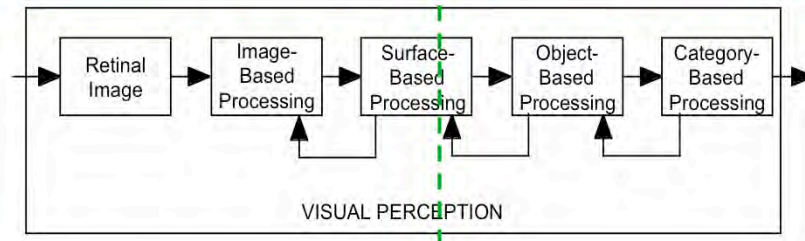


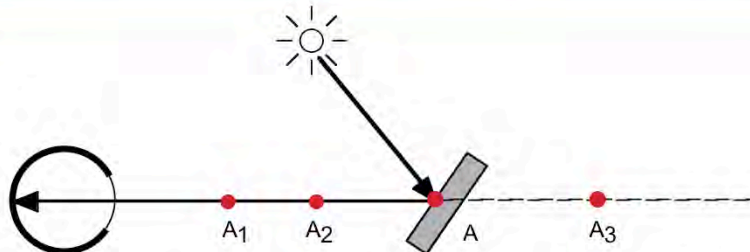
Vier Stufen der visuellen Wahrnehmung



- Helligkeits- und Farbverarbeitung insbesondere für die Kantenerkennung und Bildung von 2-D-Primiven notwendig.
- Von Linien über Regionen zur 3D-Wahrnehmung und Objekterkennung
 - Gestaltprinzipien
 - Tiefenwahrnehmung
 - Objekterkennung & 3D-Modell

- Nach der Diskussion von Gestaltprozessen, die in die Phase des *surface based processing* einzuordnen sind, beginnt die Darstellung der Tiefenwahrnehmung und Objekterkennung
- Die Darstellung folgt Palmers Abbildung (vgl. P99, p.85.)

Das Basisdilemma der Tiefenwahrnehmung



- **Inverse projection problem**

Woher weiss der Betrachter, dass das Licht (die Reflektion) von Punkt A kommt, und nicht von einem der Punkte A_1 , A_2 oder A_3 ?

- Das Problem der inversen Projektion ist grundsätzlich nicht eindeutig lösbar. D.h. wir machen Annahmen über wahrscheinliche / plausible Lösungen.

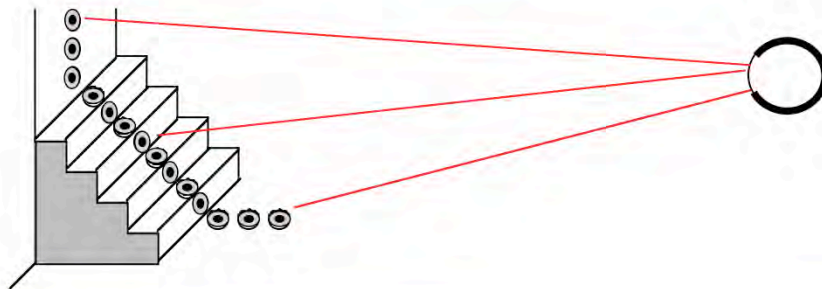
Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 147

- Das Grundproblem der Tiefenwahrnehmung:
Das retinale Bild ist zweidimensional. Aus dieser 2-D-Information soll eine Repräsentation der 3-D-Welt aufgebaut (rekonstruiert) werden.
- *Inverse projection problem*: Verschiedene Punkte der 3-D-Welt werden auf den gleichen Punkt der Retina projiziert. Werden die 3-D-Positionen der Punkte auf einer Projektionsgeraden unterschieden, so bedeutet dies, dass Tiefeninformation, also Information, die über die Orientierung der Projektionsgeraden hinausgeht, berücksichtigt wird.
- Generell gilt, dass durch Projektionsabbildungen mehrere Urbilder auf ein Bild abgebildet werden, d.h. das Resultat ist nicht eindeutig dahingehend bestimmt, aus welchem unter verschiedenen möglichen Urbildern es hervorgegangen ist.
- Abbildung nach:
Palmer, Stephen E. (1999). *Vision Science. Photons to Phenomenology*. Cambridge, MA: MIT-Press. p. 201.

Tiefenwahrnehmung: Distanz & Oberflächenorientierung

- Zwei wichtige Informationen für die Tiefenwahrnehmung:
 - Distanz des Objektes, bzw. von Objektteilen
 - Orientierung von Oberflächen des Objekts

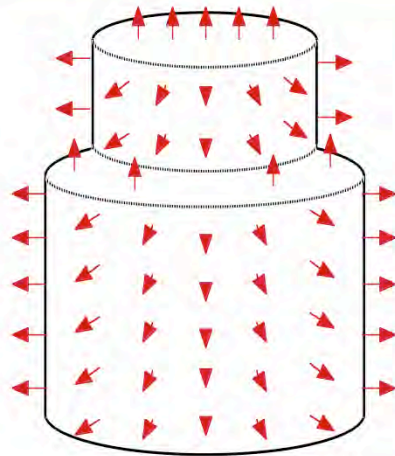


Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 148

- Wie die Abb. der vorangehenden Folie verdeutlicht, ist die Distanz zwischen BetrachterIn und Objekt, das wesentliche Kriterium, durch das verschiedene mögliche Urbilder unterschieden werden könnten.
 - Dadurch sind Verfahren zur „Distanzschätzung“ besonders wichtig für die Tiefenwahrnehmung.
- Ein weiterer wichtiger Faktor ist die Orientierung der Objektoberflächen im Hinblick auf die BetrachterIn.
- Abbildung nach:
Palmer, Stephen E. (1999). *Vision Science. Photons to Phenomenology*. Cambridge, MA: MIT-Press. p. 201.

Marrs 2.5-D-Sketch

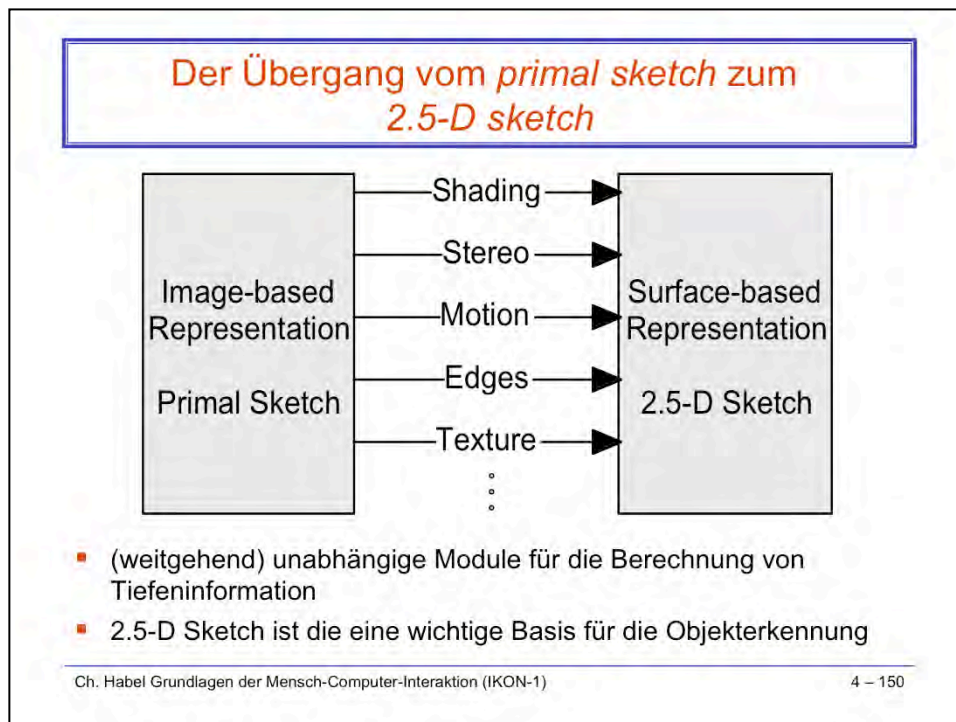


- Für jeden Punkt der Oberfläche wird durch einen Vektor die Richtung der **Oberflächennormalen** kodiert. Die Länge des Vektors kann verwendet werden, um die Distanz zum Beobachter zu kodieren.

Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

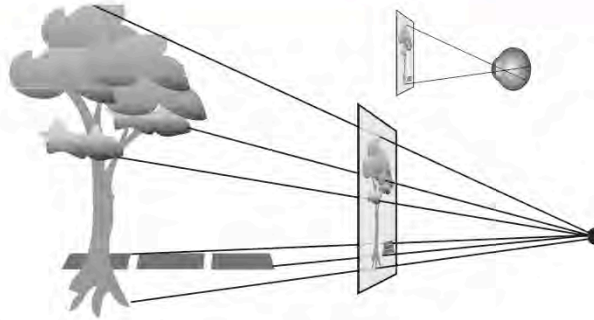
4 – 149

- Ausgehend von 2-D-Repräsentationen, die z.B. durch Projektionen auf die Retina entstanden sein können, ist die Anreicherung derartiger 2-D-Repräsentationen um Information über die Orientierung der Oberflächen eine wichtige Zwischenrepräsentation für die weitere Verarbeitung.
- Die Abb. zeigt diese zusätzliche Information in Form der Richtung der Oberflächennormalen, hier durch Pfeile dargestellt. **Oberflächennormale** ist die Senkrechte auf der Tangentialfläche.
 - Pfeile nach rechts bzw. links treten dort auf, wo die Silhouettenlinie verläuft.
 - Pfeile nach oben sind dort zu finden, wo Flächen parallel zur Ebene, auf der die BetrachterIn steht, vorliegen.
 - Pfeile, die direkt nach vorne verlaufen, treten dort auf, wo die BetrachterIn direkt auf das Objekt sieht.
- Entsprechende Repräsentationen, die als 2-D-Repräsentationen mit zusätzlicher Tiefeninformation anzusehen sind, werden als 2.5-D-sketch bezeichnet.
 Vgl. Marr, David (1982). *Vision*. New York: W.H. Freeman.
 Palmer, Stephen E. (1999). *Vision Science. Photons to Phenomenology*. Cambridge, MA: MIT-Press. p. 202f.
- Abbildung nach:
 Palmer, Stephen E. (1999). *Vision Science. Photons to Phenomenology*. Cambridge, MA: MIT-Press. p. 202.



- Abbildung nach:
Palmer, Stephen E. (1999). *Vision Science. Photons to Phenomenology*. Cambridge, MA: MIT-Press. p. 203.
- Die von David Marr vorgeschlagene Repräsentation des 2.5-D-sketch beinhaltet Tiefeninformation, d.h. die Hauptaufgabe bei der Erstellung einer derartigen 2 1/2 dimensionalen Repräsentation besteht darin, aus verschiedenen Informationsquellen Hinweise über die 3-D-Struktur der betrachteten Objekts zu erhalten und diese in den 2.5-D-sketch einzubringen.
- Die Anreicherung der 2-D-Repräsentation durch Tiefeninformation kann durch Berücksichtigung von Schatten (shading), Oberflächenstruktur (texture), die Eigenschaften der Kanten (edges), durch Stereoinformation, durch Ausnutzen von Bewegungsinformation, und weitere Quellen erfolgen. Meist werden mehrere dieser Informationsquellen ausgenutzt.
Im weiteren werden in dieser Vorlesung exemplarisch einige Aspekte der Tiefenwahrnehmung angesprochen. (Ausführlich ist dies dargestellt in Palmer, Stephen E. (1999). *Vision Science. Photons to Phenomenology*. Cambridge, MA: MIT-Press. pp. 199-253)
- Anmerkung zum Stereo-Sehen: Die Tiefenhinweise, die wir durch Sehen mit zwei Augen erhalten, sind insbesondere im Nahbereich (bis etwa 1 m) wichtig. Überprüfen lässt sich diese Behauptung durch den Vergleich von einäugigem und beidäugigem Sehen in der

Perspektive



Projektive Geometrie:

- Konvergenz von Linien, die von den BetrachterInnen ausgehen
- Abstand von der Betrachtungsebene bestimmt Grösse des Bildes

Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 151

Ware (2004)

- Ware, Colin (2004) Information Visualization: Perception for design. San Francisco, CA: Morgan Kaufmann. (2nd Edition);
Abb. entspricht Figure 8.1 (p. 261).

Hallway Illusion

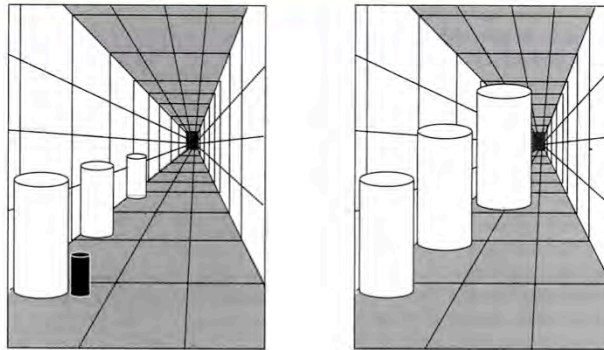
- Grössenunterschiede: „grösser ist weiter vorn“
- Textur vermittelt den Eindruck von Tiefe
 - Perspektive

Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 152

- Abbildung nach:
Palmer, Stephen E. (1999). *Vision Science. Photons to Phenomenology*. Cambridge, MA: MIT-Press. p. 316.
- An der Hallway Illusion sollen zwei für die Tiefenwahrnehmung wichtige Typen von Information illustriert werden.
 - Textur: Boden und Decke, sowie die Seiten weisen eine regelmässige Strukturierung auf, die einen Tiefeneindruck entstehen lassen, dies ist (1) eine Schar von auf einen Fluchtpunkt zulaufenden Linien, und (2) die Gruppe von Parallelen, deren Abstand sich regulär verkleinert.
 - Grösse: In der linken Abb. wird durch die unterschiedliche 2-D-Grösse der Zylinder, der Eindruck hervorgerufen, dass es sich um 3 Zylinder – annähernd – gleicher Grösse handelt, die unterschiedlich weit entfernt sind. (Der schwarze Zylinder hat die gleiche 2-D-Grösse wie der „kleinste“ weisse Zylinder. In der rechten Abb. besitzen alle drei Zylinder die gleiche 2-D-Grösse. Unsere Wahrnehmung lässt aber – ausgelöst durch die Tiefeninformation der Textur – die Zylinder unterschiedlich gross erscheinen.
- Menschliche BetrachterInnen verfügen über Erfahrungen in der Wahrnehmung, z.B. im Hinblick auf Textur und Grösse. Wir setzen diese Erfahrungen ein, wenn wir neue visuelle Eindrücke haben.

Hallway Illusion



- Größenunterschiede: „grösser ist weiter vorn“
- Textur vermittelt den Eindruck von Tiefe
 - Perspektive

Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 153

- Abbildung nach:
Palmer, Stephen E. (1999). *Vision Science. Photons to Phenomenology*. Cambridge, MA: MIT-Press. p. 316.
- An der Hallway Illusion sollen zwei für die Tiefenwahrnehmung wichtige Typen von Information illustriert werden.
 - Textur: Boden und Decke, sowie die Seiten weisen eine regelmässige Strukturierung auf, die einen Tiefeneindruck entstehen lassen, dies ist (1) eine Schar von auf einen Fluchtpunkt zulaufenden Linien, und (2) die Gruppe von Parallelen, deren Abstand sich regulär verkleinert.
 - Grösse: In der linken Abb. wird durch die unterschiedliche 2-D-Grösse der Zylinder, der Eindruck hervorgerufen, dass es sich um 3 Zylinder – annähernd – gleicher Grösse handelt, die unterschiedlich weit entfernt sind. (Der schwarze Zylinder hat die gleiche 2-D-Grösse wie der „kleinste“ weisse Zylinder. In der rechten Abb. besitzen alle drei Zylinder die gleiche 2-D-Grösse. Unsere Wahrnehmung lässt aber – ausgelöst durch die Tiefeninformation der Textur – die Zylinder unterschiedlich gross erscheinen.
- Menschliche BetrachterInnen verfügen über Erfahrungen in der Wahrnehmung, z.B. im Hinblick auf Textur und Grösse. Wir setzen diese Erfahrungen ein, wenn wir neue visuelle Eindrücke haben.



- Zimbardo explains the Ames Room illusion.
- Uploaded by [thetumblingmystic](#) on Feb 22, 2010
- <http://www.youtube.com/watch?v=hCV2Ba5wrCs>
- <http://www.zimbardo.com/>

Ames Room



Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 155

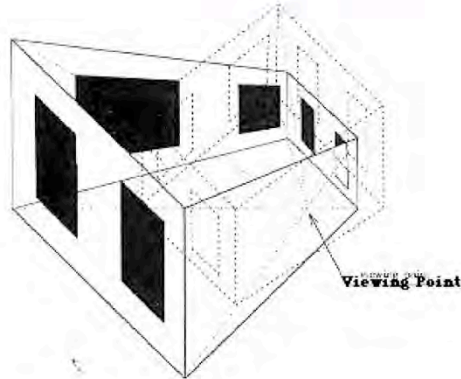
- Zum in der Vorlesung gezeigten Video: “Zimbardo explains the Ames Room illusion.” Philip Zimbardo: <http://www.zimbardo.com/>
- Video uploaded by [thetumblingmystic](http://www.youtube.com/watch?v=hCV2Ba5wrCs) on Feb 22, 2010
<http://www.youtube.com/watch?v=hCV2Ba5wrCs>
- siehe auch: Goldstein, Bruce J. (2002). Sensation and perception. Pacific Grove, CA: Wadsworth. (6th edition) [Chapt. 7 *Perceiving Depth and Size*, pp. 258f.]
- Beobachter betrachten das Innere des Raumes durch ein Guckloch (\approx Verhinderung der binokularen Wahrnehmung).
Die Wirkung (Fehlwahrnehmung) wird durch eine extreme unnatürliche Konstruktion des Raumes (siehe folgende Folie) erzielt: der Raum ist nicht quaderförmig, sondern ein schiefwinkliges Trapezoid. Dies führt zu einer Missinterpretation bei der Berechnung der Perspektive, die durch eine nicht-rechtwinklige Kachelung des Bodens (in Wirklichkeit sind es schiefwinklige Kacheln) verstärkt wird.

Siehe auch:

- http://en.wikipedia.org/wiki/Ames_room: An Ames room is a distorted room that is used to create an optical illusion. Probably influenced by the writings of Hermann Helmholtz, it was invented by American ophthalmologist Adelbert Ames, Jr. in 1934, and constructed in the following year.
- Abbildungen aus
- <http://www-graphics.stanford.edu/courses/cs99d-00/projects/JasonAnderson->

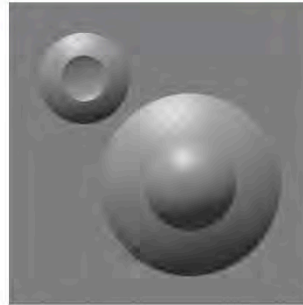
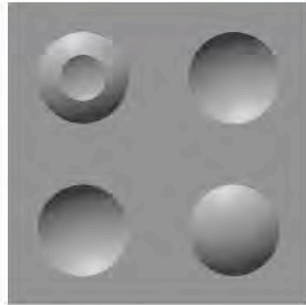
Ames Room (2)

- Guckloch verhindert Stereosehen
- Die verzerrte Geometrie des Raumes erzeugt das Bild eines Standardraumes (Quader)
 - die Ecken des Raumes scheinen äquidistant zu sein



- Zu Seckel
<http://www.illusionworks.com/>
- Abbildungen aus
- <http://www-graphics.stanford.edu/courses/cs99d-00/projects/JasonAnderson-illusions.ppt>

Shading / Schatten (1)



- Schatten bewirken die Wahrnehmung von
 - Konvexitäten
 - Konkavitäten
- Dieses basiert auf „Vorannahmen“ über die Beleuchtung

Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 157

- Abbildungen aus Ware, Colin (2004) Information Visualization: Perception for design. San Francisco, CA: Morgan Kaufmann. (2nd Edition); linke Abb. entspricht Figure 7.16 (p. 246).

Vgl. auch , Folie-6 von Lecture 6. Objects and object-based displays (<http://www.ccom.unh.edu/vislab/VisCourse/Lectures.html>)

Das Thema Schatten ist ausführlich behandelt bei Palmer, S. E. (1999). *Vision Science. Photons to Phenomenology*. Cambridge, MA: MIT-Press. p. 245.

- Schatten und Schattenwurf wird vom visuellen System häufig als Tiefeninformation „interpretiert“.
 - Die unterschiedliche Schattierung der beiden Reihen von Kreisen (linke Abb.) führt dazu, dass die oberen als konvex, d.h. kugelförmige Erhebungen, die unteren als konkav, d.h. kugelförmige Vertiefungen, interpretiert werden. (Der Effekt verändert sich entsprechend beim Umdrehen des Blattes.)
 - Der Effekt tritt nicht auf, wenn – rechte Abb. – die „Schattierung“ diskret ist.
 - Der Effekt scheint darauf zu basieren, dass wir von einer „Standardbeleuchtung“ ausgehen: Licht fällt von oben / vorne ein. Der daraus resultierende Schattenwurf, stimmt mit der oben angesprochenen Interpretation überein.
- Schattierung als Element graphischer Gestaltung, sollte stets daraufhin überprüft werden, ob die resultierende Tiefenwirkung „erwünscht“ ist. (➔)

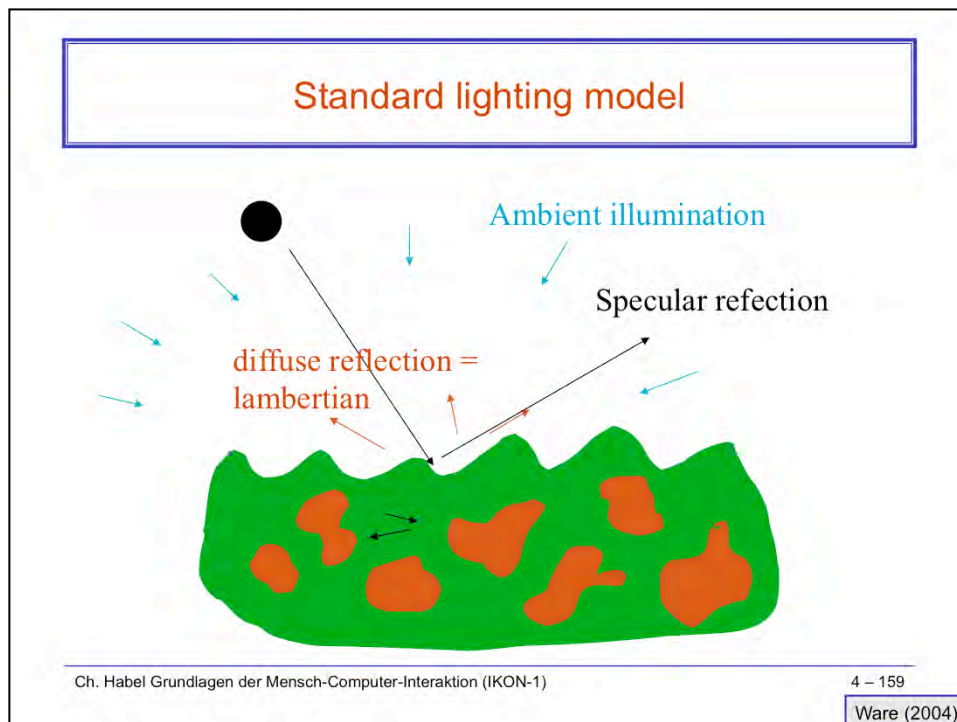
Wahrnehmung der Oberfläche und Form

- Einfaches Beleuchtungsmodell:
 - Licht kommt von oben und aus dem Unendlichen
- Shading models (Schattenmodelle)
 - Lambertian shading: Licht reflektiert an einer Oberfläche gleichmässig in alle Richtungen (diffuse reflection)
 - Specular shading: Auf einer glänzenden Oberfläche ergeben sich Glanzlichter (highlights)
 - Ambient shading: Licht kommt aus der gesamte Umgebung
 - Cast shadows: Schatten eines Objekts auf sich selbst, oder einen anderen Objekt
- Texturen – insbesondere orientierte – beeinflussen die Wahrnehmung

Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

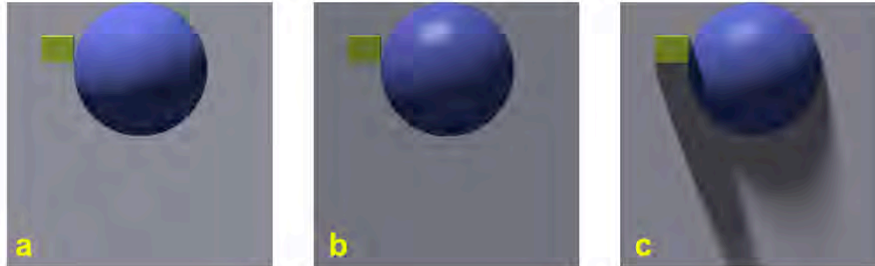
4 – 158

- Ware, Colin (2004) Information Visualization: Perception for design. San Francisco, CA: Morgan Kaufmann. (2nd Edition), p 35f & p245f



- Ware, Colin (2004) Information Visualization: Perception for design. San Francisco, CA: Morgan Kaufmann. (2nd Edition); Vgl. auch , Folie-6 von Lecture 4. Objects and object-based displays (<http://www.ccom.unh.edu/vislab/VisCourse/Lectures.html>)

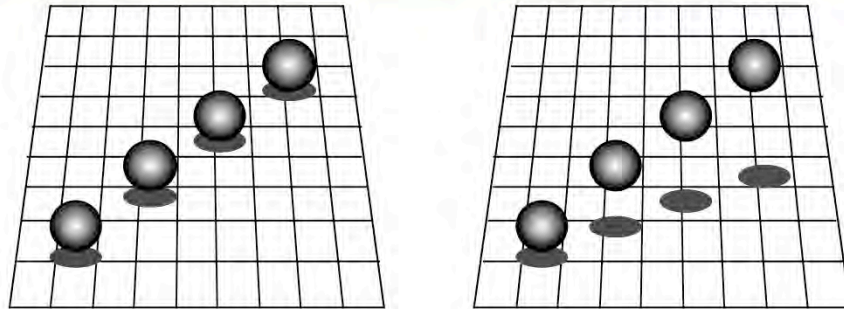
Lighting model



(a) Lambertian, (b) + specular & ambient, (c) + cast shadows

- Ware, Colin (2004) Information Visualization: Perception for design. San Francisco, CA: Morgan Kaufmann. (2nd Edition);
Abb. entspricht Figure 2.6 (p. 36).

Shading / Schatten (2)



■ Tiefenwahrnehmung und Schattenwurf

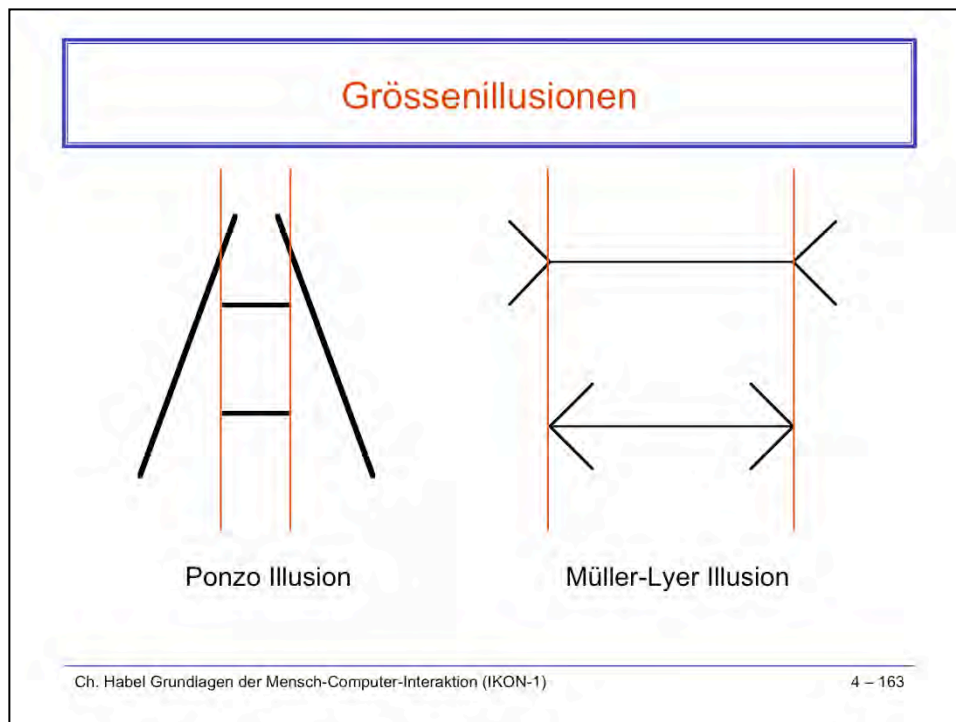
Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 161

- Abbildung nach:
Palmer, Stephen E. (1999). *Vision Science. Photons to Phenomenology*. Cambridge, MA: MIT-Press. p. 246.
- Die schattierten Kreise (Kugeln) in beiden Abbildungen sind an der gleichen Position und besitzen die gleiche Grösse.
 - In der linken Abbildungen wird eine Tiefenwirkung erzielt. Da die Objekte und ihre Schatten verbunden sind, werden die Kreise/Kugeln auf derselben Oberfläche liegende wahrgenommen. Die Textur der Grundfläche induziert Tiefenwirkung.
 - In der rechten Abbildung wird aufgrund der Unverbundenheit der drei rechten Kugeln ein „über dem Grund schweben“ wahrgenommen. Daher muss die Textur der Grundfläche nicht als Tiefeninformation wahrgenommen werden.



- Was ist Länger?
Dynamische Präsentation zur folgenden Folie



- Abbildung nach:
Palmer, Stephen E. (1999). *Vision Science. Photons to Phenomenology*. Cambridge, MA: MIT-Press. p. 8 .
- In der Ponzo-Illusion wird durch die auf einen Fluchtpunkt zulaufenden Gerade ein Tiefeneindruck erzeugt. Unsere Erfahrung der Tiefenwahrnehmung ist, dass von zwei gleichgrossen Objekten das im Bild grössere das vordere ist (mit anderen Worten, ein Objekt wird kleiner, wenn es seine Position im Raum in Richtung der Betrachterachse ändert.)
Hieraus ergibt sich, dass zwei gleichgrosse Perzepte, die auf zwei im Raum vom Betrachter unterschiedlich entfernt platzierte Objekte zurückgehen (Tiefeneindruck!), von unterschiedlich grossen Objekten verursacht werden: Das hintere ist das grössere.
- Grössen-Illusionen können z.T. bei der graphischen Gestaltung auftreten, ohne dass es beabsichtigt war.

Ponzo Illusion in der Realen Welt



Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 164

- Abbildung aus Gregory, Richard L. (2008). Emmert's Law and the moon illusion. *Spatial Vision*, 21. 407-420. Original: Fig. 1(a), p. 411.

Müller-Lyer Illusion in der Realen Welt



Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 165

- Abbildung aus Gregory, Richard L. (2008). Emmert's Law and the moon illusion. *Spatial Vision*, 21. 407-420. Original: Fig. 1(b) (cont'd), p. 412)

Ein Anwendungsfeld: Visualisierung und Wahrnehmungsprinzipien

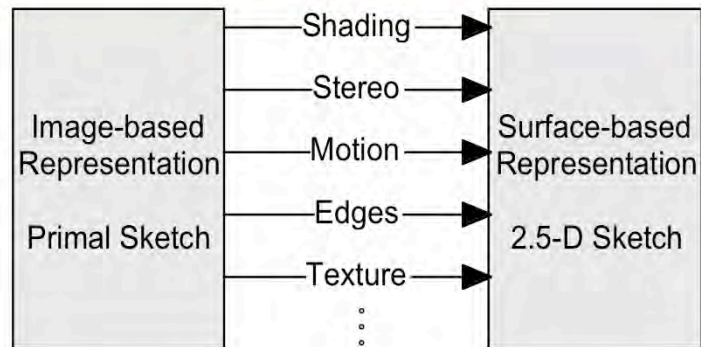
- Der Output von Visualisierungsprozessen ist der Input von Wahrnehmungsprozessen:
 - Spektrum: Von Standard-Visualisierung (Diagramm-Tools in Excel) bis zu komplexen Animationen
 - Berücksichtigung (und daher Kenntnis) von Wahrnehmungsprinzipien ist essentiell.
- Abschnitt Informationsvisualisierung in Kapitel 5

Stephen Kosslyn, ein Spezialist im Bereich der Visuellen Wahrnehmung und des Denkens über räumliche Objekte (in späteren Abschnitten der Vorlesung mehr zu seinen Arbeiten), hat in zwei Sachbüchern das Thema des Designs von Graphen behandelt.

Die wissenschaftlichen Hintergründe zu den folgenden Folien (und Beispielen) sind in den beiden Büchern ausführlich erläutert.

- Kosslyn, Stephen M. (1994). *Elements of graph design*. New York: W.H. Freeman.
- Kosslyn, Stephen M. (2006). *Graph Design for the Eye and Mind*. Oxford: Oxford University Press.

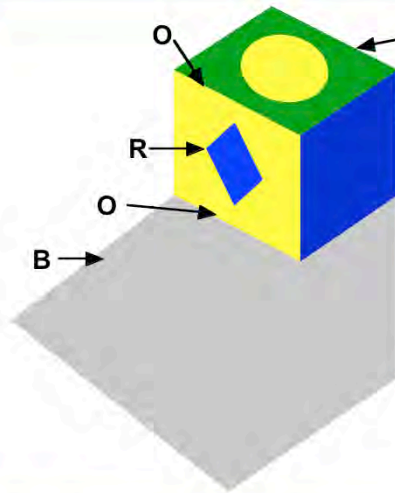
Der Übergang vom *primal sketch* zum *2.5-D sketch*



- Für die Erstellung des 2.5-D Sketchs spielen die Kanten eine zentrale Rolle.

- Abbildung nach:
Palmer, Stephen E. (1999). *Vision Science. Photons to Phenomenology*. Cambridge, MA: MIT-Press. p. 203.

Kanten-Interpretation (1): Vier Arten von Kanten



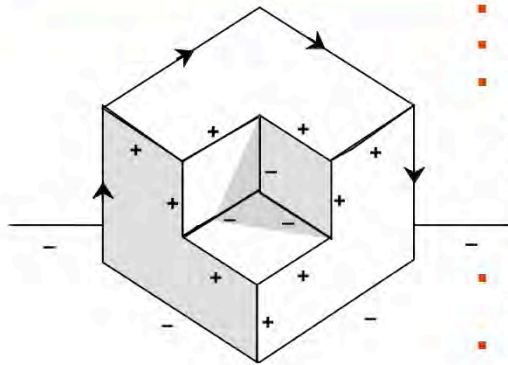
- **Orientierungskanten:** Unstetigkeiten in der Oberflächenorientierung
- **Tiefenkanten:** Räumliche Unstetigkeit zwischen Oberflächen (**Verdeckung**)
- **Beleuchtungskanten:** Unterschiede in der Beleuchtung auf einer homogenen Oberfläche
- **Reflektanzkanten:** Unterschiede in der Reflexion (Farbe, Textur, etc.)

Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 168

- Zur Kanteninterpretation:
Palmer, Stephen E. (1999). *Vision Science. Photons to Phenomenology*. Cambridge, MA: MIT-Press. p. 236–243.
- Auch bei der maschinellen Verarbeitung von Tiefeninformation ist die Kanteninterpretation zentral. Die hier vorgestellten Überlegungen gehen insbesondere auf folgende Autoren zurück:
 - Clowes, M. B., On seeing things, *Artificial Intelligence 2* (1) (1971), 79-1 16.
 - D. Huffman. Impossible objects as nonsense sentences. In B. Meltzer and D. Michie, editors, *Machine Intelligence 6*, pp. 295–324. Edinburgh U. Pr., 1971.
 - Waltz, D., (1972). Generating semantic descriptions from drawings of scenes with shadows, MAC AI-TR-271, MIT
- Ausgangspunkt ist ein 2-D-Bild, bei dem eine Kantendetektion (im Großen) erfolgreich durchgeführt werden kann. Dann gibt es vier Arten von Kanten (siehe Folie / Abb. nach Palmer p. 238):
 - Orientierungskanten sind Kanten innerhalb des Objektes bzw. zwischen Kanten, hier zwischen Würfel und Grundfläche.
 - Tiefenkanten sind Kanten zwischen Objekt und Hintergrund
 - Beleuchtungskanten sind Kanten, die z.B. durch Schattenwurf entstehen.
 - Reflektanzkanten ergeben sich u.a. durch Farbunterschiede.
- Die Zuordnung von Kanten eines Bildes zu einer dieser Typen ist zuerst einmal eine Arbeitshypothese. Es kann natürlich im Einzelfall passieren, dass wir etwas für einen Schatten halten, aber in Wirklichkeit eine Reflektanzkante vorliegt. Ebenso können wir uns irren

Kanten-Interpretation (2): Kantenmarkierung



- Konvexe Orientierungskanten +
- Konkave Orientierungskanten –
- Tiefenkanten
Pfeilmarkierung: Die nähere, verdeckende Oberfläche ist auf der rechten Seite der Kante.
- Bei n Kanten gibt es 4^n mögliche Markierungen.
- Aber nur wenige dieser Markierungen sind physikalisch möglich.

Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

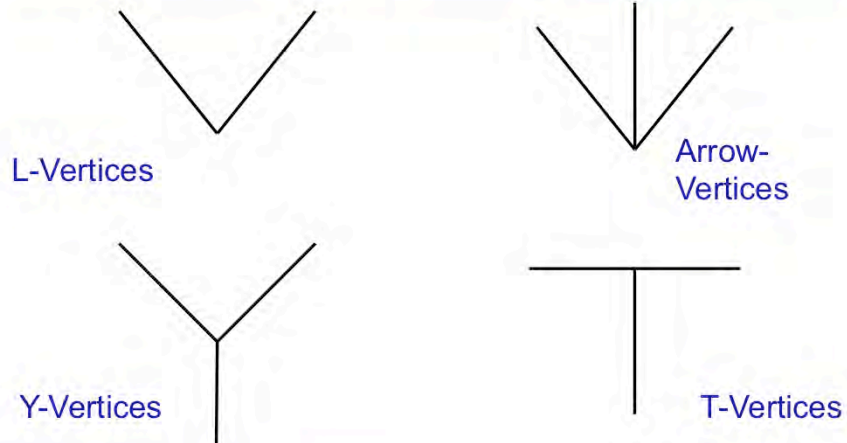
4 – 169

- Ausgangspunkt: Eine Repräsentation eines Objekts, in der nur Orientierungs- und Tiefenkanten vorkommen. [Anm. Eine Reduktion auf diese beiden Typen ist in vielen Fällen wenig problematisch.]
- Kanteninterpretation bedeutet in dieser Situation zu bestimmen, welche Subtypen vorliegen. Hierbei wird häufig eine Kantenmarkierung durchgeführt, wobei von folgenden Subtypen ausgegangen wird:
 - Orientierungskanten treten in 2 Arten auf, konvex und konkav.
 - Tiefenkanten betreffen Verdeckungskonstellationen, d.h. Konstellationen bestehend aus einem vorderen Objekt und einem hinteren Objekt (Hintergrund). Tiefenkanten können in zwei Arten verschiedener „Durchlaufrichtung“ unterschieden werden (s. Folie)

D.h., dass jede Kante (aus der Klasse der Orientierungs- und Tiefenkanten einer der vier Subtypen angehören kann).

- Die oben gezeigte Figur (Abb. nach Palmer p. 239) besitzt 20 Kanten, daher $4^{20} = 1048576$ logisch mögliche Kantenmarkierungen. Menschliche BetrachterInnen erkennen sofort eine Interpretation, nämlich die oben dargestellt.
- Die Aufgabe für eines effizientes Kanteninterpretationsverfahrens besteht darin, die Anzahl der logisch möglichen Kantenmarkierungen effizient zu reduzieren (möglichst auf eine).

Kanten-Interpretation (3): Vertex types for trihedral angles



Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 170

- Es werden vier Typen von Ecken, an denen drei Flächen (\approx trihedral) aufeinander treffen könnten, unterschieden.

Hier werden noch keine Kantenmarkierungen vorgenommen.

Die Unterscheidung von Ecken-Typen betrifft die „Perzeptuelle Erscheinung von Ecken“, nicht die „geometrische Struktur des physikalischen Gegenstandes“.

Kanten-Interpretation (4): Physikalisch mögliche Interpretationen

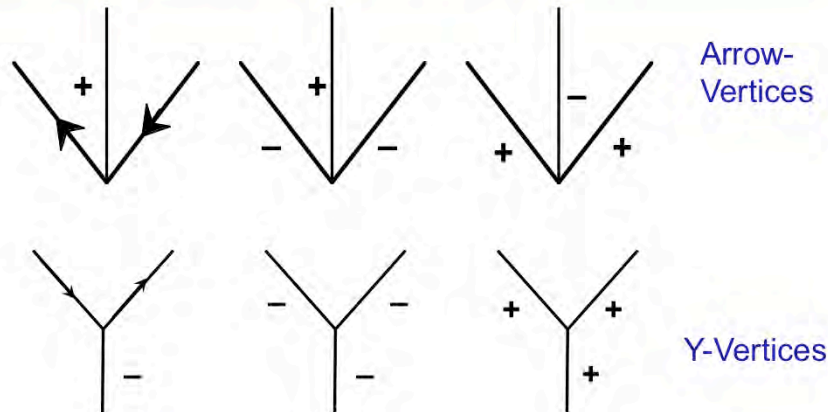
Arrow Vertices

- Mittlere Kante kann keine Tiefenkante sein
- Es existieren 3 physikalische bei $4^3 = 64$ logisch möglichen Markierungen

Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)
4 – 171

- Abb. nach Palmer p. 240 (Fig. 5.5.18)
 - Nicht alle logisch möglichen Interpretationen sind auch physikalisch möglich.
 - Beispiel: „Pfeil-Ecken“
 - Die mittlere Kante kann keine Tiefenkante sein, denn dann würde eine weitere (vierte) Kante benötigt, um das verdeckende Objekt zu bilden.
 - Die drei physikalisch möglichen Interpretationen:
 - Links: z.B., Ecke eines Würfels, der über der Grundfläche gehalten wird, vor einem Hintergrund.
 - Mitte: z.B., Ecke eines Würfels, der auf einer der Grundfläche steht.
 - Rechts: z.B., Ecke eines Würfels, von innen gesehen, etwa eine Zimmerecke.
- Vgl. die entsprechenden Ecken in der Abb. Auf der vorangehenden Folie.
- Die Interpretation des Kantengraphen eines Objekts muss logisch und physikalisch möglich sein, d.h. es ist eine konsistente Kantenmarkierung vorzunehmen.

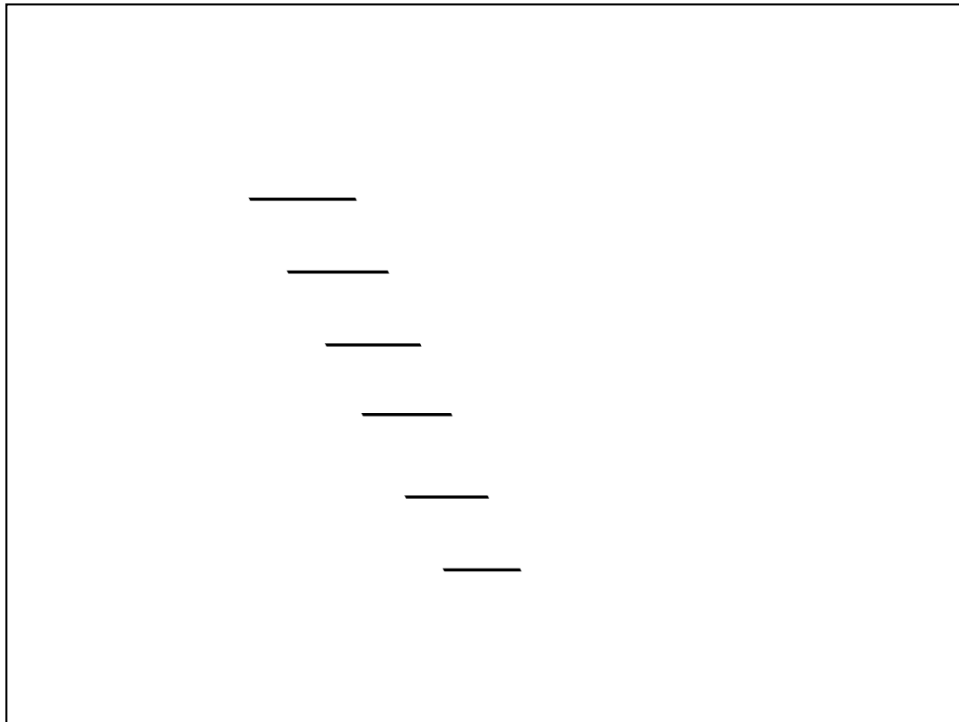
Kanten-Interpretation (5): Physikalisch mögliche Interpretationen



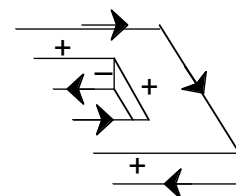
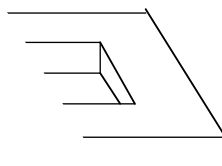
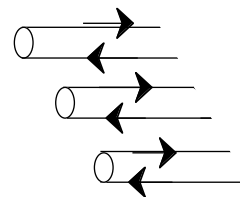
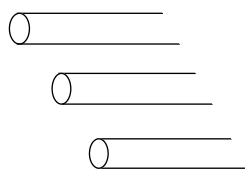
Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 172

- Abb.5.5.18 (Palmer p. 240) zeigt auch die physikalisch möglichen Interpretationen von L-vertices und T-vertices
- Gegenüberstellung von Arrow-vertices und Y-vertices
- Physikalisch und logisch konsistente Zuweisung von Vertices der verschiedenen Typen kann beim Necker-Würfel (Folie „Ambige Figuren (2)“ in zwei Weisen durchgeführt werden, d.h. zu zwei Interpretationen führen.
Die Situation beim Necker-Würfel ist – im Vergleich zu den hier verwendeten Beispielen – zusätzlich komplexer, da die einige Linien andere Linien verdecken, so dass hierdurch zusätzliche Vertex-Typen auftreten.
Vorschlag: Zum Verständnis des Phänomens ‚Kanteninterpretation‘ ist ein nochmalige Betrachtung der Folien zum Necker-Würfel (einschliesslich der Erläuterungen) nützlich.

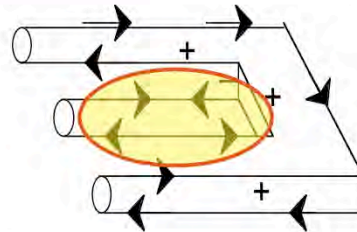
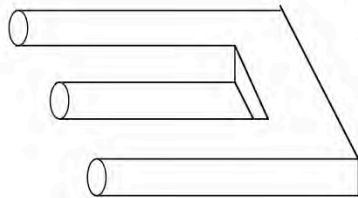


- Abb. nach Palmer p. 241
ursprünglich bei: Penrose, L.S. & Penrose, R. (1958). Impossible objects: A special type of visual illusion. *Britisch Journal of Psychology*, 49, 31–33.
- Es gibt keine konsistente, physikalisch mögliche Kantenmarkierung. Durch partielle Abdeckung lassen sich zwei konsistent markierbare Teilbereiche finden.



- Nicht jedes „unmögliches Objekt“ verfügt nur über nicht konsistente Kantenmarkierungen. Unter Umständen wird ein Objekt als „unmöglich empfunden“, weil eine konsistente Markierung sehr

Kantenmarkierungen: Impossible objects

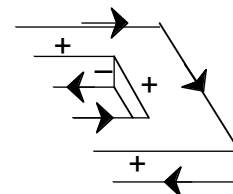
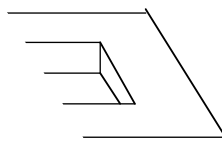
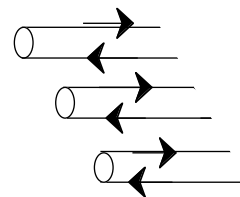
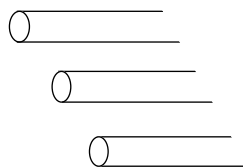


- *Devils Fork* besitzt keine konsistente, physikalisch mögliche Kantenmarkierung

Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

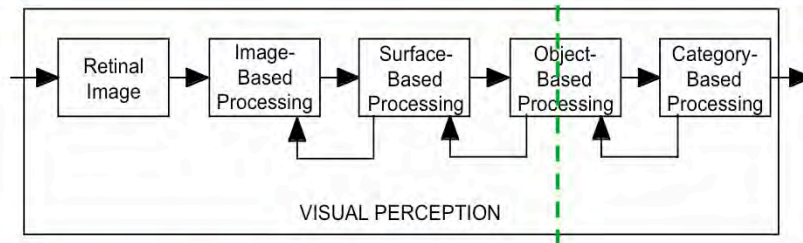
4 – 174

- Abb. nach Palmer p. 241
ursprünglich bei: Penrose, L.S. & Penrose, R. (1958). Impossible objects: A special type of visual illusion. *Britisch Journal of Psychology*, 49, 31–33.
- Es gibt keine konsistente, physikalisch mögliche Kantenmarkierung. Durch partielle Abdeckung lassen sich zwei konsistent markierbare Teilbereiche finden.



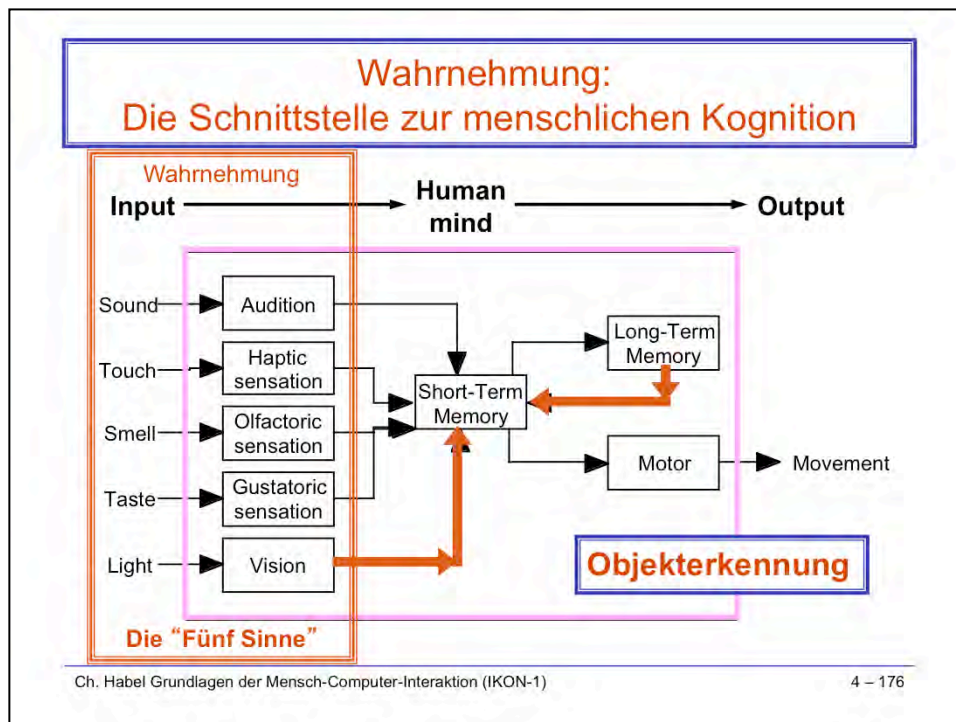
- Nicht jedes „unmögliche Objekt“ verfügt nur über nicht konsistente Kantenmarkierungen. Unter Umständen wird ein Objekt als „unmöglich empfunden“, weil eine konsistente Markierung sehr

Vier Stufen der visuellen Wahrnehmung



- Retinal image: 2-D Projektion der Umwelt
- Image based processing: Erkennen von Bildatomen, z.B. Kanten
- Surface based processing: 2-D-Primitive: Regionen,...
- Object based processing: 3-D-Primitive,
- Category based processing: Erkennen, Beziehung zum Wissen

- Die Darstellung folgt Palmers Abbildung (vgl. P99, p.85.)
- Wir befinden uns in der Vorlesung gegenwärtig beim Übergang von der vorletzten zur letzten Phase.



- Diese Folie ist eine Variante der Folie mit gleichem Titel zu Beginn des 3. Kapitels, in der hier insbesondere die kognitive Aufgabe der Objekterkennung hervorgehoben ist. Die Grundstruktur der Abbildung entspricht Figure 2.3.6 in Palmer, Stephen E. (1999). *Vision Science. Photons to Phenomenology*. Cambridge, MA: MIT-Press. (p. 75).

Objekterkennung / Objektmodelle

- Objekterkennung basiert auf Wissen über die

- Objektteile

- und deren räumliche Anordnung zueinander

- typische Ansichten

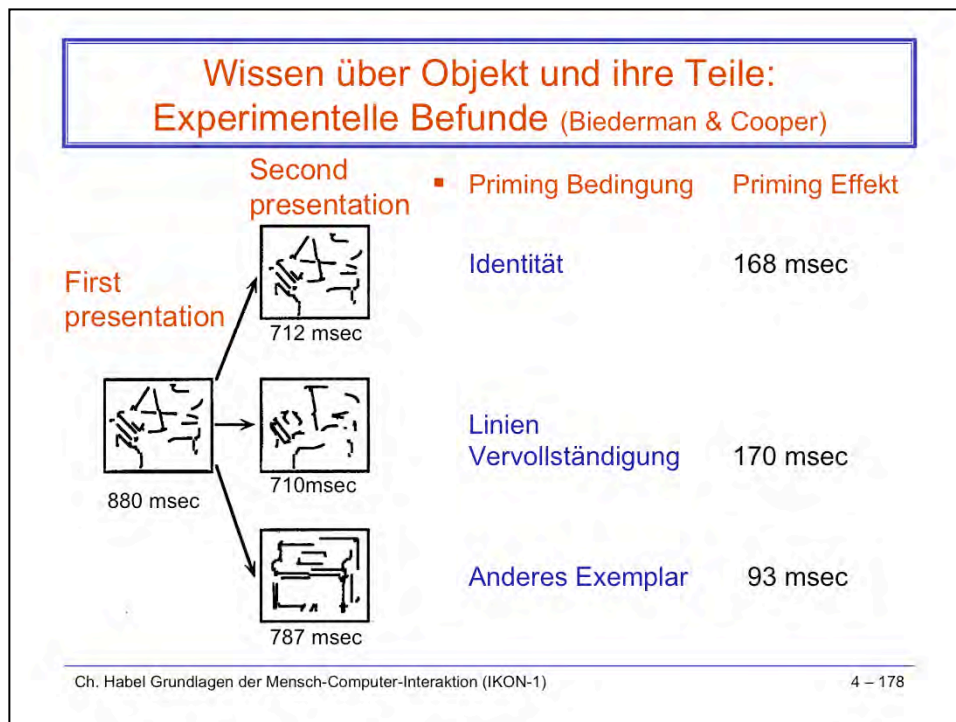
- insbesondere auch Konturen

→ Verschiedene Ansätze fokussieren auf verschiedene Typen von Wissen, das bei der Objekterkennung verwendet wird:

- Recognition by components (Biederman)

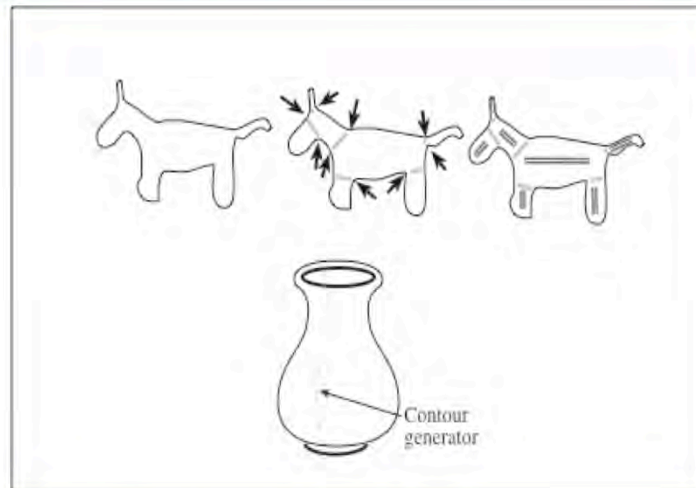
- Viewpoint specific recognition

- Objektwahrnehmung basiert auf unserem Vorwissen über Objekte. Im weiteren wird exemplarisch der Ansatz von Irving Biederman vorgestellt. Ausführliche Darstellungen der Probleme der Objekterkennung geben:
 - Biederman, I. (1995). Visual object recognition. In Stephen M. Kosslyn & Daniel N. Osherson (eds.), *Visual Cognition – An Invitation to Cognitive Science (2nd ed.) Vol. 2.* (pp. 121–165). Cambridge, MA: MIT Press.
 - Palmer, Stephen E. (1999). *Vision Science. Photons to Phenomenology.* Cambridge, MA: MIT-Press. p. 416–452.
- Die im folgenden erläuterten Untersuchungen verwenden als Versuchsmaterial Strichzeichnungen bzw. Umrisszeichnungen.
 - Die Resultate können auf andere Arten von Abbildungen und Wahrnehmungen übertragen werden (z.B. Photos, reale Objekte).
 - Die Wahrnehmung von Zeichnungen etc. spielt im Bereich der Präsentation von Objekten in der Mensch-Maschine-Interaktion eine grosse Rolle.



- Abbildung aus:
Palmer, Stephen E. (1999). *Vision Science. Photons to Phenomenology*. Cambridge, MA: MIT-Press. p. 427.
- Präsentiert wurden in diesem Experiment Strichzeichnungen von Objekten. Die Aufgabe der Versuchspersonen (VP) bestand darin, eine Klassifikation der Objekte vorzunehmen, d.h. so schnell wie möglich zu sagen, was für ein Objekt sie sehen.
Die Strichzeichnungen waren derart, dass jeweils nur die Hälfte der Konturlinien vorhanden war; daher konnten Paare komplementärer Zeichnungen konstruiert werden, so dass in der einen Zeichnung jeweils die Teillinien existierten, die in der anderen fehlten (Reihe „second presentation“ oberste und mittlere Abb.)
 - Im ersten Durchgang wurden den VPs jeweils eine der Variante pro Objekt präsentiert und dabei die Klassifikationszeit bestimmt. (vgl. Abb.: 880 msec./ Objekt: Flügel)
 - Im zweiten Durchgang wurden den VPs drei unterschiedliche Zeichnungen zur Klassifikation vorgelegt. (Jede VP erhielt nur eine dieser Varianten. Es wurden 3 VP-Gruppen gebildet; Varianten wurden gleichmässig auf die Gruppen verteilt.)
 - Untersucht wurde, inwieweit die Präsentation des ersten Durchgangs die Klassifikation im zweiten Durchgang **erleichtert** (Priming effect). Zwischen den beiden, komplementären Zeichnungsvarianten des gleichen Objekte bestand kein signifikanter Erleichterungs-unterschied. Die Präsentation einer Zeichnung eines verwandten Exemplars (hier Klavier) führte zu eine wesentlich geringere Erleichterung.

Konturen und strukturelle Skelette



Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 179

- Ware, Colin (2004) Information Visualization: Perception for design. San Francisco, CA: Morgan Kaufmann. (2nd Edition); Fig. 7.7 & Fig. 7.8 (p. 236)
Vgl. auch, Folie-13 von Ware 's
Lecture 6. Objects and object-based displays (<http://www.ccom.unh.edu/vislab/VisCourse/Lectures.html>)
- Die hier verwendeten Abbildungen und insbesondere die zugrunde liegenden Theorien der visuellen Wahrnehmung gehen auf David Marr und Kollegen zurück:
 - Marr, D. & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three-dimensional shape. In *Proc. of the Royal Society, Series B*, 200. (pp. 269–294).
 - Marr, David (1982). *Vision*. New York: W.H. Freeman.

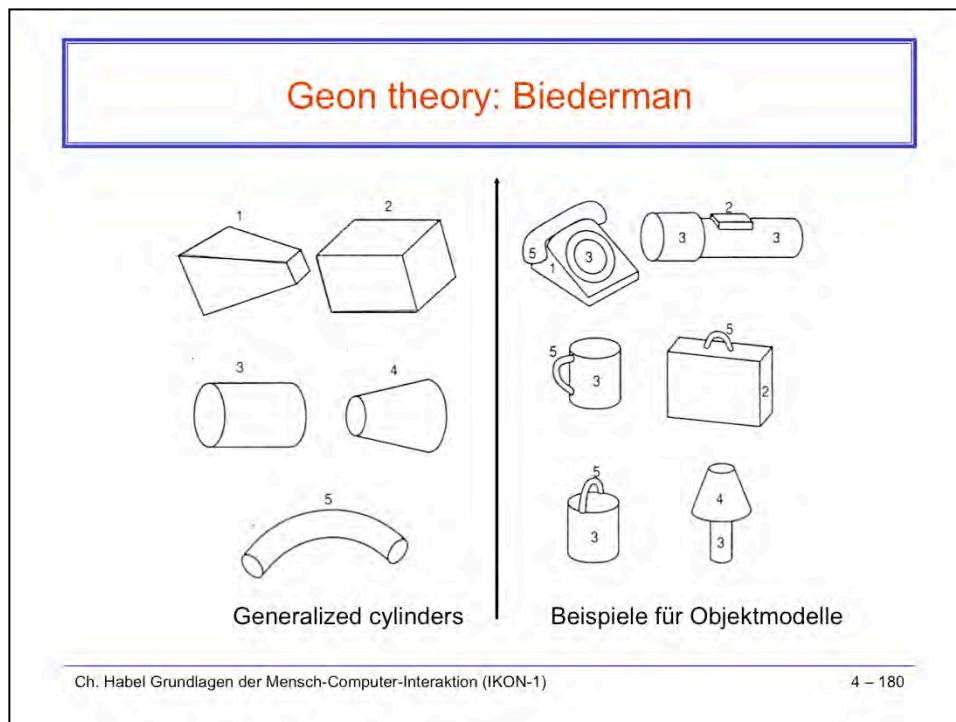
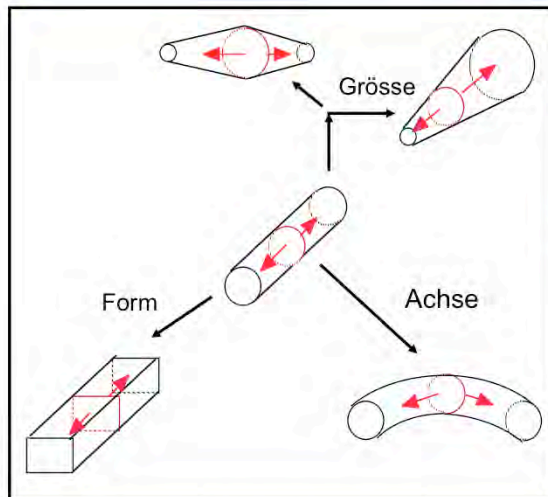


Abbildung aus:

Biederman, I. (1995). Visual object recognition. In Stephen M. Kosslyn & Daniel N. Osherson (eds.), *Visual Cognition – An Invitation to Cognitive Science (2nd ed.)* Vol. 2. Cambridge, MA: MIT Press. P. 140

- Wie wird nun das Wissen über die Form eines Objekts bzw. die Teile, aus denen ein Objekt besteht, gespeichert? Aufbauend auf Ansätzen aus dem Informatikbereich „Computer Vision“ schlagen viele KognitionswissenschaftlerInnen vor, eine Darstellung durch Geons (generalized cons / generalisierte Zylinder) zu verwenden.
- Die linke Abbildung zeigt einige Typen generalisierter Zylinder (ausführlicher siehe die folgende Folie)
- Biederman argumentiert dafür, dass die meisten Objekte, die wir in unserer Umwelt vorfinden, als Aggregate von wenigen (meist 3 – 4) Geons beschrieben werden können. (Beachte: Für eine Klassifikation / Objekterkennung kommt es im ersten Verarbeitungsschritt nicht darauf an, Details zu verwenden, sondern darauf, charakteristische Formen (grober Massstab) zu berücksichtigen. (Vgl. hierzu auch Folie „Hierarchie von 3-D Modellen“).
- Die rechte Abbildung zeigt einige Beispiele von Artefakten als Aggregate von Geons.
- Aus Geons aufgebaute Objektmodelle, die von Biederman als wesentliche Voraussetzung für die Objekterkennung angenommen werden, sind Teil des Gedächtnisses.
- Geon-artige Darstellungen entsprechen in ihrem Abstraktionsniveau dem, was schnell verarbeitbar ist. Hieraus lassen sich Leitlinien für die schematische Darstellung von Objekten, etwa in der Mensch-Computer-Interaktion ableiten. (siehe im Abschnitt Problemlösung dieser Vorlesung.)

Geons (2)



Attribute der Geon-Konstruktion:

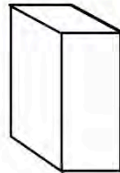
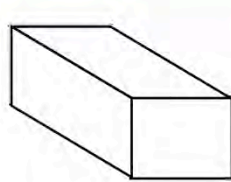
- Form der generierenden Fläche
- Form der Achse
- Grösse: Konstanz vs. Expansion vs. Kontraktion

Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 181

- Abbildung aus:
Biederman, I. (1995). Visual object recognition. In Stephen M. Kosslyn & Daniel N. Osherson (eds.), *Visual Cognition – An Invitation to Cognitive Science (2nd ed.)* Vol. 2. Cambridge, MA: MIT Press. P. 141
- Geons sind parametrisierte geometrische Objekte. Ein Geon ist konstruierbar / beschreibbar, durch den Teil des 3-dim-Raumes, der bei der Transformation eines 2-dim Querschnittes längs einer Achse im 3-d-Raum „überstrichen“ wird. Die verwendeten Parameter sind:
 - Achse und ihre Form: gerade vs. gekrümmte Achsen
 - Querschnitt: Hier wird davon ausgegangen, dass eine kleine Klasse von 2-dim. Grundformen verwendet wird. In der Abbildung sind die beiden Formen „kreis“ und „rechteck“ dargestellt.
 - Grösse des Querschnitts. Diese kann sich während der erzeugenden Transformation verändern. Es wird davon ausgegangen, dass für effizient verarbeitbare Geons einer der folgenden Typen von Querschnittsveränderungen vorliegt: keine Veränderung, konstante bzw. gleichmässige Vergrößerung/Verkleinerung bzw. erst Vergrößerung, dann Verkleinerung (oder umgekehrt)
 - Aufgabe: Überlegen Sie, wie die für Bananen typische Form durch ein Geon dargestellt werden kann.
 - Geons sind eine wichtige Konzeption für die Computer Graphik / Visualisierung.

Warum sind Geons perzeptuell grundlegend?

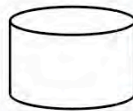


- Konstanz von Eigenschaften unter veränderten Bedingungen der Wahrnehmung:
→ nonaccidental properties

Quader

- 3 Mengen paralleler Kanten
- 1 innere Y-Ecke
- 3 äussere Pfeil-Ecken

Zylinder



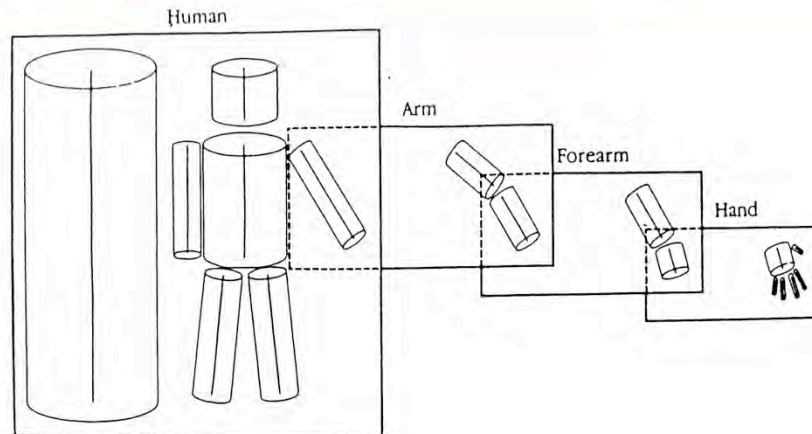
- 2 parallele gerade Kanten
- 2 parallele gekrümmte Kanten
- 2 tangentielle Y-Ecken

Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 182

- Abbildung nach:
Palmer, Stephen E. (1999). *Vision Science. Photons to Phenomenology*. Cambridge, MA: MIT-Press. p. 436.
- Klassifikation / Wiedererkennung von Objekten muss unter veränderten Wahrnehmungsbedingungen funktionieren. Das bedeutet, dass wir bei der Klassifikation solche Indizien / Kriterien einsetzen sollten, die sich bei Veränderung (etwa der Betrachtungsperspektive) nur gering, bzw. in gut rekonstruierbarer Weise, verändern, d.h. im wesentlichen invariant sind.
- Die obigen Abbildungen zeigen, dass etwa Quader und Zylinder unter verschiedenen Perspektiven Invarianz im Hinblick auf die folgenden Eigenschaften aufweisen:
 - Mengen paralleler Kanten (betrifft sowohl gerade als auch gekrümmte Kanten)
 - Anzahl von Y-Ecken / Pfeil-Ecken

Hierarchie von 3-D Modellen: Marr

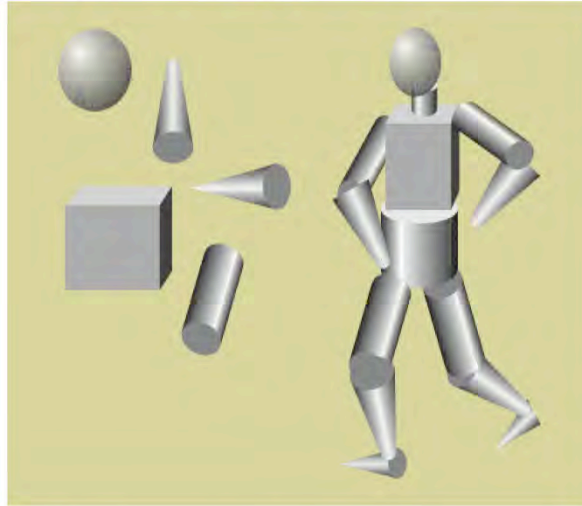


Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 183

- Für die Erkennung / Klassifikation werden schematische / abstrahierte Formbeschreibungen benötigt. Und nur auf diese kognitive Aufgabe bezieht sich die oben formulierte Annahme, dass für eine Objektbeschreibung ein Aggregat von wenigen Geons ausreichend sei.
- Komplexe räumliche Darstellungen von Objekten basieren auf anderen Darstellungsmechanismen. Hier wird überwiegend davon ausgegangen, dass hierarchische Repräsentationen eingesetzt werden.
So kann die Form eines Menschen in der größten Darstellungsstufe durch einen Zylinder repräsentiert werden (dies entspricht dem, was wir bei schlechten Sichtbedingungen, etwa Nebel, Dämmerung, grosse Distanz, wahrnehmen können). In der ersten Verfeinerung Unterscheiden wir 6 Zylinder, nämlich für Kopf, Rumpf und 2 Paare von Extremitäten. Die Arme können als Kombination von jeweils 2 Zylindern dargestellt werden,
Prinzip der sukzessiven Verfeinerung
- Abbildung aus: Marr, David (1982). *Vision*. New York: W.H. Freeman. P. 306

Geons als Bausteine der Generierung von Modellen



Ch. Habel Grundlagen der Mensch-Computer-Interaktion (IKON-1)

4 – 184

- Geons werden in der Informatik nicht nur für die Objekterkennung eingesetzt, sondern auch für Visualisierung von Objekten. Der Einsatz von Geon-basierten Objektmodellen ist nützlich für die Beschreibung der „Grobstruktur“ von Objekten, einschliesslich der Modellierung von Bewegungen (etwa im Fall der geometrischen Beschreibung eines Menschen, der Bewegungen, des Kopfes, der Arme und der Beine). Die „Feinstruktur“ insbesondere die Oberfläche kann dann in einer zweiten Phase berechnet werden (Haut, Haare, Kleidung, etc.)
- Ware, Colin (2004) Information Visualization: Perception for design. San Francisco, CA: Morgan Kaufmann. (2nd Edition); Fig. 7.5 (p. 235) Vgl. auch , Folie-11 von Lecture 6. Objects and object-based displays (<http://www.ccom.unh.edu/vislab/VisCourse/Lectures.html>)

Grundlagen der Mensch-Computer Interaktion

4. Kapitel

Der Mensch

- Ein- und Ausgabe: Wahrnehmung und Handeln
 - Wahrnehmung: Modalitäten der Wahrnehmung
 - Visuelle Wahrnehmung
 - Haptische Wahrnehmung & Hand-Arm-Motorik
- Gedächtnis
- Verarbeitung von Information & Informationsnutzung