Goldstein Kapitel 5

Einleitung:

2004 fand ein Rennen für robotische Fahrzeuge statt, sie sollten ohne menschliches Eingreifen 142 Meilen durch die Wüste zurücklegen und dabei Hindernisse erkennen und umgehen. Das beste Fahrzeug kam 7,4 Meilen weit.

2005 erreichten bei einem ähnlichen Rennen schon 23 Fahrzeuge das Ziel.

* Fortschritt erkennbar, aber Computern Objektwahrnehmung beizubringen ist offensichtlich sehr schwierig.

Schwierigkeiten bei der Objektwahrnehmung:

1. Abbild des Stimulus auf der Retina ist mehrdeutig („Problem der inversen Projektion“) 🡪 unendlich viele verschiedene Objekte erzeugen das gleiche Bild auf der Netzhaut.

2. Objekte können unscharf oder teilweise verdeckt sein.

3. Objekte sehen aus verschiedenen Blickwinkeln unterschiedlich aus. Menschen ist trotzdem im Allgemeinen klar, dass es sich um ein und das selbe Objekt handelt (Fähigkeit zur Blickwinkelinvarianz).

4. Oft ist es schwer zu entscheiden, ob ein Farbunterschied durch unterschiedliche Lichteinstrahlung oder durch unterschiedliches Material hervorgerufen wird.

All diese Probleme bereiten Computern erhebliche Schwierigkeiten, wogegen Menschen sie mit Leichtigkeit lösen und sich der „Rechenleistung“ die dafür benötigt wird nicht einmal bewusst sind!

* **Wieso sind Menschen Computern in dieser Hinsicht so weit überlegen?**

**Strukturalismus** nach Wilhelm Wundt:

(Wundt gründete 1879 das erste experimentalpsychologische Labor.)

„Wahrnehmung kommt durch die Kombination elementarer Erfahrungen zustande (so wie viele Punkte gemeinsam ein Bild erzeugen)“

Dieser Ansatz beeinflusste die **Gestaltpsychologie**:

Blinken nacheinander zwei Bilder oder Lichter an verschiedenen Stellen auf, so wird Bewegung wahrgenommen, obwohl faktisch keine vorhanden ist 🡪 Scheinbewegung (z.B. Leuchtreklame mit durchlaufendem Schriftzug)

Solche Scheinwahrnehmungen waren ein Problem für den Strukturalismus, das sie eigentlich nicht vorhanden sind und nicht auf elementaren Wahrnehmungen basieren.

Die Gestaltpsychologie suchte nach neuen Erklärungen. Grundsatz: **Das Ganze ist mehr als die Summe seiner Teile!**

🡪 Aufstellung von Regeln der Wahrnehmungsorganisation

Gestaltprinzipien

Wie organisieren wir kleine Teile zu einem großen Ganzen?

**9 Gestaltprinzipien:**

**1. Das Prinzip der Prägnanz / der guten Gestalt**

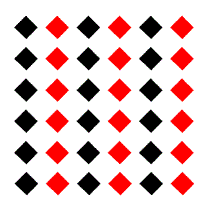
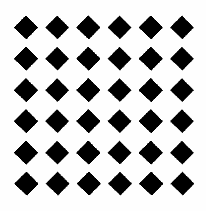
Jedes Reizmuster wird so gesehen, dass die resultierende Struktur so einfach wie möglich ist.

*Beispiel: Das Olympia-Symbol wird als verschlungene Ringe gesehen und nicht als Anordnung kompliziertere Formen.*

**2. Das Prinzip der Ähnlichkeit**

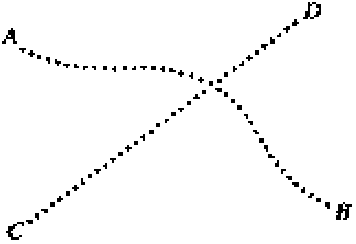
Ähnliche Dinge erscheinen zu Gruppen geordnet.

*Beispiel: In der linken Abbildung werden entweder Spalten oder Zeilen oder beides wahrgenommen (nicht festgelegt), in der rechten dagegen nimmt man durch die Farbgebung automatisch Spalten war. Dieses Gesetz greift auch bei auditiven Wahrnehmungen.*



**3. Das Prinzip des guten Verlaufs**

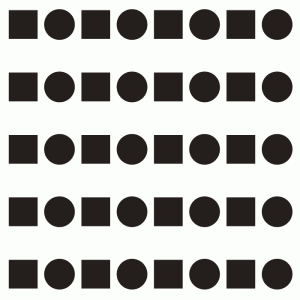
Punkte, die als gerade oder sanft geschwungene Linien gesehen werden, wenn man sie verbindet, werden als zusammengehörig wahrgenommen. Linien werden tendenziell so gesehen, als folgten sie dem einfachsten Weg.



*Beispiel: A ist mit B verbunden und nicht mit D.*

**4. Das Prinzip der Nähe**

Dinge, die sich nahe beieinander befinden, erscheinen als zusammengehörig.



*Beispiel: Man sieht Zeilen, obwohl das Prinzip der Ähnlichkeit Spalten suggerieren würde 🡪 Prinzip der Nähe ist hier stärker.*

**5. Das Prinzip des gemeinsamen Schicksals**

Dinge, die sich in die gleiche Richtung bewegen, erscheinen als zusammengehörig.

*Beispiel: Vögel, die in die gleiche Richtung fliegen, bilden einen Schwarm.*

**6. Das Prinzip der Vertrautheit**

Dinge bilden mit größerer Wahrscheinlichkeit Gruppen, wenn die Gruppen vertraut erscheinen oder etwas bedeuten.



*Beispiel: In dem Bild sind 12 Gesichter versteckt (viel Spaß bei Suchen :P). Die Landschaftselemente, die diese bilden werden aber erst als Gruppe wahrgenommen, wenn das Gesicht erkannt, also die Elemente als Gesicht organisiert wurden. Dann fällt es sogar schwer, sie noch als einzelne Steine, Bäume etc. zu sehen.*

**7. Das Prinzip der gemeinsamen Region**

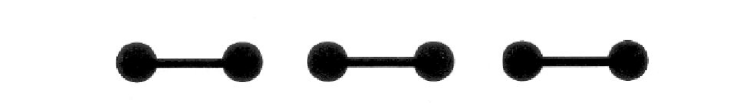
Elemente, die innerhalb einer gemeinsamen Region liegen, werden zusammengruppiert.



*Beispiel: Obwohl die Punkte in den Ellipsen weiter auseinander liegen als der Punkt in der jeweils nächsten Ellipse, werden sie durch die Umrandung als eine Gruppe wahrgenommen.*

**8. Das Prinzip der Verbundenheit von Elementen**

Verbundene Elemente innerhalb einer Region mit gemeinsamen visuellen Charakteristiken wie Helligkeit, Farbe, Textur oder Bewegung werden als Einheit gesehen.



*Beispiel: Selbes Prinzip wie oben.*

**9. Das Prinzip der zeitlichen Synchronizität**

Visuelle Ereignisse, die zur selben Zeit auftreten, werden als zusammengehörig wahrgenommen.

*Beispiel: Gleichzeitig blinkende Lichter werden als zusammengehörig wahrgenommen.*

**Gruppierungen auf Neuronaler Ebene:**

Bilder mit Gestaltprinzipien lösen eine stärkere Neuronale Antwort im Cortex aus, als ungeordnete.

(Beispiel mit Affe im Buch, ist ohne die Abbildung aber nicht erklärbar)

**Gestaltprinzipien = Gesetze?**

Früher wurden die Prinzipien als Gestaltgesetze bezeichnet, heute sieht man sie aber als „Heuristiken“.

Heuristik = Faustregeln, die die scheinbar naheliegendste Lösung für ein Problem liefern.

Stehen im Gegensatz zu

Algorithmen = Prozedur, die mit Sicherheit zur Lösung des Problems führt.

*Beispiel: Katze im Haus verschwunden. Algorithmus: Akribisch Zimmer für Zimmer durchsuchen. Heuristik: Die Lieblingsverstecke der Katze absuchen.*

Nachteil der Heuristik: Unter Umständen wird das Ziel nicht erreicht. Vorteil: In großen Prozentsatz der Fälle wird das Ziel erreicht und das auch deutlich schneller als mit dem Algorithmus.

Da das Überleben in der Umwelt häufig von der Schnelligkeit der Wahrnehmung und Objekterkennung abhängt, ist es sinnvoll, dass das menschliche Wahrnehmungssystem heuristisch arbeitet.

Die Gestaltprinzipien beschreiben, wie Teile eines Stimulus zu einem Objekt zusammengruppiert werden. Wie aber ist es im umgekehrten Fall? Wie erkennen wir, was Objekt und was Hintergrund ist? 🡪 Figur-Grund-Trennung / Problem der Perzeptuellen Gliederung

Perzeptuelle Gliederung

Merkmale von Figuren:

* Leichter im Gedächtnis zu behalten, dinghafter
* Steht vor dem Hintergrund
* Die Kontur wird als zur Figur gehörig wahrgenommen (Besitz der Kontur)

Merkmale des Grunds

* Ungeformtes Material
* Hinter der Figur (logischerweise)

Faktoren, die die Sicht als Figur begünstigen:

* Der untere Teil einer Darstellung,
* symmetrische Formen,
* flächenmäßig kleinere Formen,
* vertikal oder horizontal ausgerichtete Formen,
* bedeutungshaltige Formen (die an etwas Bekanntes erinnern)

werden bevorzugt als Figur wahrgenommen.

Auf neuronaler Ebene:

An Affen wurden Versuche gemacht: Ein Muster aus gleich ausgerichteten kleinen Strichen wurde dargeboten und die Feuerrate eines Neurons untersucht. Waren die Striche überall gleich, feuerte das Neuron nicht. Waren die Striche rund um und in dem rezeptiven Feld des Neurons aber anders ausgerichtet als die Übrigen und daher Teil einer „Figur“ vor einem Hintergrund, so feuerte das Neuron (obwohl sich an der Stimulus in seinem rezeptiven Feld nichts geändert hatte).

🡪 Wie kann das Neuron Informationen erhalten, die außerhalb seines rezeptiven Felds liegen und sich davon beeinflussen lassen? („Kontextuelle Modulierung“)

🡪 Diese Informationen müssen von einer „höheren“ Ebene im Wahrnehmungsprozess kommen, es gibt also neben der Informationsweiterleitung von „niedrigen“ zu „hohen“ Ebenen (feedforward-Prozesse) auch einen Informationsfluss in die andere Richtung (Feedback)!

Die moderne Forschung zur Objektwahrnehmung

Heutige Wahrnehmungspsychologen nutzen gestaltspychologische Konzepte noch als Ausgangspunkt für weitere Forschung, gehen aber darüber hinaus:

1. Messung ist wichtiger als Beschreibung
2. Konzentrieren sich auf die Mechanismen, die für die Wahrnehmung verantwortlich sind. Gestaltprinzipien erklären nicht, *wie* die beschriebenen Effekte zustande kommen.

**Welche neuronale Verschaltung ist Grundlage dieser Leistungen?**

4 Grundlegende Fragen:

**1. Warum antwortet das visuelle System auf bestimmte Stimuli stärker als auf andere?**

Das menschliche Wahrnehmungssystem ist auf die Umwelt angestimmt durch a) die Evolution über Jahrtausende hinweg und b) die erfahrungsabhängige Plastizität des Gehirns, die durch alltägliche Erfahrungen Neuronen formt.

Regelmäßigkeiten in der Umwelt führen zu angepasster Wahrnehmung:

Viele gerade Konturen, vorwiegend horizontale und vertikale, auch gekrümmte Konturen eher „glatt“, nicht zackig 🡪 halb verdeckte Objekte laufen eher gerade weiter, bilden außerhalb des Blickfelds keine überraschenden Windungen 🡪 Prinzip des guten Verlaufs

Menschen sind empfindlicher für Horizontalen und Vertikalen (Obliqueneffekt), im Visuellen Cortex mehr Neuronen die auf horizontale und vertikale Linien antworten.

**2. Muss eine Figur vom Grund getrennt werden, bevor wir Objekte erkennen können?**

These der Gestaltpsychologen: Trennung von Figur und Grund ist einer der ersten Schritte im Wahrnehmungsprozess.

ABER: Die Figur-Grund-Konstellation wird von der Bedeutungshaltigkeit eines Stimulus beeinflusst (Flächen, die etwas darstellen, werden eher als Figur gesehen als willkürlich geformte) 🡪 Der Erkennungsprozess muss entweder vor oder zu dem selben Zeitpunkt wie die Figur-Grund-Trennung erfolgen.

**3. Wie erkennen wir Objekte aus verschiedenen Blickwinkeln?**

2 Theorien:

1) Modell der Zerlegung in elementare Teilkörper:

Dreidimensionale Objekte werden als Kombination von einfachen dreidimensionalen Körpern („volumetrische Merkmale“) gesehen (Ein Mensch besteht zum Beispiel aus einer Anzahl von Zylindern).

Weiterentwicklung dieses Modells von Irving Biederman: Rekognition-by-Components-(RBC-)Theorie

Er nennt volumetrische Merkmale „Geons“ (geometric icons), schlägt 36 davon vor, aus denen bereits fast alle Objekte zusammengebaut werden können.

Die Geons haben blickwinkelinvariante Merkmale, z.B. sind die 3 parallelen Kanten eines Quaders aus fast allen Blickwinkeln sichtbar. Nur selten aus wenigen Perspektiven sieht man sie nicht, Biederman bezeichnet dies als „zufälliges Merkmal“ von Geons, da es nur in seltenen Fällen auftritt.

Wenn zufällige Merkmale auftreten erschwert das die Objekterkennung.

Außerdem besitzen Geons blickwinkelinvariante Unterscheidbarkeit, d.h. sie sind fast immer von anderen Geons zu unterscheiden.

Die Grundidee ist, dass wir Objekte erkennen, sobald wir ihre Geons wahrnehmen können (Prinzip der Rekonstruktion elementarer Teilkörper)

Pro:

Zeigt, dass wie Objekte an einer recht kleinen Anzahl einfacher Teilkörper bereits erkennen können.

Contra:

Man kann Pferde von Zebras unterscheiden, obwohl sie in etwa den selben Körperbau haben 🡪 andere Merkmale wie Textur, Bewegung und Tiefe werden in Biedermans Theorie vernachlässigt.

2) Modelle der kanonischen Ansichten:

Abgespeicherte zweidimensionale Ansichten von Objekten aus unterschiedlichen Blickwinkeln ermöglichen die Erkennung.

Pro:

Erklärung dafür, dass die Blickwinkelinvarianz nicht immer gegeben ist, gerade bei unbekannten Objekten ist sie manchmal nicht vorhanden (es sind ja auch noch keine Bilder davon abgespeichert).

Die „Wahrheit“ liegt vermutlich irgendwo zwischen den beiden Ansätzen, wir nutzen wahrscheinlich Elemente von beiden Methoden für die Erkennung dreidimensionaler Objekte.

**4. Wie verarbeitet das Gehirn Informationen über Objekte?**

Bei Objekterkennung starke Aktivität im inferotemporalen (IT) Cortex und auch anderen Arealen.

3 Versuche dazu:

1) Versuch mit einem Affen (schon wieder, die armen Affen!):

Bekam für das linke Auge das Bild einer Sonne, für das rechte das eines Schmetterlings dargeboten 🡪 binokulare Rivalität, führt dazu, dass man ein paar Sekunden das eine und ein paar Sekunden das andere wahrnimmt.

Affe zog an Hebel, wenn er den Schmetterling sah, so dass man wusste, wann er was wahrnahm. Ein Neuron wurde beobachtet, bei dem zuvor festgestellt wurde, dass es auf das Schmetterlingsbild hin feuerte.

Zog der Affe am Hebel, d.h. nahm er den Schmetterling wahr, feuerte auch das Neuron, sah der Affe das Sonnenscheinbild, hörte das Neuron auf.

* Der Stimulus auf den Netzhäuten des Affen veränderte sich nicht, d.h. Veränderungen in der Wahrnehmung passieren im Gehirn, wie man auch an der veränderten Feuerrate des Neurons erkennen kann.

2) Versuch mit Menschen, die Harrison Ford erkennen sollten.

Was geschieht im Gehirn, wenn wir ein Gesicht identifizieren?

Problem: Von Person zu Person sehr große Schwankungen bezüglich der genauen Position eines bestimmten Areals im Gehirn. Daher: Vortest mit Stimuli mit und ohne Gesichtern, dadurch wurden die Regionen, die auf Gesichter reagierten, bestimmt. Nun konnte man sich auf dieses Areal konzentrieren.

Personen bekommen für 50 msec ein Bild: Entweder Harrison Ford, eine andere Person oder eine Zufallstextur (ohne Gesicht). Da die Wahrnehmung eines Stimulus immer für 250 msec nach dem Verschwinden des Stimulus bestehen bleibt (Persistenz des Sehens), erschien nach dem Bild ein Maskierungsreiz aus zufälligen Linien (so ist der Stimulus nur 50 und nicht 300 msec zu sehen), dann sollten die Personen sagen, was sie gesehen hatten (Harrison Ford, eine andere Person oder kein Gesicht).

Ergebnis: Die neuronale Aktivierung war am größten, wenn Harrison Ford erkannt wurde, niedriger wenn er nicht erkannt wurde und kaum vorhanden wenn gar kein Gesicht erkannt wurde.

* Neuronales Antwortverhalten hängt nicht nur von dem Stimulus ab, sondern auch von der Reaktion der Versuchsperson.

Identifikation eines Objekts: Starke Antwort

Entdecken des Objekts: Schwächere Antwort

Fehlschalgen des Entdeckens: Keine Antwort

3) Objektidentifikation, wieder an Affen:

Bekannt, dass im IT Cortex Areale für die Formwahrnehmung liegen, von dort gelangen Informationen zum präfrontalen (PF) Areal des Frontallappens.

Untersuchung des Antwortverhaltens von Neuronen in beiden Arealen:

Durch Morphing wurden 2 Kategorien von Stimuli erstellt: Katze, Hund und Mischformen mit mal mehr, mal weniger Anteil Katze bzw. Hund.

Affe wurde trainiert zwischen Katze und Hund zu unterscheiden.

Verzögerte Übereinstimmungsprüfung: 1. Beispielphase: Beispielstimulus wird dargeboten, entweder Katze oder Hund. 2. Verzögerungsphase: Kurze Verzögerung ohne Stimulus. 3. Testphase: Teststimulus (Hund oder Katze) 🡪 Frage: gehörten beide Stimuli zur selben Kategorie?

Ergebins:

IT-Neuron antwortet in Phase 1 stärker auf Hund als auf Katze, in Phase 3 aber auf beides ähnlich.

PF-Neuron antwortet in Phase 1 auf beide Stimuli etwa gleich, stärker auf Hund in Phase 2 und 3.

* IT-Neuronen antworten unterschiedlich auf die Darbietung von Katzen und Hunden 🡪 Formwahrnehmung. Visuelle Neuronen, antworten auf Wahrnehmung.
* Das PF-Neuron antwortet unterschiedlich wenn der Affe sich an Hunde oder Katzen erinnert und dann eine Entscheidung über den Stimulus trifft. Verhaltensbezogene Neuronen, steuern Verhalten.

**All die Versuche befassen sich mit der Frage, welchem Aspekt der Objektwahrnehmung ein bestimmtes Areal dient.**

Die Intelligenz der menschlichen Objektwahrnehmung:

Warum sind Menschen Computern bei der Objektwahrnehmung so überlegen?

* Intelligenz!
* Verhaltensbezogene Intelligenz, die auf unserem Wissen über die Umwelt basiert
* Physiologische Intelligenz, die auf der Abstimmung unseres perzeptuellen Systems auf die Umwelt basiert

**Frühe Ideen einer perzeptuellen Intelligenz:**

Likelihood-Prinzip: Wir nehmen das Objekt wahr, das die größte Wahrscheinlichkeit dafür aufweist, das von uns empfangene Reizmuster verursacht zu haben (bei mehrdeutigen Reizen)

**Moderne Ideen zur perzeptuellen Intelligenz:**

Versuch von Palmer zeigt, dass über eine Szenerie vermitteltes Wissen die Wahrnehmung beeinflussen kann.

Licht-von-oben-Heuristik: Wir nehmen automatisch an, dass Licht von oben kommt, da dies in unserer Umwelt meist der Fall ist.

Fazit:

Computern ist es zwar schon gelungen, das Wüstenrennen zu bestehen, also einer einprogrammierten Strecke zu folgen und Hindernissen in gewissem Rahmen auszuweichen, doch damit sie andere Aufgaben wie Gesichtserkennung aus verschiedenen Blickwinkeln etc. meistern können, müssen sie die menschliche perzeptuelle Intelligenz kopieren, die wir ganz automatisch nutzen.