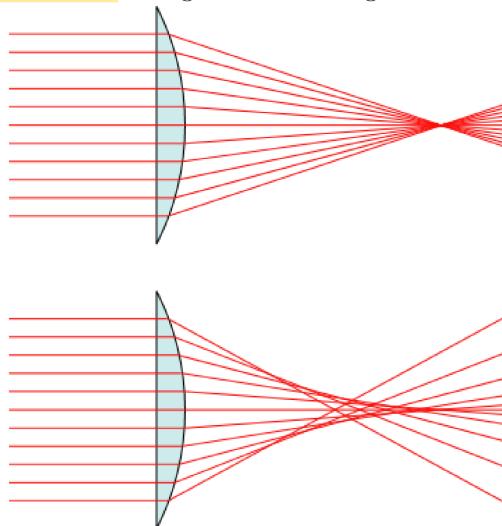


Abbildung 15: Optische Aberration

- Die *optische Aberration* tritt auf, wenn Licht von einem Punkt auf einem Objekt nicht in einem einzigen Punkt nach der Transmission durch die Linse zusammenfällt.
  - Die *Chromatische Aberration* ist eine Art von Störung, bei der die Linse nicht im Stande ist, alle Farben im gleichen Konvergenzpunkt zu fokussieren. Diese Aberration tritt auf, da Linsen unterschiedliche Brechungsindizes für verschiedene Wellenlängen des Lichtes aufweisen. Chromatische Aberration manifestiert sich durch "Farbränder/-säume entlang der Abgrenzung zwischen dunklen und hellen Bildteilen.

- Die *sphärische Aberration* tritt auf, da Kugeloberflächen nicht die ideale Form für Linsen sind, aber beim Schleifen von Glas die einfachste Form darstellen. Diese Aberration führt dazu, dass Strahlen, die weiter am Rand der Linse auftreffen in einem etwas anderen Punkt fokussiert werden, als Strahlen die nahe der Achse liegen.
  - Die *Linsenverzerrung* ist eine Abweichung einer geradlinigen Projektion, gerade Linien in einer Szene bleiben nicht gerade im Bild, diese können nach außen (tonnenförmig) oder nach innen (kissenförmig) verzerrt sein.

# Weißabgleich

- **Jede Kamera** geht bei der Aufnahme eines Fotos zunächst von einem **Durchschnittsbild** aus
  - Die Kamera versucht, ein Bild zu erstellen, das dem **Durchschnitt** in Bezug auf Belichtung, Farbigkeit und mehr entspricht – was nicht immer korrekt ist

- **Weißabgleich** hilft, die **Farbstimmung** oder **Farbtemperatur** anzupassen
- Der Weißabgleich sagt der Kamera, welche Farbe als **weiß** bzw. als **18%-iges neutrales Grau** angesehen werden soll
- **Farbtemperatur** wird in **Kelvin** gemessen, z.B. hat **Licht von Leuchtstoffröhren** ca. 4000 Kelvin, **normale Glühbirnen** etwa 3200 Kelvin und im **Schatten** um 7000 Kelvin
- **Automatischer Weißabgleich:**
  - Die Kamera sucht eine neutrale Fläche und passt alle Farben an diese an
  - Dies funktioniert nicht immer perfekt, daher gibt es **verschiedene Anpassungsmöglichkeiten**

## Optionen für den Weißabgleich

- Der **Weißabgleich** ist meist auf **Automatik** eingestellt
- **Voreingestellte Lichtsituationen** wie z.B. **Kunstlicht, Sonne, Schnee** oder **Schatten** stehen zur Auswahl
- **Manueller Weißabgleich:**
  - Ein weißes Blatt oder eine **18% Graukarte** wird fotografiert
  - Dieses Bild wird als **Referenz** verwendet, um den Weißabgleich korrekt zu erzeugen
- **Software:** Der Weißabgleich kann auch **nachträglich** angepasst werden
  - Hierbei wird ein **weißer Bereich** im Bild ausgewählt, und die Software berechnet den Weißabgleich
  - Beim **RAW-Format** funktioniert dies besonders gut, da der in der Kamera vorgenommene Weißabgleich nur als Vorschlag gespeichert wird und nachträglich geändert werden kann

No white balance



White balance



## 2. Bildaufnahme