

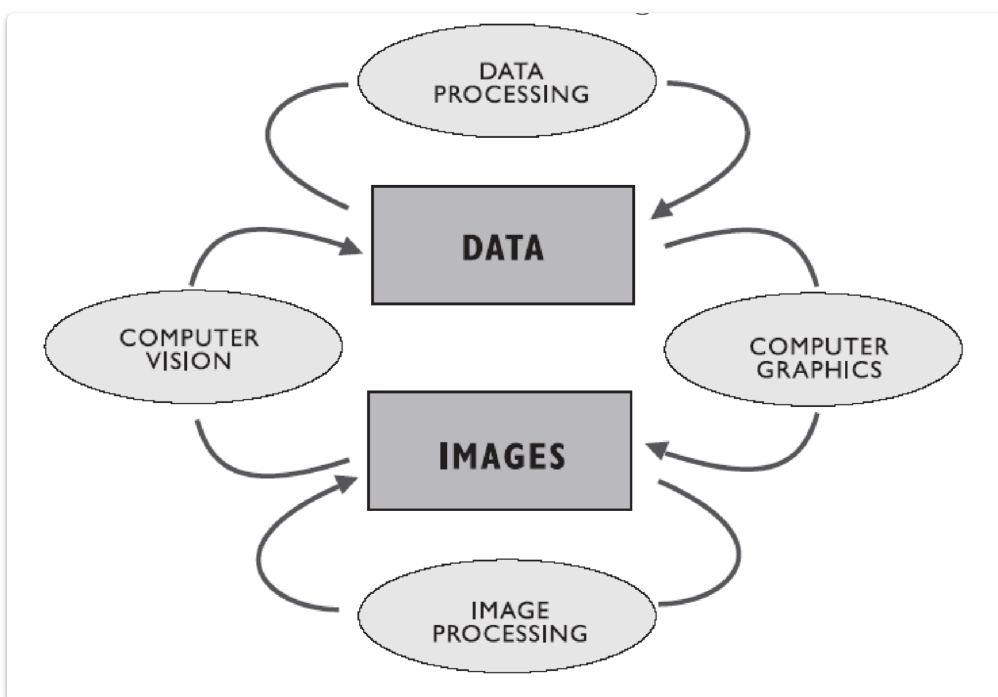
# 1. Einführung in Computer Vision

Quellen:

- EVC\_Skriptum\_CV, p.4
- EVC\_Skriptum\_CV, p.5
- EVC\_Skriptum\_CV, p.6
- EVC\_Skriptum\_CV, p.7

## Die drei Teilbereiche

- **Computergrafik:**
  - Erzeugt Bilder aus Daten (z.B. Diagramme, 3D-Modelle)
  - Ziel: Realistische Bildgenerierung.
- **Computer Vision (Maschinelles Sehen):**
  - Extrahiert Semantik aus Bildern (also genau solche Diagramme / Modelle).
  - Ziel: Nachbildung des Sehvorgangs, Erkennung von Objekten, Bewegung, etc.
  - Teilbereich: Mustererkennung (z.B. optische Zeichenerkennung).
- **Bildverarbeitung:**
  - Verbessert Bilder für bestimmte Aufgaben (z.B. Rauschunterdrückung, Qualitätsverbesserung).
  - Ergebnis: Ein verändertes Bild.
  - Unterschied zur Bildbearbeitung: Bildverarbeitung ist algorithmisch, Bildbearbeitung ist interaktiv.



# Definition Computer Vision

- Automatische Ableitung von Struktur und Eigenschaften einer 3D-Welt aus 2D-Bildern.
- Ziel: Computer sollen visuelle Daten wahrnehmen, verarbeiten und verstehen wie Menschen.
- Anwendungsbereiche:
  - Industrielle Bildverarbeitung (z.B. Flaschenerkennung).
  - Künstliche Intelligenz (z.B. autonome Roboter).
  - Gesichtserkennung.
- Computer Vision vs. Machine Vision: Computer Vision ist die Kerntechnologie zur Bildanalyse, Machine Vision kombiniert die Bildanalyse mit anderen Technologien.
- Teilgebiete: Szenenrekonstruktion, Objekterkennung, Video-Tracking, etc.
- Beispiel Gesichtserkennung:

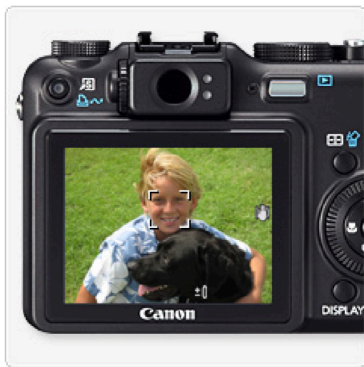


Abbildung 2: Gesichtserkennung

## 3. Menschliches Sehen vs. Computer Vision

- Computer Vision versucht, menschliches Sehen zu reproduzieren, aber:
  - Menschliches Sehen ist Ergebnis von Millionen Jahren Evolution
  - Menschliches Sehen ist fehlerhaft (optische Täuschungen).
  - Wenig Wissen über die genauen Prozesse des menschlichen Sehens.
  - Die Frage besteht, ob es erstrebenswert ist, das Maschinen genauso sehen wie Menschen.
- Das menschliche Sehen ist nicht unfehlbar. Optische Täuschungen und Mehrdeutigkeiten beweisen dies.
- In CV Herleitung durch geometrischen Eigenschaften, Materialeigenschaften und Beleuchtung
- **Begrenztes Wissen:**
  - Unser Verständnis der komplexen Prozesse im Gehirn, die nach der Reizaufnahme durch das Auge ablaufen, ist begrenzt.

- **Fehlerhaftigkeit des menschlichen Sehens:**
  - Optische Täuschungen, Mehrdeutigkeiten und Inkonsistenzen zeigen, dass unser Sehen nicht perfekt ist.
- **Subjektivität der Wahrnehmung:**
  - Die Drehung eines Bildes kann unsere Wahrnehmung komplett verändern.
- **Helmholtz' Theorie:**
  - Wir sehen, was wir erwarten oder was unter normalen Bedingungen zu erwarten wäre.
  - Unser Gehirn konstruiert aktiv unsere visuelle Realität.
- **Toleranzunterschiede:**
  - Wir akzeptieren die Unvollkommenheit des menschlichen Sehens, sind aber bei Maschinen weniger tolerant.
- **Die Frage:**
  - Wollen wir, dass Maschinen genauso "sehen" wie wir, mit all den damit verbundenen Fehlern und Verzerrungen?

## 2. Optische Täuschungen und ihre Bedeutung

- **Einblicke in die Natur des Sehens:**
  - Optische Täuschungen geben uns wertvolle Einblicke in die Funktionsweise unseres visuellen Systems.
- **Kontrollierte Halluzination?:**
  - Die Frage, ob unser Sehen eher einer "kontrollierten Halluzination" gleicht, wird aufgeworfen.
- **Zuverlässigkeit der Wahrnehmung:**
  - Die Täuschungen verdeutlichen, dass unser Vertrauen in unsere visuellen Fähigkeiten nicht immer gerechtfertigt ist.

## 3. Konsequenzen für die Computer Vision

- **Nachbildung des menschlichen Sehens:**
  - Die Schwierigkeiten bei der Nachbildung des menschlichen Sehens werden deutlich.
- **Erwartungen an Maschinen:**
  - Wir müssen uns fragen, welche Art von "Sehen" wir von Maschinen erwarten.
- **Grundlagenforschung:**
  - Die Erforschung optischer Täuschungen hilft, die Grundlagen des menschlichen Sehens besser zu verstehen, was wiederum in der Entwicklung von Computer Vision Systemen hilfreich ist.

## 4. Weitere Informationen:

- Weitere Infos zu wie wir sehen findet man hier:
  - 2. Bildaufnahme

- 4. Farbe

## 4. Plenoptische Funktion

- Beschreibt die Menge aller Lichtstrahlen in einem Raum.
- Parameter zum Beschreiben der Beleuchtung:
  - Position ( $V_x, V_y, V_z$ ).
  - Einfallswinkel ( $\theta, \phi$ ).
  - Wellenlänge ( $\lambda$ ).
  - Zeit ( $t$ ).

Wir wollen nun die Parameter zur Beschreibung beleuchteter Umgebungen betrachten. Im ersten Schritt nehmen wir eine Schwarzweissfotografie einer Lochkamera an, die uns darüber Auskunft gibt, wie groß die Intensität des Lichtes von einem einzigen Blickpunkt aus zu einem bestimmten Zeitpunkt, gemittelt über die Wellenlängen des Spektrums des sichtbaren Lichtes, ist. Diese Lichtintensitätsverteilung  $P$  kann beim Durchgang der Strahlenbündel durch die Linse gemessen und anhand von Kugelkoordinaten  $P(\theta, \phi)$  oder kartesischen Koordinaten  $P(x, y)$  parametrisiert werden.

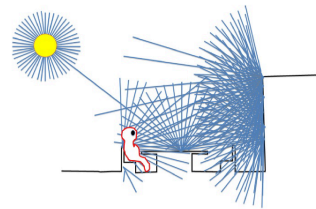


Abbildung 4: Plenoptische Funktion

Ein Farbbild enthält zusätzliche Informationen, welche die Veränderung der Lichtintensität bezüglich der Wellenlänge  $\lambda$  berücksichtigt (daher  $P(\theta, \phi, \lambda)$ ). Ein Farbvideo oder Film erweitert die Informationen um die Dimension der Zeit  $t$ :  $P(\theta, \phi, \lambda, t)$ . Schlussendlich zeigen farbige holografische Filme die komplette wahrnehmbare Lichtintensität von jeder Betrachtungsposition  $V_x, V_y$  und  $V_z$  an:  $P(\theta, \phi, \lambda, t, V_x, V_y, V_z)$ . Solch eine vollständige Repräsentation impliziert die Beschreibung aller möglichen Bilder, die von einem bestimmten Raum-Zeit Stück der Welt aufgenommen werden können (unter Vernachlässigung der Polarisation und momentanen Phase des einfallenden Lichtes). Es ist zu beachten, dass die plenoptische Funktion keine zusätzlichen Parameter zur Spezifizierung der "Blickrichtung" des Auges benötigt. Die Veränderung der Blickrichtung ohne Veränderung der Position des Auges hat keine Auswirkung auf die Verteilung des Lichtes eines Strahlenbündels beim Auftreffen auf die Pupille. Nur die relative

Einfallposition des Lichtes auf der Netzhaut wird dadurch verändert. Die Messung der plenoptischen Funktion erfolgt durch die imaginäre Platzierung eines idealen Auges an jeder möglichen  $(V_x, V_y, V_z)$  Position und beinhaltet die Messung der Intensität der Lichtstrahlen, für jeden möglichen Einfallswinkel  $(\theta, \phi)$ , für jede Wellenlänge  $\lambda$ , zu jedem Zeitpunkt  $t$ , die durch das Zentrum der Pupille gehen. Die Winkel  $(\theta, \phi)$  können immer relativ zu einer optischen Achse, die parallel zur  $V_z$ -Achse liegt, beschrieben werden. Die resultierende Funktion hat die Form:  

$$p = P(\theta, \phi, \lambda, t, V_x, V_y, V_z).$$

- Die plenoptische Funktion beschreibt alle möglichen Bilder, die von einem bestimmten Raum-Zeit-Abschnitt der Welt aufgenommen werden können.

## 5. Camera Obscura / Lochkamera

- Einfachstes Kameraprinzip.
- war die erste Spycam
- Funktionsweise: Lichtstrahlen werden durch eine kleine Öffnung auf eine Bildebene projiziert.
- Ergebnis: Verkleinertes, seitenverkehrtes Abbild.
- Diente als Grundlage für die Entwicklung der Fotografie.

Die Lochkamera besteht aus einer geschlossenen Box mit einer winzigen Öffnung an der Vorderseite ("Pinhole") und der Bildebene an der gegenüberliegenden Rückseite (siehe Abbildung 5). Lichtstrahlen, die von einem Objektpunkt vor der Kamera ausgehend durch die Öffnung einfallen, werden geradlinig auf die Bildebene projiziert, wodurch ein verkleinertes und seitenverkehrtes Abbild der sichtbaren Szene entsteht. Durch den Einsatz von Spiegeln kann das projizierte Bild in die richtige Position gedreht werden. Eine portable Version der Lochkamera stellte eine Kiste mit einem angewinkelten Spiegel dar, mit denen ein Bild in richtiger Ausrichtung auf ein durchscheinendes Papier, welches auf einer Glasoberfläche liegt, projiziert wird. Die Lochkamera kann als einfache Kamera ohne Linse und einer einzigen, kleinen Apertur (Öffnung) angesehen werden.

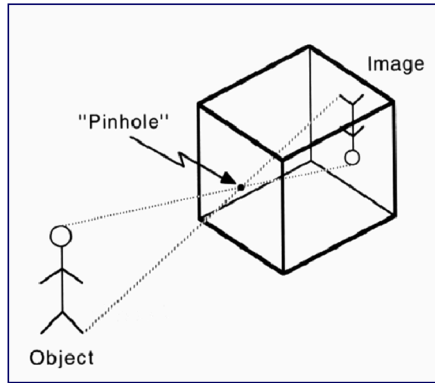


Abbildung 5: Modell der *pinhole camera*

