

---

# Visual Computing

---

Zusammenfassung  
Fabian Damken

---



TECHNISCHE  
UNIVERSITÄT  
DARMSTADT

---

---

## Inhaltsverzeichnis

---

<b>1</b>	<b>Einführung</b>	<b>7</b>
1.1	Visual Computing	7
1.1.1	3D-Internet	7
1.1.2	Skalierbare Objektmodellierung/-erkennung	7
1.1.3	Big Data, Visual Analytics	7
1.1.4	Scene Understanding	7
1.2	Generalisierte Dokumente	7
1.2.1	Retro-Digitalisierung, Digital Creation	7
1.2.2	Generative Modeling Language	7
<b>2</b>	<b>Wahrnehmung</b>	<b>8</b>
2.1	Human-Computer-Interaction	8
2.2	Überblick	8
2.2.1	Menschliche Informationsverarbeitung	8
2.3	Wahrnehmung	10
2.3.1	Das Auge	10
2.3.2	Vorverarbeitung visueller Informationen	11
2.3.3	Informationsextraktion	13
2.4	Aufmerksamkeit	15
2.4.1	Limitierung der Wahrnehmung	15
2.4.2	Das Gedächtnis und „Gateway to Memory“	15
<b>3</b>	<b>Objekterkennung und Bayes</b>	<b>17</b>
3.1	Computer Vision	17
3.1.1	Lochkamera	17
3.1.2	Digitale Bilder	17
3.1.3	Konfluenz	17
3.1.4	Fallstudie	17
3.1.5	Intuitionen	17
3.1.6	Bildaufbau	17
3.2	Bayesian Decision Theory	17
3.2.1	Konzepte und Bayes Theorem	17
3.2.2	Problemstellung	17
3.2.3	Entscheidungsregel	17
3.2.4	Naive Bayes Classifier	17
3.3	Probability Density Estimation	17
3.4	Gesichtsdetektion	17
3.4.1	Appearance-Based Methods	17
3.4.2	Sliding Window Approach	17
3.4.3	Beispiel: Gesichtsdetektion	17
3.4.4	Naive Bayes Classifier	18
3.4.5	Erkennungsarten	18
3.5	Diskussion und Ausblick	18

<b>4</b>	<b>Fouriertheorie</b>	<b>19</b>
4.1	Mathematische Grundlagen . . . . .	19
4.1.1	Vektorraum . . . . .	19
4.1.2	Basis eines Vektorraums . . . . .	19
4.1.3	Krummlinige Koordinatensysteme . . . . .	19
4.1.4	Andere Räume . . . . .	19
4.1.5	Komplexe Zahlen . . . . .	19
4.1.6	Gerade/Ungerade Funktionen . . . . .	19
4.2	Fourier-Reihe . . . . .	19
4.2.1	Dirichlet-Bedingungen . . . . .	19
4.2.2	$2\pi$ -periodische Funktion . . . . .	19
4.2.3	Skalarprodukt, Orthogonale Basis . . . . .	19
4.2.4	Berechnung der Koeffizienten $a_n, b_n$ . . . . .	19
4.2.5	Beispiel: Rechteck-Schwingung . . . . .	19
4.3	Fourier-Transformation . . . . .	19
4.3.1	Mathematische Herleitung . . . . .	19
4.3.2	Zusammenfassung . . . . .	19
4.3.3	Beispiel: Rechteckimpuls . . . . .	19
4.3.4	Fourier-Darstellung . . . . .	20
4.3.5	Transformationspaare . . . . .	20
4.4	2D-Fourier-Transformation . . . . .	20
4.5	Faltung . . . . .	20
4.5.1	Anwendung: Filter . . . . .	20
4.6	Abtastung . . . . .	20
4.6.1	Diskrete Abtastung . . . . .	20
4.6.2	Abtasttheorie . . . . .	20
4.6.3	Abtasttheorem von Whittaker-Shannon . . . . .	20
<b>5</b>	<b>Bilder</b>	<b>21</b>
5.1	Bildverbesserung . . . . .	21
5.1.1	Histogramm . . . . .	21
5.1.2	Pixeloperationen . . . . .	21
5.1.3	Kontrastspreizung . . . . .	21
5.1.4	Histogrammausgleich . . . . .	21
5.1.5	Mittelung . . . . .	21
5.2	Bildfilterung . . . . .	21
5.2.1	Ortsraum . . . . .	21
5.2.2	Frequenzraum . . . . .	22
5.2.3	Vergleich: Orts- und Frequenzraum-Filter . . . . .	22
5.3	Bildkompression . . . . .	22
5.3.1	Harmonische Transformation . . . . .	22
<b>6</b>	<b>Bildverarbeitung</b>	<b>23</b>
6.1	Deblurring . . . . .	23
6.1.1	Inverser Operator . . . . .	23
6.1.2	Rekonstruktion . . . . .	23
6.1.3	Rauschen . . . . .	23
6.1.4	Korrekt gestellte Probleme . . . . .	23
6.2	Einschrittverfahren . . . . .	23
6.2.1	Wiener Filter . . . . .	23

6.2.2	Ansatz mit mehreren Komponenten	23
6.3	Mehrschrittverfahren (Iterative Methoden)	23
6.3.1	Variationsableitung	23
6.3.2	Alternativen	23
6.3.3	Perona-Malik	23
6.3.4	Eingeschränkte Evolution: Totale Variation	24
<b>7</b>	<b>Grafikpipeline</b>	<b>25</b>
7.1	Hardware	25
7.1.1	Large-Scale-Computing	25
7.1.2	Personal/Desktop Computing	25
7.1.3	Networked Computing	25
7.1.4	Mobile Computing	25
7.1.5	Collaborative Computing	25
7.1.6	Virtual Reality	25
7.1.7	Augmented Reality	25
7.1.8	Ambient/Invisible	25
7.1.9	Wearable/Ubiquitous	25
7.2	Computergrafik	25
7.2.1	Geschichte	25
7.2.2	Virtuelle Charaktere	26
7.3	Grafikpipeline	26
7.4	Eingabe grafischer Daten	26
7.5	Repräsentation von 3D-Daten	26
7.5.1	Grafische Primitive	26
7.5.2	Transformationen	26
7.6	Räumliche Datenstrukturen	26
7.6.1	Hüllkörper (Bounding Volumes)	26
7.6.2	Hüllkörperhierarchien	26
7.6.3	Raumunterteilung	26
7.7	Geometrieverarbeitung	26
7.7.1	Modelltransformation (Orientierung, Position)	26
7.7.2	Simulation der Beleuchtung	26
7.7.3	Perspektivische Transformation und Clipping (Abschneiden)	27
7.7.4	Culling (Verdeckungsrechnung im Objektraum)	27
7.7.5	Projektion	27
7.8	Rasterisierung	27
7.8.1	Scan-Konvertierung	27
7.8.2	Verdeckungsrechnung	27
7.9	Ausgabe	27
7.10	Beispiele	27
<b>8</b>	<b>Transformationen</b>	<b>28</b>
8.1	Affine Abbildungen	28
8.1.1	Eigenschaften	28
8.1.2	Translation	28
8.1.3	Homogene Koordinaten	28
8.1.4	Matrizenschreibweise	28
8.2	Skalierung, Scherung, Rotation	28
8.2.1	Skalierung	28

8.2.2	Scherung . . . . .	28
8.2.3	Rotation . . . . .	28
8.2.4	Nicht-Kommutativität von Transformationen . . . . .	28
8.2.5	Rechenaufwand . . . . .	28
8.3	Projektion . . . . .	28
8.3.1	Perspektive Projektion . . . . .	28
8.3.2	Parallele Projektion . . . . .	29
8.3.3	Ein-, Zwei- und Dreipunktperspektive . . . . .	29
8.4	3D-Interaktion . . . . .	29
8.4.1	Manipulatoren . . . . .	29
<b>9</b>	<b>3D-Visualisierung</b>	<b>30</b>
9.1	3D-Daten . . . . .	30
9.2	Triangulation von Punktwolken . . . . .	30
9.2.1	Ideal Triangulation . . . . .	30
9.2.2	Voronoi-Diagramm und Delaunay-Triangulation . . . . .	30
9.3	Indirekte Volumenvisualisierung . . . . .	30
9.3.1	3D-Volumen und Nachbarschaft . . . . .	30
9.3.2	2D: Konturlinien . . . . .	30
9.3.3	3D: Isoflächen . . . . .	30
9.3.4	2D: Marching Squares . . . . .	30
9.3.5	3D: Marching Cubes . . . . .	30
9.3.6	Große Polygonmodelle und Performanz . . . . .	30
9.4	Direkte Volumenvisualisierung . . . . .	31
9.4.1	Density Emitter Model . . . . .	31
9.4.2	Volumen-Rendering-Gleichung . . . . .	31
9.5	Volumen-Rendering-Pipeline . . . . .	31
9.5.1	Pipeline . . . . .	31
9.5.2	Transferfunktion . . . . .	31
<b>10</b>	<b>Szenengraphen am Beispiel X3DOM</b>	<b>32</b>
10.1	Strukturierung von 3D-Szenendaten . . . . .	32
10.2	Szenengraph . . . . .	32
10.3	X3DOM . . . . .	32
<b>11</b>	<b>Informationsvisualisierung</b>	<b>33</b>
11.1	Informationsdesign . . . . .	33
11.1.1	Referenzmodell von Card . . . . .	33
11.2	Datentypen . . . . .	33
11.2.1	1D-Daten, Zeitreihen . . . . .	33
11.2.2	2D-Daten . . . . .	33
11.2.3	mD-Daten (multidimensional) . . . . .	33
11.2.4	Hierarchien . . . . .	33
11.2.5	Graphen/Netzwerke . . . . .	33
11.3	Kuchendiagramm (1D) . . . . .	33
11.4	Balkendiagramm (1D) . . . . .	33
11.5	Liniendiagramm (Zeitreihe) . . . . .	33
11.6	Scatterplot (2D, 3D) . . . . .	33
11.7	Scatterplotmatrix (nD) . . . . .	33
11.8	Parallele Koordinaten (3D, nD) . . . . .	33

11.9 Node-Link-Diagramm (Hierarchien, Graphen) . . . . .	34
11.10 Treemap (Hierarchien) . . . . .	34
11.11 Zusammenfassung . . . . .	34
<b>12 Farbe</b>	<b>35</b>
12.1 Dimensionalität . . . . .	35
12.2 Wahrnehmungskorrelate . . . . .	35
12.3 Reproduktion . . . . .	35
12.4 Berechnung von Farbattributen . . . . .	35
12.4.1 Das Auge . . . . .	35
12.4.2 Spektrale Charakterisierung des Auges . . . . .	35
12.4.3 Spektralwertfunktion . . . . .	35
12.4.4 Cone Fundamentals . . . . .	35
12.5 Objektfarben, Lichtmatrix und XIEXYZ-Farbraum . . . . .	35
12.6 Metamerie . . . . .	35
12.7 Gegenfarbtheorie . . . . .	35
12.8 Stevenssche Potenzfunktion . . . . .	35
12.9 CIELAB Farbraum . . . . .	35
12.10 Technische Farbräume . . . . .	35
12.10.1 Geräte RGB . . . . .	35
12.10.2 Geräteunabhängige RGB . . . . .	35
12.10.3 YCbCr . . . . .	35
12.10.4 HSI/HSV/HSL . . . . .	36
12.10.5 CMY/CMYK . . . . .	36
12.11 Komplexität von Farbe . . . . .	36
12.11.1 Chromatische Adaptation . . . . .	36
12.11.2 Farbwahrnehmungsphänomene . . . . .	36
12.11.3 Farbwahrnehmungsmodelle . . . . .	36
12.11.4 Kontrastsensitivität . . . . .	36
<b>13 User Interfaces</b>	<b>37</b>
13.1 Interaktion . . . . .	37
13.1.1 Möglichkeiten . . . . .	37
13.1.2 Designprozess . . . . .	37
13.2 GUI: Benutzeroberflächen . . . . .	37
13.2.1 Das WIMP-Interface . . . . .	38
13.2.2 Menübasierte Programme . . . . .	38
13.2.3 GUI-Anwendungen und Event-basiertes Programmieren . . . . .	38
13.3 3D-Interaktion . . . . .	38
<b>14 Multimedia Information Retrieval</b>	<b>39</b>
14.1 Inhaltsbasierte Suche . . . . .	39
14.1.1 Mathematische Beschreibung . . . . .	39
14.1.2 Retrieval Ergebnis . . . . .	39
14.2 Distanzmaße . . . . .	39
14.3 Query-Modalitäten . . . . .	39
14.3.1 Text . . . . .	39
14.3.2 Example . . . . .	39
14.3.3 Example-Bilder . . . . .	39
14.3.4 Sketch . . . . .	39

---

14.4 Explorative Suche . . . . .	39
14.4.1 Research Data . . . . .	39

---

---

## **1 Einführung**

---

### **1.1 Visual Computing**

---

1.1.1 3D-Internet

---

1.1.2 Skalierbare Objektmodellierung/-erkennung

---

1.1.3 Big Data, Visual Analytics

---

1.1.4 Scene Understanding

---

### **1.2 Generalisierte Dokumente**

---

1.2.1 Retro-Digitalisierung, Digital Creation

---

1.2.2 Generative Modeling Language

---



---

## 2 Wahrnehmung

---

### 2.1 Human-Computer-Interaction

---

Abbildung 2.1 zeigt den klassischen Zyklus der *Human-Computer-Interaction* (HCI), d. h. der Interaktion zwischen Mensch und Maschine. Dabei dient insbesondere die visuelle Interaktion und Kommunikation über das Auge eine große Rolle.

### 2.2 Überblick

---

Der Mensch hat fünf grundlegende Sinne: Sehen, Hören, Fühlen, Schmecken und Riechen, wobei das Sehen, Hören und Fühlen derzeit dominant sind. Der heute sicherlich relevanteste Sinn ist dabei das Sehen und das menschliche Auge. Da die meisten erzeugten Bilder der Kommunikation von und zum Menschen dienen sollen, ist es gut, das menschliche visuelle System zu kennen, um den Informationstransfer optimal zu gestalten (der Monitorausgang ist nicht das Ende des Informationsflusses).

Hören und Fühlen sind dabei relevant für die Informationsaufnahme und Interaktion mit der realen Welt (außerhalb der Mensch-Maschine-Interaktion).

Bei der Gestaltung von Kommunikation gibt es zwei große Probleme:

- Die Wahrnehmung ist nicht objektiv.
- Das visuelle System ist stark nichtlinear (es ist keine einfache Interpolation oder Extrapolation von Versuchsergebnissen möglich).

---

#### 2.2.1 Menschliche Informationsverarbeitung

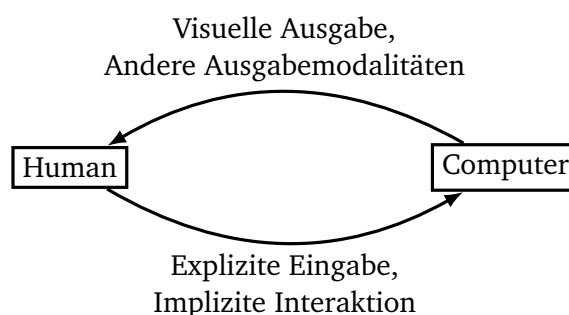
---

Abbildung 2.2 zeigt die drei Stufen der menschlichen Informationsverarbeitung:

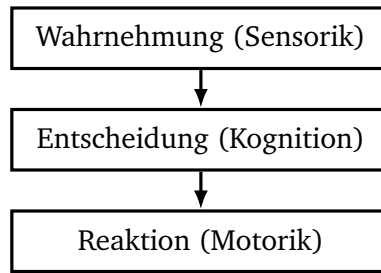
- *Wahrnehmung* von Eindrücken durch die Sinne,
- *Entscheidungsfindung* im Gehirn und
- *Reaktion* durch den Körper.

Dabei verhält sich die Ausführungszeit additiv und die Funktionen werden durch neurologisch getrennte Gehirnteile ausgeführt, die „elektronisch“ verbunden sind.

Dabei braucht jede Bearbeitung in den einzelnen Stufen unterschiedlich lange und die benötigten Zeiten können verwendet werden, um die Performanz abzuschätzen, bzw. vorherzusagen (bspw. für die Bildfrequenz von Filmen, die maximale Morserate, ...). Typische Zeiten sind in Tabelle 2.1 abgebildet.



**Abbildung 2.1:** Klassischer Zyklus der Human-Computer-Interaction (HCI).



**Abbildung 2.2:** Modulares Drei-Stufenmodell der menschlichen Informationsverarbeitung.

Untersystem	Durchschnitt	Bereich
Wahrnehmung (Perzeption)	100 ms	50 ms to 200 ms
Entscheidung (Kognition)	70 ms	25 ms to 170 ms
Reaktion (Motorik)	70 ms	30 ms to 100 ms

**Tabelle 2.1:** Typische Bearbeitungszeiten der Untersysteme der menschlichen Informationsverarbeitung.

---

## Eingabe (Wahrnehmung)

---

Die Untersysteme der Wahrnehmung,

- Visuell (Sehen)
- Akustisch (Hören)
- Haptisch (Fühlen)

können dabei (theoretisch) parallel arbeiten.

### Klangwahrnehmung

Die Hauptkomponenten von Klängen sind

- Klangfarbe,
- Tonlage und
- Lautstärke.

Diese werden durch verschiedene Mechanismen wahrgenommen und Informationen (z. B. der Ursprung eines Geräuschs) extrahiert.

### Berührungswahrnehmung

Die Hauptkomponenten der Haptik sind

- Fühl- und Tastsinn (Temperatur, Schmerz, Druck, Oberflächen) und
- Propriozeption (Wahrnehmung der Bewegung und Lage der eigenen Körperglieder).

Dabei interagiert die Haptik stark mit Sehen und Hören, was bei sich widersprechenden Informationen Illusionen hervorrufen kann. Ein User-Interface-Designer nutzt Illusionen dabei gezielt aus, um bestimmte Informationen zu vermitteln.

---

## Ausgabe (Reaktion)

---

Die Untersysteme der Reaktion,

- Artikulation (Sprechen)
- Motorisch (Bewegen)

können dabei (theoretisch) parallel arbeiten.

Die motorische Ausgabe kann dabei auf verschiedene Weisen angewandt werden:

- Diskret (Schalter) oder
- Kontinuierlich (Heben).

Sie ist dabei beschränkt durch Geschwindigkeit, Stärke, Koordinationsvermögen, Wendigkeit, . . . . Neurologisch ist die motorische Ausgabe dabei mit dem haptischen System verbunden (Reflexe).

Das *Muskelgedächtnis* hilft dabei, relevante Positionen im Raum (z. B. die Gangschaltung im Auto) zu lernen.

---

## 2.3 Wahrnehmung

---

---

### 2.3.1 Das Auge

---

---

#### Reiz und Licht

---

Einer äußerer, visueller Reiz (Licht) erzeugt beim Menschen eine physikalische Rezeption des äußeren Reizes (Input). Dies geschieht durch einen Sensor (bspw. das Auge) und die Reizung produziert ein neuro-physiologisches Signal. Dieses wird anschließend verarbeitet und interpretiert.

Physikalisch ist ein solcher Reiz elektromagnetische Strahlung. Dabei wird monochromatisches, d. h. einfarbiges, Licht durch die Angabe der Frequenz  $\nu$ , bzw. der Wellenlänge  $\lambda$ , beschrieben. Diese beiden Größen sind durch die Beziehung

$$\nu\lambda = c, \quad c \approx 3 \times 10^8 \text{ m s}^{-1}$$

miteinander verknüpft, wobei  $c$  die Ausbreitungsgeschwindigkeit des Lichts ist.

Das menschliche Auge kann dabei Frequenzen im Wellenlängenbereich 380 nm to 750 nm wahrnehmen. Kleinere Wellenlängen haben z. B. Ultraviolett-Licht, Röntgen- und  $\gamma$ -Strahlung. Darüber liegende Wellenlängen haben z. B. Infrarot-Licht und Rundfunk-Wellen.

---

#### Das visuelle System

---

Das menschliche Auge ist aufgebaut aus:

- Hornhaut (Kornea)
- Linse (zur Scharfstellung)
- Iris (Blendenmechanismus)
- Retina (Netzhaut)
  - Blinder Fleck: Hier geht der Sehnerv ab.
  - Fovea Centralis (Gelber Fleck): Bereich mit der höchsten Auflösung.

---

## Photorezeptoren

---

Die Photorezeptoren (welche auf der Retina platziert sind), bestehen aus:

- Stäbchen
  - Hauptsächlich außerhalb der Fovea.
  - Das Empfindlichkeitsmaximum liegt bei 498 nm („grün“).
- Zapfen
  - Vor allem in der Fovea platziert.
  - Es gibt drei Zapfentypen für Farbsehen.
  - Das Empfindlichkeitsmaximum dieser Zapfen liegt bei 420 nm („blau“), 534 nm („grün“) und 564 nm („rot“).

---

## Skotopisches und Photopisches Sehen

---

- Nachtsehen (skotopisch): Dominanz der Stäbchen.
- Tagsehen (photopisch): Dominanz der Zapfen.

---

## Zapfenverteilung

---

---

### 2.3.2 Vorverarbeitung visueller Informationen

---

---

## Signalverarbeitung in der Retina

---

Neben den Photorezeptoren gibt es noch weitere Zellen zur Signalverarbeitung in der Retina:

- Horizontale Zellen  
Kombination von mehreren Rezeptoren einer Region.
- Amakrin-Zellen  
Zeitliche Verarbeitung.
- Bipolar-Zellen  
Informationsfilter (Sammeln, Gewichten und Weiterleiten).
- Ganglien-Zellen  
Integration Informationen (z. B. Kontrastwahrnehmung).

---

## Helligkeit

---

- *Helligkeit (brightness)* entspricht der wahrgenommenen Menge an Licht, das von einer selbstleuchtenden Lichtquelle ausgeht.
- *Helligkeit (lightness)* entspricht der wahrgenommenen Menge an Licht, das von einer reflektierenden Oberfläche ausgeht.
  - Dies ist keine absolute Wahrnehmungsgröße und abhängig von

- \* Reizstärke (Leuchtdichte)
  - \* Vorherige Leuchtdichte (Adaption)
  - \* Umgebungsleuchtdichte
  - \* Größe (Fläche) des Reizes
- Somit subjektiv!

- Dies wirft einige nicht so einfach zu beantwortende Fragen auf, z. B.: Was ist weiß? Was ist schwarz? Was ist mittelgrau?
- Der Hell-Dunkel-Kontrast ist dabei eine wichtige Empfindungsgröße zum Form- und Objektsehen. Daher muss der Unterschied groß genug sein (für kleine Details mindestens 3 : 1, besser 10 : 1).

### Kontrast als Reizverhältnis

Für den Kontrast gibt es verschiedene Definitionen, z. B. (dabei ist  $L$  stets die Leuchtdichte):

$$m = k = \frac{L_{\max} - L_{\min}}{L_{\max} + L_{\min}}$$

oder

$$K = \frac{L_R - L_H}{L_H} = \frac{\delta L}{L_H}$$

wobei  $L_R$  die Leuchtdichte des Vordergrunds und  $L_H$  die Leuchtdichte des Hintergrunds darstellt.

---

### Erkennung von Details

---

Die Erkennung kleiner Details ist begrenzt durch

- Optische Eigenschaften des Auges, z. B. Beugungserscheinungen,
- Abtastung durch Rezeptoren und
- nervöse Verarbeitung.

Zwei mögliche Maße zur „Erkennbarkeit“ sind:

- Kontrastempfindlichkeit
- Schwellenkontrast

### Kontrastempfindlichkeit

Die Kontrastempfindlichkeit ist die Auflösung des menschlichen Auges im Frequenzraum. Veränderliche Intensität kann dabei mit Sinus-förmigen Mustern gemessen werden.

---

### Frühe Wahrnehmung

---

Das Auge nimmt einige Veränderungen der Umgebung schneller wahr als andere. Um die Aufmerksamkeit auf etwas zu lenken, können beispielsweise

- Farbe,
- Richtung,

- 
- Bewegung,
  - Größe,
  - Beleuchtung/Schattierung

variiert werden.

---

### 2.3.3 Informationsextraktion

---

Ein reiner Reiz ist noch keine *Wahrnehmung*. Dazu kommen noch andere Faktoren wie Kontext, Erwartungen, Adaption. Das Messen der tatsächlichen Wahrnehmung ist leider sehr schwierig, weshalb häufig nur statistische Aussagen auf Basis von User-Tests getätigt werden können.

Dabei wird erschwert, dass die Wahrnehmung nicht immer der Realität entspricht. Es wird hingegen das Bild durch einen Wahrnehmungsprozess im Gehirn produziert. Dabei wird die menschliche Wahrnehmung adaptiert, bspw. dreht sich das Bild bei einem Kopfstand.

---

### Raumwahrnehmung

---

Die Wahrnehmung des Raums (Raumwahrnehmung) enthält unter anderem

- Tiefenwahrnehmung,
- Entfernungs- und Distanzwahrnehmung und
- Ausrichtung des Körpers im Raum.

Daran sind viele Wahrnehmungssysteme beteiligt:

- Vestibuläres System (im Innenohr)
- Haptisch-somatisches System (Tasten und Berühren)
- Auditives Sehen (Gehört)
- Propriozeptives System (Eigenwahrnehmung)
- Visuelles System

Dabei ist die Raumwahrnehmung auch mit einem Auge (Monokular) möglich (tatsächlich sind 5 % to 10 % aller Menschen stereoblind und 20 % haben eine Stereo-Schwäche).

Tatsächlich ist die Raumwahrnehmung ein sehr komplexer Prozess, der auch heute nur zu Teilen verstanden wird. Dabei fließen noch viele weitere Phänomene ein, z. B. Größenkonstanz, Annahme starrer Körper oder Vektion. Letzteres ist dabei die scheinbare Eigenbewegung bei einem statischen Vordergrund als Referenzrahmen und einem bewegtem Hintergrund.

---

### Depth Cue Theorie

---

Die Annahme der *Depth Cue Theorie* ist, dass die Raumwahrnehmung des visuellen Systems auf Hinweisreizen (sogenannten *Depth Cues*) basiert. Diese werden in drei Kategorien eingeteilt:

1. Binokulare Depth Cues (mit zwei Augen)
  - Disparität/Parallaxe
  - Akkomodation (Krümmung der Augenlinsen)

- 
- Konvergenz (die Augen nach innen drehen)

## 2. Pictorial Depth Cues (mit einem Auge)

- Linearperspektive
- Verdeckung
- Texturgradient
- Fokus und Blur
- Atmosphärische Tiefe
- Vertraute Größe
- Höhe im Gesichtsfeld
- Beleuchtung
- Schattenwurf
- Luminanzänderung
- Transluzenz
- Schattierung

## 3. Dynamische Depth Cues (Animation)

- Bewegungsparallaxe
- Kinetischer Tiefeneffekt
- Interposition
- Bewegung von Highlights

### Stereoskopie

Bei der Stereoskopie nehmen beide Augen ein leicht unterschiedliches Bild wahr, woraus die Entfernung zu einem Objekt berechnet werden kann.

### Pictorial Depth Cues

#### Linearperspektive:

**Texturgradient:** Sind als parallel angenommene Linien nicht mehr parallel, so ergibt sich eine scheinbare Tiefe (als wenn kariertes Papier um einen Ball gerollt und von oben betrachtet wird).

**Fokus und Blur:** Das Auge fokussiert einen Punkt und produziert somit eine Tiefenschärfe. Daran kann erahnt werden, welche Objekte im Vorder- oder Hintergrund sind.

**Atmosphärische Tiefe:** Anhand der Atmosphäre (z. B. durch Nebel ausgelöst) wird erkannt, was vermutlich im Hintergrund liegt. So kann zum Beispiel bei einem Foto von einem Berg geschätzt werden, dass der Boden niedriger ist, wenn Wolken über diesem hängen.

**Schattenwurf:** Annahme: Beleuchtung von oben und Vorhandensein einer Grundebene. Dann kann durch den Abstand von Schatten zum Objekt erahnt werden, wie weit dieses vom Boden entfernt ist.

### Dynamische Depth Cues

#### Motion Parallax:

**Raumwahrnehmung durch Bewegung:** Wird z. B. eine schaukelnde Vase von oben betrachtet, so bewegt sich die Öffnung charakteristisch, sodass eine Wahrnehmung der Tiefe entsteht.

#### Kinetic Depth Effect, Structure from Motion:

---

## Auswertung von Depth Cues

Unterschiedliche Depth Cues haben im Allgemeinen einen unterschiedlichen Informationsgehalt. Dabei sind sie nicht redundant, sondern additiv. Durch ein kompliziertes Zusammenspiel (flexible Gewichtung, Dominanz eines Depth Cue) bildet sich das Gehirn ein Bild. Dabei bildet es sich allerdings kein tatsächliches 3D-Modell, sondern verwendet sie unterschiedlichen Cues für verschiedene Aufgaben. Diese können z. B. sein:

- Einschätzen von Objektgrößen
- Einschätzen von Entfernungen
- Verfolgung von Pfaden
- Navigation
- Einschätzen der Eigenbewegung
- Abschätzung der Kollisionszeit

---

## 2.4 Aufmerksamkeit

---

### 2.4.1 Limitierung der Wahrnehmung

---

Die initiale Reizaufnahme hat viele Limitierungen, sodass nur ein Bruchteil des äußeren Reizes zur kognitiven Verarbeitung zur Verfügung steht. Dabei sind Aufmerksamkeit und externe Faktoren wichtige Einflüsse auf die tatsächliche Wahrnehmung. Die Wahrnehmung ist dabei eher eine partielle Hypothese, die auf Basis unvollständiger Informationen generiert wurde. Es wird dabei periodisch aktualisiert aufgrund von Beobachtungen, d. h. die Hypothese wird gegen sensorische Daten getestet. Durch eine dynamische Suche des visuellen Systems wird nach der besten Hypothese/Interpretation/Modell gesucht.

---

### 2.4.2 Das Gedächtnis und „Gateway to Memory“

---

Das Gehirn kann sich auf bestimmte Dinge fokussieren und den Rest ignorieren. Dabei gibt es drei verschiedene Arten der Aufmerksamkeit:

- *Gewählte Aufmerksamkeit* (selective): Zwischen mehreren Möglichkeiten wird eine zu fokussierende Sache aktiv ausgewählt.
  - Das Auge folgt den Objekten von Interesse.
  - Der Kopf folgt den Klängen von Interesse.
  - Es gibt nur einen einzigen „Ort der Aufmerksamkeit“.
- *Geteilte Aufmerksamkeit* (divided): Ein Versuch durch „Multitasking“ mehrere Dinge zu fokussieren.
  - Entweder „gleichzeitig“ oder durch schnelles Umschalten (time multiplexing).
  - Dies wirkt sich negativ auf die Verarbeitung aus, wenn die Aufgaben überfordernd sind.
  - Die Aufgaben beeinträchtigen sich gegenseitig.
- *Erfasste Aufmerksamkeit* (captured): Ein äußerer Reiz zieht alle Aufmerksamkeit auf sich.
  - Im Gegensatz zur gewählten Aufmerksamkeit wird der „Ort“ nicht aktiv ausgewählt.



---

– Dies geschieht z. B. wenn man von einem Tier angefallen wird.

Das menschliche Gedächtnis ist in mehrere „Teilgedächtnisse“ aufgeteilt. Voran steht das *Arbeitsgedächtnis*, auf das ein schneller Zugriff (ca. 70 ms) möglich ist, welches aber einen schnellen Verfall hat (nach ca. 200 ms). Nach wenigen Sekunden wird der Inhalt jedoch an das Langzeitgedächtnis weitergegeben. Es stellt sozusagen das „Schmierblatt“ des Gehirns da.

Das Langzeitgedächtnis ist langsamer (ca. 100 ms), dafür aber auch sehr viel größer (die genaue Größe ist unbekannt). Das Langzeitgedächtnis hat dabei drei Hauptaufgaben:

- Informationen speichern und sich an diese erinnern,
- Informationen abrufen und
- Informationen vergessen.

---

---

### **3 Objekterkennung und Bayes**

---

#### **3.1 Computer Vision**

---

3.1.1 Lochkamera

---

3.1.2 Digitale Bilder

---

3.1.3 Konfluenz

---

3.1.4 Fallstudie

---

3.1.5 Intuitionen

---

3.1.6 Bildaufbau

---

#### **3.2 Bayesian Decision Theory**

---

3.2.1 Konzepte und Bayes Theorem

---

3.2.2 Problemstellung

---

3.2.3 Entscheidungsregel

---

3.2.4 Naive Bayes Classifier

---

#### **3.3 Probability Density Estimation**

---

#### **3.4 Gesichtsdetektion**

---

3.4.1 Appearance-Based Methods

---

3.4.2 Sliding Window Approach

---

3.4.3 Beispiel: Gesichtsdetektion

---

---

---

3.4.4 Naive Bayes Classifier

---

3.4.5 Erkennungsarten

---

**3.5 Diskussion und Ausblick**

---

---

---

## 4 Fouriertheorie

---

### 4.1 Mathematische Grundlagen

---

4.1.1 Vektorraum

---

4.1.2 Basis eines Vektorraums

---

4.1.3 Krummlinige Koordinatensysteme

---

4.1.4 Andere Räume

---

4.1.5 Komplexe Zahlen

---

4.1.6 Gerade/Ungerade Funktionen

---

### 4.2 Fourier-Reihe

---

4.2.1 Dirichlet-Bedingungen

---

4.2.2  $2\pi$ -periodische Funktion

---

4.2.3 Skalarprodukt, Orthogonale Basis

---

4.2.4 Berechnung der Koeffizienten  $a_n, b_n$

---

4.2.5 Beispiel: Rechteck-Schwingung

---

### 4.3 Fourier-Transformation

---

4.3.1 Mathematische Herleitung

---

4.3.2 Zusammenfassung

---

4.3.3 Beispiel: Rechteckimpuls

---

---

---

4.3.4 Fourier-Darstellung

---

4.3.5 Transformationspaare

---

**4.4 2D-Fourier-Transformation**

---

**4.5 Faltung**

---

4.5.1 Anwendung: Filter

---

**4.6 Abtastung**

---

4.6.1 Diskrete Abtastung

---

4.6.2 Abtasttheorie

---

4.6.3 Abtasttheorem von Whittaker-Shannon

---

---

---

## 5 Bilder

---

### 5.1 Bildverbesserung

---

#### 5.1.1 Histogramm

---

#### 5.1.2 Pixeloperationen

---

Bildnegativ

---

Binärisierung/Thresholding

---

Graufenfensterung

---

#### 5.1.3 Kontrastspreizung

---

#### 5.1.4 Histogrammausgleich

---

#### 5.1.5 Mittelung

---

### 5.2 Bildfilterung

---

#### 5.2.1 Ortsraum

---

Lineare Filterung (Faltung)

---

Tiefpass-Filter

---

Mittelwert-Filter

---

Gauss-Filter

---

Median-Filter

---

Hochpass-Filter

---

---

Diskretisierte Ableitungen

---

**Laplacian-Filter**

**Laplacian of Gaussian Filter**

---

Bilateral Filter

---

5.2.2 Frequenzraum

---

Idealer Tiefpass-Filter

---

Gaußscher Tiefpass-Filter

---

Idealer Hochpass-Filter

---

5.2.3 Vergleich: Orts- und Frequenzraum-Filter

---

### **5.3 Bildkompression**

---

5.3.1 Harmonische Transformation

---

JPEG

---

**Schritt 1: Umwandlung in den YCC-Farbraum**

**Schritt 2: Farb-Subsampling**

**Schritt 3: Diskrete Kosinustransformation**

**Schritt 4: Quantisierung**

**Schritt 5: Kodierung der Koeffizienten**

---

MH17

---

---

---

## 6 Bildverarbeitung

---

### 6.1 Deblurring

---

6.1.1 Inverser Operator

---

6.1.2 Rekonstruktion

---

6.1.3 Rauschen

---

6.1.4 Korrekt gestellte Probleme

---

### 6.2 Einschrittverfahren

---

6.2.1 Wiener Filter

---

6.2.2 Ansatz mit mehreren Komponenten

---

Scale-Space-Ansatz

---

Gaußscher Scale-Space

---

### 6.3 Mehrschrittverfahren (Iterative Methoden)

---

6.3.1 Variationsableitung

---

6.3.2 Alternativen

---

6.3.3 Perona-Malik

---

Nichtlineare Diffusionsgleichungen

---

Die Perona-Malik-Gleichung

---

Implementierung

---



---

Einfluss und Wahl von  $k$

---

Stopzeit

---

#### 6.3.4 Eingeschränkte Evolution: Totale Variation

---

Distance Penalty

---

Basisenergien

---

Totale Variation

---

Erweiterungen

---

---

---

## **7 Grafikpipeline**

---

### **7.1 Hardware**

---

7.1.1 Large-Scale-Computing

---

7.1.2 Personal/Desktop Computing

---

7.1.3 Networked Computing

---

7.1.4 Mobile Computing

---

7.1.5 Collaborative Computing

---

7.1.6 Virtual Reality

---

7.1.7 Augmented Reality

---

7.1.8 Ambient/Invisible

---

7.1.9 Wearable/Ubiquitous

---

### **7.2 Computergrafik**

---

7.2.1 Geschichte

---

Die Anfänge

---

60er Jahre

---

70er Jahre

---

80er Jahre

---

90er Jahre

---

---

---

2000 bis 2005

---

2006 bis 2020

---

7.2.2 Virtuelle Charaktere

---

### **7.3 Grafikpipeline**

---

#### **7.4 Eingabe grafischer Daten**

---

#### **7.5 Repräsentation von 3D-Daten**

---

7.5.1 Grafische Primitive

---

7.5.2 Transformationen

---

#### **7.6 Räumliche Datenstrukturen**

---

7.6.1 Hüllkörper (Bounding Volumes)

---

7.6.2 Hüllkörperhierarchien

---

7.6.3 Raumunterteilung

---

Achsenparallele Gitter (Grids)

---

Quadtree/Octree

---

Binary Space Partition

---

### **7.7 Geometrieverarbeitung**

---

7.7.1 Modelltransformation (Orientierung, Position)

---

7.7.2 Simulation der Beleuchtung

---

---

---

Phong-Beleuchtungsmodell

---

7.7.3 Perspektivische Transformation und Clipping (Abschneiden)

---

Painters Algorithmus

---

7.7.4 Culling (Verdeckungsrechnung im Objektraum)

---

7.7.5 Projektion

---

## **7.8 Rasterisierung**

---

7.8.1 Scan-Konvertierung

---

Rasterisierung von Linien (Bresenham-Algorithmus)

---

Rasterisierung von Polygonen (Scanline Algorithmus)

---

7.8.2 Verdeckungsrechnung

---

z-Buffer-Algorithmus

---

## **7.9 Ausgabe**

---

## **7.10 Beispiele**

---

---

---

## 8 Transformationen

---

### 8.1 Affine Abbildungen

---

#### 8.1.1 Eigenschaften

---

#### 8.1.2 Translation

---

#### 8.1.3 Homogene Koordinaten

---

Translation

---

#### 8.1.4 Matrizenschreibweise

---

### 8.2 Skalierung, Scherung, Rotation

---

#### 8.2.1 Skalierung

---

#### 8.2.2 Scherung

---

#### 8.2.3 Rotation

---

Rotation um beliebige Achse

---

Rotation um beliebige Raumachse

---

#### 8.2.4 Nicht-Kommutativität von Transformationen

---

#### 8.2.5 Rechenaufwand

---

### 8.3 Projektion

---

#### 8.3.1 Perspektive Projektion

---

Kanonisches Sichtvolumen

---

---

---

Allgemeine perspektivische Transformation

---

8.3.2 Parallele Projektion

---

8.3.3 Ein-, Zwei- und Dreipunktperspektive

---

## **8.4 3D-Interaktion**

---

8.4.1 Manipulatoren

---

---

---

## **9 3D-Visualisierung**

---

### **9.1 3D-Daten**

---

Terrain

Laser Scanning

Range Images

Medizinische Bilddaten

Schneiden

Wetter

Klimaforschung

---

### **9.2 Triangulation von Punktwolken**

---

9.2.1 Ideal Triangulation

---

9.2.2 Voronoi-Diagramm und Delaunay-Triangulation

---

### **9.3 Indirekte Volumenvisualisierung**

---

9.3.1 3D-Volumen und Nachbarschaft

---

9.3.2 2D: Konturlinien

---

9.3.3 3D: Isoflächen

---

9.3.4 2D: Marching Squares

---

9.3.5 3D: Marching Cubes

---

9.3.6 Große Polygonmodelle und Performanz

---

Culling von Geometrie

---

Meshreduktion

---

Mesh-Glättung

---

---

## **9.4 Direkte Volumenvisualisierung**

---

### 9.4.1 Density Emitter Model

---

### 9.4.2 Volumen-Rendering-Gleichung

---

## **9.5 Volumen-Rendering-Pipeline**

---

### 9.5.1 Pipeline

---

**Abtastung**

**Klassifikation und Beleuchtung**

**Komposition**

**Back-to-Front-Komposition:**

**Front-to-Back-Komposition:**

---

### 9.5.2 Transferfunktion

---



---

---

## **10 Szenengraphen am Beispiel X3DOM**

---

### **10.1 Strukturierung von 3D-Szenendaten**

---

### **10.2 Szenengraph**

---

### **10.3 X3DOM**

---

---

---

## **11 Informationsvisualisierung**

---

### **11.1 Informationsdesign**

---

#### 11.1.1 Referenzmodell von Card

---

### **11.2 Datentypen**

---

#### 11.2.1 1D-Daten, Zeitreihen

---

#### 11.2.2 2D-Daten

---

#### 11.2.3 mD-Daten (multidimensional)

---

#### 11.2.4 Hierarchien

---

#### 11.2.5 Graphen/Netzwerke

---

### **11.3 Kuchendiagramm (1D)**

---

### **11.4 Balkendiagramm (1D)**

---

### **11.5 Liniendiagramm (Zeitreihe)**

---

**Problem: Viele Zeitreihen**

**Problem: Länge**

---

### **11.6 Scatterplot (2D, 3D)**

---

**Problem: Overplotting**

---

### **11.7 Scatterplotmatrix (nD)**

---

### **11.8 Parallele Koordinaten (3D, nD)**

---

**Problem: Overplotting**

**Problem: Viele Dimensionen**

---

---

## **11.9 Node-Link-Diagramm (Hierarchien, Graphen)**

---

**Problem: Layout**

**Problem: Viele Knoten**

---

## **11.10 Treemap (Hierarchien)**

---

**Problem: Überlappung**

**Problem: Größendarstellung**

---

## **11.11 Zusammenfassung**

---

---

---

## **12 Farbe**

---

### **12.1 Dimensionalität**

---

### **12.2 Wahrnehmungskorrelate**

---

### **12.3 Reproduktion**

---

### **12.4 Berechnung von Farbattributen**

---

#### 12.4.1 Das Auge

---

#### 12.4.2 Spektrale Charakterisierung des Auges

---

#### 12.4.3 Spektralwertfunktion

---

#### 12.4.4 Cone Fundamentals

---

### **12.5 Objektfarben, Lichtmatrix und XIEXYZ-Farbraum**

---

### **12.6 Metamerie**

---

### **12.7 Gegenfarbtheorie**

---

### **12.8 Stevenssche Potenzfunktion**

---

### **12.9 CIELAB Farbraum**

---

### **12.10 Technische Farbräume**

---

#### 12.10.1 Geräte RGB

---

#### 12.10.2 Geräteunabhängige RGB

---

#### 12.10.3 YCbCr

---

---

---

12.10.4 HSI/HSV/HSL

---

12.10.5 CMY/CMYK

---

## **12.11 Komplexität von Farbe**

---

12.11.1 Chromatische Adaptation

---

Modellbildung

---

12.11.2 Farbwahrnehmungsphänomene

---

Simultankontrast

---

Crispening Effekt

---

Stevens Effekt

---

Hunt Effekt

---

12.11.3 Farbwahrnehmungsmodelle

---

CIECAM02

---

12.11.4 Kontrastsensitivität

---

S-CIELAB (Spacial-CIELAB)

---

iCAM (Image Color Appearance Model)

---

---

---

## **13 User Interfaces**

---

### **13.1 Interaktion**

---

#### 13.1.1 Möglichkeiten

---

Kommandozeile

---

Menüs

---

Formulare

---

Fragen und Antworten

---

Direkte Manipulation

---

3D-Umgebungen

---

Natürliche Sprache

---

Gesten

---

#### 13.1.2 Designprozess

---

Wasserfallmodell

---

Spiralmodell

---

V-Modell

---

Dynamic Systems Development Method (DSDM)

---

Design Process Model

---

### **13.2 GUI: Benutzeroberflächen**

---

---

---

### 13.2.1 Das WIMP-Interface

---

Fenster-Komponenten

---

**Multiple Document Interface (MDI)**

**Single Document Interface (SDI)**

**Tabbed Document Interface**

---

Dialogboxen

---

**Checkboxen**

**Radio Buttons**

**Listboxen**

**Comboboxen**

**Spinner**

**Slider**

**Weiteres**

---

### 13.2.2 Menübasierte Programme

---

Untermenüs

---

Auswahl (if-then-else-Struktur)

---

**Verschachtelte Entscheidungsstrukturen**

---

Die case-Struktur

---

Modularisierung

---

### 13.2.3 GUI-Anwendungen und Event-basiertes Programmieren

---

Graphical User Interfaces (GUIs)

---

Event-Handler

---

Nutzerinteraktionen

---

Das Delegationsmodell

---

## 13.3 3D-Interaktion

---

---

---

## **14 Multimedia Information Retrieval**

---

### **14.1 Inhaltsbasierte Suche**

---

#### 14.1.1 Mathematische Beschreibung

---

#### 14.1.2 Retrieval Ergebnis

---

### **14.2 Distanzmaße**

---

### **14.3 Query-Modalitäten**

---

#### 14.3.1 Text

---

#### 14.3.2 Example

---

#### 14.3.3 Example-Bilder

---

#### 14.3.4 Sketch

---

### **14.4 Explorative Suche**

---

#### 14.4.1 Research Data

---