

Visuelle Welt

Günther Kebeck

Vorwort

„Von dem, was du erkennen und messen willst,
musst du Abschied nehmen, wenigstens auf eine Zeit.
Erst wenn du die Stadt verlassen hast, siehst du,
wie hoch sich ihre Türme über die Häuser erheben.“

Friedrich Nietzsche

Distanz als Mittel der Erkenntnis – für eine Forscherin oder einen Forscher ein vertrauter Gedanke. Doch aus Abstand resultiert auch eine Reduktion von Komplexität, und eine zu starke Vereinfachung birgt die Gefahr der Verfälschung. Was könnte die angemessene Distanz sein, um einen komplexen Forschungsgegenstand wie das menschliche Sehen darzustellen?

Dieses Buch möchte ein Verständnis für die erstaunlichen Leistungen unseres visuellen Systems wecken. Ein Blick aus der Entfernung ist sicher notwendig. Doch sollte er – in einer Abwandlung des Zitats von Friedrich Nietzsche – mit „Spaziergängen durch die Stadt“ verbunden sein. Denn die Erfahrung lehrt, dass nicht der Abstand selbst, sondern die Abstandsvariation die Erkenntnis fördert. Bei der Leserin oder dem Leser wird das Interesse vorausgesetzt, sich gleichermaßen mit theoretischen Überlegungen wie mit konkreten Beispielen zu beschäftigen. Im Mittelpunkt stehen die „Gesetze des Sehens“, die Ordnung der visuellen Welt. Grundlegende Prinzipien der visuellen Informationsverarbeitung, die sich in so unterschiedlichen Bereichen wie der Form-, Raum- oder Bewegungswahrnehmung finden.

In gewisser Weise ist die Wahrnehmungsforschung gegenüber anderen Disziplinen im Vorteil. Nicht selten kann sie zeigen, was sie behauptet. Bilder und eigene Erfahrungen erlauben es, das gewünschte Phänomen hervorzurufen. Es ist deshalb bei der Lektüre dieses Buches besonders wichtig, dem Sehen zunächst mehr zu vertrauen als dem Wissen. Die Fälle, in denen unsere Wahrnehmung von unserem Wissen abweicht, sind besonders produktiv. Sie geben einen Einblick in das faszinierende Zusammenspiel von Licht, Auge und Gehirn. Wahrnehmungsforschung

ist vom Ansatz her interdisziplinär. Das macht es den Leser*innen nicht immer leicht, weil sie sich zunächst mit einer für sie fremden Begrifflichkeit vertraut machen müssen. Der Lohn besteht darin, über die hier angesprochenen Fragen eigenständig nachdenken zu können.

Dieses Buch ist Resümee einer langjährigen Forschungstätigkeit. Einige der hier vorgestellten Überlegungen wurden zuvor einzeln und in anderer Form publiziert. Dies gilt insbesondere für die Abschnitte I.3 (vgl. *Die Ordnung der visuellen Welt*), IV.1 - IV.3 (vgl. *Gliederung und Gruppierung*), VIII.1 - VIII.3 (vgl. *Experimentelle Ästhetik*) sowie die Abgrenzung von Bild- und Umgebungswahrnehmung (vgl. *Bild und Betrachter*). Für die sorgfältige Lektüre von Abschnitten oder auch des gesamten Manuskriptes danke ich meinen Kolleginnen und Kollegen Dr. Max Bruchmann, Dr. Paul Bürkner, Prof. Dr. Markus Lappe, Dr. Hennig Schroll, Prof. Dr. Ricarda Schubotz, Prof. Dr. Dirk Vorberg und Prof. Dr. Pienie Zwitserlood. Sie haben durch ihre wohlwollenden und kritischen Anmerkungen einen wesentlichen Beitrag zur jetzt vorliegenden Fassung geleistet. Leonie Schröder und Philipp Jabold unterstützten mich bei der Korrektur des Textes sowie der Erstellung der Abbildungen und Verzeichnisse. Ihnen allen sei herzlich gedankt.

Münster, im März 2019

Günther Kebeck

Inhalt

Vorwort	2
Kapitel I: Gesetze des Sehens	6
1. Beschreibungen.....	10
2. Aufgaben.....	17
3. Prinzipien.....	21
3. Optimale Wahrnehmung	33
4. Ideale Beobachterin und idealer Beobachter.....	46
Kapitel II: Visuelle Reize und neuronale Prozesse	53
1. Ökologische Optik.....	55
2. Optische Bilder.....	59
3. Transduktion und Transformation	63
4. Skotopisches und photopisches Sehen.....	66
5. Retinotopie Organisation.....	70
6. Modularität und Mehrdimensionalität.....	74
Kapitel III: Aufbau der visuellen Welt	79
1. Kontraste.....	80
2. Texturen.....	82
3. Figuren	85
4. Objekte	86
5. Szenen	90
Kapitel IV: Gliederung und Gruppierung	94
1. Gestaltgesetze	94
2. Transformationsgüte.....	96
3. Gliederung des Ganzfeldes	99
Kapitel V: Phänomene und Heuristiken	129
1. Amodalität	131
2. Transparenz.....	135

3. Multistabilität.....	138
4. Schatten.....	144
5. Farbausbreitung.....	151
6. Biologische Bewegung.....	154
7. Kausalität	160
8. Visuelle Illusionen	165
Kapitel VI: Beständigkeit der visuellen Welt.....	170
1. Helligkeits- und Farbkonstanz.....	172
2. Größenkonstanz	181
3. Formkonstanz	184
Kapitel VII: Analyse und Konstruktion.....	190
1. Autonome Fahrzeuge.....	192
2. Maschinelle und biologische Wahrnehmung	202
3. Modelle der visuellen Informationsverarbeitung.....	204
4. Bildbasierte Verarbeitung.....	208
5. Oberflächenbasierte Verarbeitung	215
6. Objektbasierte Verarbeitung.....	219
7. Kategorienbasierte Verarbeitung.....	223
8. Ambiguitätsreduktion durch Ergänzung.....	230
Kapitel VIII: Ästhetische Wahrnehmung.....	244
1. Ordnung und Komplexität.....	246
2. Gestaltgüte	251
3. Schönheit als Produkt der Informationsverarbeitung	258
4. Wahrnehmung von Kunstwerken.....	288
Glossar.....	293
Literatur	305
Abbildungen.....	319

Kapitel I: Gesetze des Sehens

Obwohl Sehen der vielleicht grundlegendste Erkenntnisprozess des Menschen ist, sind wir nicht geübt, ihn zu reflektieren. Das Sehen ist uns so selbstverständlich, dass wir keinen Anlass haben zu fragen, wie unsere *visuelle Welt* entsteht. Diese hat ihren Ausgangspunkt in der physikalischen Welt. Die Grundlage der Stimulation ist das von den Oberflächen reflektierte Licht. Zu dieser Außenwelt haben wir aber keinen direkten Zugang. Der Lichtreiz muss zunächst in ein neuronales Signal umgewandelt werden. Dieses Aktivitätsmuster und nicht der physikalische Reiz bildet die Grundlage für den Perzeptionsvorgang. Dennoch sollte es Korrespondenzen geben zwischen der Außenwelt und unserem Erleben.

Die *retinalen Bilder* der beiden Augen sind keinesfalls klar und eindeutig. Sie sind unvollständig, instabil und mehrdeutig (vgl. Abbildung 1). Doch wir erleben die Welt als vollständig, stabil und in sich kohärent. Wir bewegen uns in ihr sicher und schnell. Wir führen unsere Handlungen präzise aus und sind in der Lage, neue Erkenntnisse zu gewinnen. Wie ist das möglich? Wie erstellen wir ein in fast allen Situationen unseres Lebens adäquates Bild der Wirklichkeit?

Für den Prozess des Sehens gibt es feste Regeln, die sich zwischen Individuen kaum unterscheiden. Entgegenkommende Autofahrer*innen müssen sicher sein, dass ihre visuellen Welten in allen wesentlichen Belangen übereinstimmen und ihre Wahrnehmung den tatsächlichen Straßen-, Beleuchtungs- und Entfernungsverhältnissen entspricht. Jede andere Vorstellung hätte ein Chaos zur Folge und wäre lebensgefährlich. Das menschliche Sehen ist evolutionär geprägt. Seine Regeln haben sich über einen langen Zeitraum als Anpassung an die Lebensbedingungen entwickelt. Die individuelle Erfahrung spielt nur eine sekundäre Rolle.

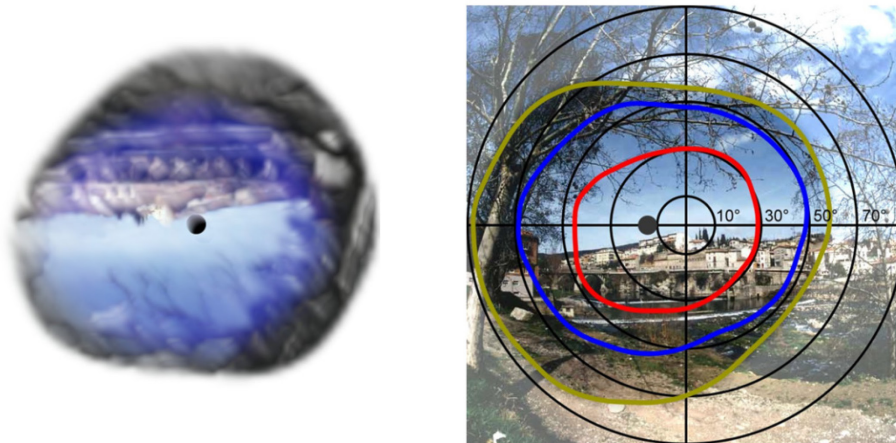


Abbildung 1: Vergleich von retinalem Bild (links) und wahrgenommenem Bild (rechts). Beide „Bilder“ sind Simulationen. In Wirklichkeit existieren lediglich neuronale Korrelate. Berücksichtigt wird bei der Simulation des „retinalen Bildes“, dass es invertiert ist, noch keine Konturen enthält, die Sensitivität für verschiedene Wellenlängen unterschiedlich ist und es an der Stelle des „blinden Flecks“ keine Reizung gibt.

Die Wahrnehmung von Bildern, die derzeit einen so großen Stellenwert hat, ist aus evolutionärer Perspektive eher unbedeutend, ein Spezialfall, für den unser visuelles System nicht geschaffen wurde. Im Alltag dominierte lange Zeit die Umgebungswahrnehmung. Dort haben wir es nicht mit Bildern, sondern mit Objekten zu tun. Viele dieser Objekte bewegen sich ebenso wie die Beobachterin oder der Beobachter. Das Sehen ist eingebunden in funktionale Zusammenhänge. So dient es der Orientierung im Raum, der eigenen Fortbewegung und der Ausführung oder Kontrolle von Handlungen. Deshalb weist die visuelle Wahrnehmung des Menschen enge Beziehungen zu den übrigen Sinnesmodalitäten und zur Motorik auf.

Vor allem die Forschung der letzten dreißig Jahre hat gezeigt, dass die über die Sinnesorgane eingehenden Informationen lediglich die Datengrundlage für eine Eigenleistung des Gehirns bilden. Die Wahrnehmung entsteht über eine Reihe von Zwischenschritten. Das

Ergebnis lässt sich nicht alleine aus den Eigenschaften der physikalischen Welt vorhersagen. „Gesetze des Sehens“ setzen ein Verständnis über die Arbeitsweise des Gehirns voraus. Die Frage von Koffka (1935, S.79) „Why do things look as they do?“ ist auch heute keine triviale Frage. Sie lässt sich nur mit einer näheren Untersuchung der Prozesse der visuellen Informationsverarbeitung beantworten (für detaillierte Darstellungen der einzelnen Phasen vgl. Chalupa und Werner, 2004; Werner und Chalupa, 2013).

Es wird sich zeigen: Das Sehen ist eine erstaunliche Leistung, ein streng gesetzmäßiger Vorgang von hoher Produktivität. Die Forschung verfügt aktuell über ein breites und gut gesichertes Wissen in Bezug auf viele der zugrundeliegenden Mechanismen, vor allem auf Teilprozesse. So verstehen wir den Strahlengang des Lichts im Auge oder die Reizverarbeitung in der Retina. Wir haben detaillierte Vorstellungen davon, wie Kanten oder Konturen entstehen, wann wir Bewegung sehen und wann nicht. Durch neue Forschungsmethoden steigt dieses Wissen gegenwärtig sprunghaft. Aber es ist bisher vor allem ein Wissen im Detail. Was fehlt, ist ein *universelles Modell* des Sehens, eine Vorstellung, wie unser Bild von der Welt entsteht. Hier stehen wir erst am Anfang.

Ein Modell des Sehens lässt sich nicht einfach aus der Empirie ableiten. Es ergibt sich nicht als Summe alltäglicher Beobachtungen oder experimenteller Befunde. Vielmehr ist es eine Theorie. In der langen Geschichte der Wahrnehmungsforschung hat es vielfältige Versuche gegeben, eine solche umfassendere Theorie zu entwickeln. Die „Theorie der unbewussten Schlüsse“ von Hermann von Helmholtz in der Mitte des 19. Jahrhunderts, die „Gestalttheorie“ zu Beginn des 20. Jahrhunderts und die „ökologische Optik“ von James Jerome Gibson in der zweiten Hälfte des 20. Jahrhunderts sind markante Beispiele. Ihnen ist es gelungen, der Forschung neue Blickwinkel auf die grundlegenden Fragen zu eröffnen: Wie nehmen wir Objekte, wie Räumlichkeit, wie Bewegung wahr? Bei aller

historischen Wertschätzung begegnen wir diesen Ansätzen heute auch mit Skepsis.

Die Forschung hat mithilfe verschiedener Messtechniken der funktionellen Neurobiologie neue Fragen aufgeworfen und Einsichten über die zugrundeliegenden Prozesse gewonnen: An welcher Stelle im Gehirn wird zu welchem Zeitpunkt welche Information verarbeitet? Die Befunde sind zum Teil noch widersprüchlich oder schwer einzuordnen. Aktuelle Zusammenfassungen des Forschungsstandes gelingt es heute, entweder Modelle zu formulieren, die eine hohe Auflösung, aber einen geringen Geltungsbereich haben (z. B. sich auf die Mustererkennung oder bestimmte Formen der Bewegungswahrnehmung beziehen) oder Rahmenmodelle aufzustellen, die zwar einen großen Geltungsbereich haben, dafür jedoch eine sehr allgemeine Beschreibung liefern. Ein Modell, das beide Anliegen verknüpft, ist derzeit nicht in Sicht.

Auch dieses Buch unternimmt keinen Versuch, ein universelles Modell des Sehens zu entwerfen. Das bestehende Wissen soll stattdessen neu gegliedert und zusammengefasst werden. Durch eine veränderte Ordnung und Perspektive, so die Hoffnung, könnte das Verständnis des Sehprozesses wachsen und als Vorarbeit für den Entwurf eines Modells dienen, dessen Konturen am Horizont bereits erkennbar sind. Der Zielsetzung, einen Einblick in den Stand der Forschung zu geben und gleichzeitig den Diskurs zwischen unterschiedlichen Fachrichtungen wie Biologie, Physiologie, Psychologie, Physik, Informatik und Philosophie zu fördern, dient die Verwendung einer Sprache, die nicht nur von Spezialisten verstanden wird.

Traditionelle Lehrbücher gliedern ihre Texte nach den Sinnesmodalitäten, innerhalb der visuellen Modalität nach Themen wie Helligkeit, Farbe, Form, Größe, Tiefe und Bewegung. Diese etablierten Themenfelder korrespondieren mit Merkmalen des Wahrnehmungsreizes bzw. mit den

phänomenalen Qualitäten der Reize. Eine solche Systematik nach Stimulusmerkmalen bzw. Qualitäten hat Vorteile. Sie entspricht den historischen Forschungsfeldern. Wahrnehmungsforscher*innen beschäftigen sich selten gleichzeitig mit Objektwahrnehmung und Farbwahrnehmung. Die einzelnen Themenfelder stehen in der Geschichte der Wahrnehmungsforschung nebeneinander. Hinzu kommt, dass es in den letzten Jahrzehnten viele Hinweise auf einen modularen Aufbau des visuellen Systems gab. Vor allem in frühen Verarbeitungsstufen werden einzelne Reizmerkmale getrennt verarbeitet. Von daher legitimiert sich eine separate Darstellung. Nicht zuletzt entspricht diese horizontale Gliederung unserem Sprachgebrauch und unserer Alltagserfahrung. Im Vordergrund stehen Objekte und ihre Qualitäten. Die Außenwelt wird in Termini phänomenaler Eigenschaften beschrieben: Das Objekt ist „klein“, „blau“ und „weit entfernt“.

Doch um zu verstehen, warum wir die Welt so sehen, wie wir sie sehen, ist diese Betrachtungsweise hinderlich. Sie stellt die Forschungsergebnisse additiv nebeneinander. Eine vertikale Gliederung, wie sie hier erprobt wird, soll die Zusammenhänge in den Vordergrund stellen. Frei nach dem Motto Kurt Lewins: „Nichts ist so praktisch wie eine gute Theorie“, ist das Ziel eine *theoretische Wahrnehmungspsychologie* auf empirischer Basis. Vollständigkeit ist nicht intendiert. In der exemplarischen Darstellung werden die Beispiele so gewählt, dass sie unterschiedliche Aspekte der Arbeitsweise des visuellen Systems und der Theoriebildung innerhalb der Wahrnehmungsforschung veranschaulichen.

1. Beschreibungen

Ein Text über das Sehen benötigt eine verbindliche Begrifflichkeit. Die Wahrnehmungsforschung verfügt bisher über keine einheitliche Terminologie. Selbst bei Verwendung derselben Begriffe kommt es nicht selten zu einem unterschiedlichen Verständnis. Zunächst sollen daher die wichtigsten Konzepte eingeführt und erläutert werden.

Reize

Die Forschung zur visuellen Wahrnehmung beginnt nicht beim Subjekt. Ausgangspunkt ist das Licht. Seine Energie wird im Sinnesorgan in ein neuronales Signal umgewandelt, das vom Gehirn weiterverarbeitet wird. Anders als die Vorstellung bedarf die Wahrnehmung einer permanenten Reizung des Sinnesorgans. Wahrnehmung bedeutet zunächst *Analyse der Stimulation*. Zur Beschreibung der Stimulation hat es sich als sinnvoll erwiesen, zwischen distalen und proximalen Reizen zu unterscheiden. Als *distale Stimuli* werden die Objekte der Außenwelt bezeichnet. Im Mittelpunkt steht das von den Oberflächen dieser Objekte reflektierte Licht. Der *proximale Stimulus* ist die Verteilung der Lichtenergie am Sinnesorgan. Er beschreibt die Intensitäten und Wellenlängen, die das Auge erreichen und für die das Auge empfindlich ist. Für die *Stimulation* und damit für die weitere Informationsverarbeitung sind letztlich nur die proximalen Verhältnisse entscheidend. Distaler und proximaler Stimulus können sich erheblich unterscheiden. Objekte können durch andere Objekte verdeckt oder unterschiedlich weit vom Auge entfernt sein. Die Analyse der Reize ist Bestandteil der Wahrnehmungsforschung. Ihre physikalische Beschreibung alleine reicht nicht aus, weil sie das Sinnessystem nicht einbezieht. Vielmehr ist eine *ökologische Optik* im Sinne von James Jerome Gibson notwendig. Sie untersucht den Reiz in seiner Bedeutung für den Organismus und stellt Fragen wie: Kann eine Oberfläche betreten oder ein Gegenstand hochgehoben werden?

Prozesse

Die Arbeitsweise des visuellen Systems wird durch Prozesse beschrieben. Wie wird die gegebene Stimulation analysiert? Wie wird auf einer unvollständigen und mehrdeutigen Datengrundlage eine in sich stimmige visuelle Welt erstellt? Welche Regeln steuern diesen Vorgang? Zu diesen neuronalen Prozessen gibt es auf dem Wege der Introspektion keinen Zugang, es ist aussichtslos, Proband*innen nach ihnen zu befragen.

Während die Wahrnehmung phänomenal als ein einheitlicher Prozess erscheint, muss sie unter einer naturwissenschaftlichen Perspektive in räumliche wie zeitliche Verarbeitungsschritte gegliedert werden. Beschreibungen auf dieser Ebene der Prozesse beziehen sich häufig auf einzelne Reizmerkmale. Sie weisen direkte Beziehungen zu den neurophysiologischen Grundlagen und den psychophysischen Experimenten der Wahrnehmungsforschung auf. Das Wissen über die neuronalen Mechanismen kann in Form von *Heuristiken* zusammengefasst werden (vgl. auch Hoffman, 2001). Darunter werden Regeln verstanden, die häufig, aber nicht immer zu einer richtigen Lösung führen. Außer in der Lösungswahrscheinlichkeit unterscheiden sie sich in der Anwendbarkeit. Lässt sich eine Heuristik auf viele Ereignisse anwenden, dann wird sie als *leistungsstarke Heuristik* bezeichnet.

Heuristiken sind Rekonstruktionen der Forschenden über die zugrundeliegenden Mechanismen. Forschung ist die ständige Entdeckung neuer und die Modifizierung und Präzisierung bekannter Heuristiken. Die Beschreibung physiologischer Mechanismen durch Heuristiken kann ein unterschiedliches Auflösungs-niveau haben. Es reicht von der Ebene einzelner Neurone oder Neuronenverbände bis zum Zusammenspiel von Hirnregionen. Welches Auflösungs-niveau gewählt wird, hängt von der Zielsetzung der Forschung ab, welches erreicht wird, von ihren Methoden. Die Anwendung von Heuristiken in Computerprogrammen (Computer Vision) erlaubt es heute, „sehende Maschinen“ zu konstruieren, die erstaunlich leistungsfähig sind. Ein eindrucksvolles Beispiel sind autonome Automobile, auf die später noch ausführlicher eingegangen wird.

Phänomene

Phänomene sind die subjektive Seite der Wahrnehmung. Sie sind das Resultat der Prozesse. Sie werden auch als *Perzepte* bezeichnet. Sie sind Bestandteile des Bewusstseins. Beobachtet werden kann das „Was“ (das

Phänomen), nicht das „Wie“ (der Prozess). Auf diese Weise entsteht der Eindruck einer Einheitlichkeit, einer Wirklichkeitstreue und einer Unmittelbarkeit des Erlebens. Die im Gehirn generierte Wirklichkeit erscheint als Außenwelt. Wir bemerken nicht, dass die visuelle Welt in unserem Kopf entsteht.

Die Alltagstheorie der Wahrnehmung ist ein „naiver Realismus“. Dieser geht davon aus, dass die phänomenale und die physikalische Beschreibung der Welt übereinstimmen, dass die visuelle Welt lediglich ein Abbild der Außenwelt ist. Die Leistung des Gehirns, die Innenwelt als Außenwelt erscheinen zu lassen, ist im Alltag höchst zweckmäßig. Sie vermittelt den Eindruck eines unmittelbaren Zugangs zur Außenwelt. Eine bestimmte Klasse von Wahrnehmungserlebnissen ist in der Lage, auch auf der phänomenalen Ebene eine Irritation zu erzeugen und das Alltagsmodell der Wahrnehmung infrage zu stellen. Diese Gruppe von Erlebnissen wird oft und etwas unglücklich als „Täuschungen“ bezeichnet – treffender ist der Ausdruck „visuelle Illusionen“. Die physikalische und die phänomenale Beschreibung weichen hier so deutlich voneinander ab, dass der Widerspruch unübersehbar ist. Ihre Aufdeckung führt in der Regel nicht dazu, das Alltagsmodell der Wahrnehmung zu verwerfen. Vielmehr werden die Abweichungen als zufällige Fehler des Systems interpretiert. In Wirklichkeit handelt es sich nicht um ein Malheur, sondern um einen Einblick in den produktiven Charakter der Wahrnehmung, bei dem wichtige neuronale Mechanismen und die verborgenen Konstruktionsregeln der visuellen Welt sichtbar werden.

Eine besondere Rolle in der Geschichte der Wahrnehmungsforschung spielt die Beschreibung der funktionalen Beziehung zwischen Reizen und Phänomenen. Die *psychophysische Perspektive* bezeichnet das Verhältnis von Reizstärke zu Empfindungsstärke. Welche Reizstärke ist notwendig, um eine Empfindung auszulösen? Wie groß muss der Unterschied zwischen zwei Reizen sein, damit ein Unterschied in der Stärke der

Empfindung entsteht? Dabei verschiebt sich das Interesse aufgrund methodischer Entwicklungen in den Neurowissenschaften zunehmend von einer *äußeren Psychophysik*, bei der das Verhältnis der Reizeigenschaften zum Perzept im Mittelpunkt steht, zu einer *inneren Psychophysik*, die das Perzept in ein Verhältnis zur Erregungsverteilung im Gehirn setzt.

Prinzipien

Prinzipien beschreiben die visuelle Informationsverarbeitung auf einer sehr abstrakten Ebene. Sie sind unabhängig von spezifischen Reizmerkmalen und gelten zum Beispiel für die Form- wie für die Bewegungswahrnehmung. Sie erklären das Zusammenspiel der Prozesse, indem sie Begründungen für ihre Effizienz liefern. Es sind theoretische Begriffe, die neuronalen Mechanismen einen Sinn geben. Bei Prozessen ist es plausibel zu fragen: Wann und wo finden sie statt? Bei Prinzipien ist eine solche „Verortung“ und „Verzeitlichung“ nicht möglich. Das System selbst bedarf ihrer nicht. Prinzipien sind erläuternde Konstrukte der Forscherin oder des Forschers, Hilfen, die das Verständnis erleichtern. Sie beantworten die Frage nach dem „Wozu“ oder dem „Warum“. Eine Liste dieser Prinzipien ist weder vollständig noch trennscharf. Zwischen den Prinzipien gibt es mehr oder weniger große Überlappungen und ihre Auswahl hängt von der Zielsetzung der Beschreibung ab.

Modelle

Modelle verdichten unser Wissen über die visuelle Informationsverarbeitung. Sie geben uns eine Vorstellung von einem an sich unanschaulichen Prozess und bringen die naturwissenschaftlichen Beobachtungen in einen deduktiven Zusammenhang. Da über die Introspektion kein Zugang zum Sehprozess gelingt, ist eine Beschreibung in Form von Modellen notwendig, um eine Vorstellung von seinem Ablauf zu gewinnen. In der aktuellen Wahrnehmungsforschung finden sich vor

allem physiologische und computationale Modelle. Letztere sind Modelle, die den Vorgang des Sehens durch ein Computerprogramm simulieren. Nach der hier vertretenen Auffassung sind Modelle *Theorien über das Sehen*. Daher wird nicht zwischen Theorien und Modellen unterschieden. Wie auch die Beschreibung der Prozesse durch Heuristiken, kann die Beschreibung durch Modelle oder Theorien ein unterschiedliches Auflösungs-niveau haben. Die Spanne reicht von einem Modell, das die Vorgänge auf der Ebene einzelner Neurone beschreibt, bis zu einem Modell, das den gesamten Prozess des Sehens umfasst. Unabhängig davon, auf welcher Ebene das Modell eine Beschreibung vornimmt, weist es gemäß der allgemeinen Modelltheorie drei Merkmale auf: Abbildung, Verkürzung und Pragmatismus (vgl. Stachowiak, 1973). Demnach ist das Modell immer ein Abbild von etwas, eine Repräsentation natürlicher oder künstlicher Originale. Es erfasst nicht alle Attribute des Originals, sondern nur diejenigen, die dem Konstrukteur oder der Konstrukteurin relevant erscheinen. Schließlich ist das Modell dem Original nicht von sich aus zugeordnet. Es ist eine Interpretation und es bedarf der Interpretation.

Zwischen den fünf diskutierten Beschreibungsebenen gelten folgende Zusammenhänge: Ausgangspunkt ist die *Stimulation*. Sie wird gebildet durch den Teil des Umgebungslichtes, der die Retina erreicht und für den das Auge empfindlich ist. Ihre Beschreibung erfolgt in Termini der Physik, erweitert um eine ökologische Optik, die den Bezug zum Sinnesorgan herstellt. Die *Phänomene* sind die subjektive Seite der Wahrnehmung. Sie zeigen die Leistungen des visuellen Systems. Diese sind erklärungsbedürftig. Hierzu untersucht die Wahrnehmungsforschung die *Prozesse* der visuellen Informationsverarbeitung. Diese naturwissenschaftlichen Beobachtungen und Experimente bilden das Fundament. Das Wissen über die neuronalen Mechanismen kann in Form von *Heuristiken* zusammengefasst werden. Sie beschreiben, wie die Leistungen des visuellen Systems zustande kommen. Die *Prinzipien* sind übergeordnete theoretische Begriffe, die der Forscher einführt, um zu

erklären, warum einzelne Prozesse oder Mechanismen sinnvoll sind, bzw. welche Funktion sie bei der Analyse der Stimulation haben. Die Prinzipien müssen mit dem Wissen über die neuronalen Mechanismen konsistent sein. *Modelle* schließlich fassen unser Wissen über das Sehen in anschaulicher Weise zusammen. Sie sind selbst Theorien über das Sehen. Die Geschichte der Wahrnehmungsforschung lässt sich als eine Geschichte solcher Theorien beschreiben. Die Beschreibung auf den einzelnen Ebenen verfolgt unterschiedliche Ziele.

Da derzeit kein anerkanntes universelles Modell existiert, ist es notwendig, die eigene Position zumindest in Grundzügen zu skizzieren und einige Überlegungen, die bisher zwischen den Zeilen anklangen, zu explizieren. Angesichts der Schwierigkeit des Problems ist es nicht überraschend, dass wichtige Fragen auch nach mehr als einhundertfünfzig Jahren intensiver naturwissenschaftlicher Wahrnehmungsforschung kontrovers diskutiert werden: Wie entsteht aus der zweidimensionalen Abbildung auf der Retina eine dreidimensionale Welt? Warum wird die Welt als einheitlich erlebt, wenn doch viele ihrer Merkmale in getrennten Schritten verarbeitet werden? Warum erscheinen die Veränderungen, und damit die Zeit, als kontinuierlich, wenn dem visuellen System nur diskrete Momentaufnahmen zur Verfügung stehen? Wie kommt es, dass Objekte wiedererkannt werden, die aus einer bestimmten Perspektive noch nie gesehen wurden? Wie funktioniert das Zusammenspiel zwischen den verschiedenen Sinnesmodalitäten?

Je grundlegender die Fragen, umso kontroverser die Positionen. Eine Rekonstruktion der Argumente deckt zwei unterschiedliche Auffassungen über die visuelle Welt auf. Die erste lautet: Die visuelle Welt ist das Resultat der *Analyse einer Stimulation*. Hier wird betont, dass in der Stimulation bereits alle Merkmale enthalten sein müssen, die später auch die bewusste Wahrnehmung bestimmen. Die zweite Position lautet: Die visuelle Welt ist das Ergebnis einer *Konstruktion auf der Basis neuronaler Prozesse im Gehirn*. Demnach ist die Stimulation zwar notwendig, aber sie bildet

lediglich den Ausgangspunkt. Mit Hilfe des Gehirns wird die visuelle Welt nach eigenen Gesetzmäßigkeiten geschaffen.

Die vom Autor vertretene Auffassung lautet: Sehen ist Analyse plus Konstruktion. Der suggerierte Gegensatz ist ein künstlicher, anstelle eines „oder“ ist ein „und“ wahrscheinlicher: *Sehen ist die Konstruktion eines visuellen Modells auf der Basis einer Analyse permanenter Stimulation*. Die gleichberechtigte Verwendung der beiden Begriffe Analyse und Konstruktion macht deutlich, dass die Wahrnehmung, im Gegensatz zur Vorstellung, immer durch die Stimulation gesteuert wird. Durch Objekt-, Eigen- und Augenbewegungen verändert sich diese ständig. Zu Recht kann von einer *dynamischen Stimulation* des Sinnesorgans als Grundlage der weiteren Informationsverarbeitung gesprochen werden. Eine wesentliche Funktion der Wahrnehmung besteht darin, aus dieser Stimulation eine stabile Welt zu erzeugen. Dies ist die Aufgabe des Gehirns. Sie lässt sich nicht dadurch lösen, dass Kopien der Außenwelt gespeichert werden. Vielmehr müssen Informationen selektiert, ergänzt und bewertet werden. Hierbei spielen Heuristiken eine zentrale Rolle. Wegen der produktiven Aspekte wird dieser Prozess als Konstruktionsprozess bezeichnet. Eine angemessene Vorstellung dieses Prozesses ist, dass Wahrnehmende ein *visuelles Modell der Welt* erstellen, so wie Forscher*innen ein *Modell der visuellen Wahrnehmung* erstellen.

2. Aufgaben

Heute verbinden wir das Sehen oft mit Tätigkeiten wie dem Autofahren, dem Fernsehen oder dem Lesen. Evolutionsgeschichtlich sind diese Situationen eher unbedeutend. Wichtig für die phylogenetische Entwicklung des Systems waren zunächst Situationen, die der Orientierung, der Nahrungsaufnahme, der täglichen Arbeit sowie dem menschlichen Zusammenleben dienten. In diese funktionalen Zusammenhänge ist das Sehen untrennbar eingebunden. Deshalb weist es

enge Beziehungen zur Körpermotorik sowie den übrigen Sinnessystemen auf. Seine vielfältigen Aufgaben könnte es ohne diese Kooperationen nicht erfüllen. Sie lassen sich in sechs Bereiche gliedern: Lokomotion und Navigation, Identifikation von Objekten und Personen, Vorbereitung und Kontrolle von Handlungen, Interpretation von Situationen, Gewinnung von Erkenntnissen, ästhetischer Genuss und ästhetische Erfahrung. Zwischen diesen Bereichen gibt es vielfältige Überlappungen.

(a) Lokomotion und Navigation

Lokomotion und Navigation sind eng verbunden und stehen für die Bestimmung der eigenen Position im Raum sowie der Steuerung und Kontrolle der Fortbewegung. Es sind komplexe Anforderungen, für die wir Routinen entwickelt haben, sodass sie in den meisten Fällen keine Aufmerksamkeitszuwendung mehr benötigen. Besonders an der Lokomotion wird deutlich, dass das visuelle System es nicht nur mit räumlichen Koordinaten, sondern auch mit zeitlichen Abfolgen zu tun hat. Deshalb ergibt es Sinn, im Zusammenhang mit dem Sehen von *visuellen Ereignissen* als einer grundlegenden Untersuchungseinheit zu sprechen. Der Begriff des Ereignisses betont die Strukturiertheit eines Vorgangs in der Zeit und seine Nicht-Umkehrbarkeit. Wie die anderen Bereiche auch zerfallen Lokomotion und Navigation in sehr unterschiedliche Aufgaben: Bestimmung von Entfernungen, Identifikation von Oberflächen und ihren Neigungen, Einschätzung von Beleuchtungsverhältnissen – um nur einige Beispiele zu nennen.

(b) Identifikation von Objekten und Personen

Während uns die Aufgaben des Sehvorgangs in Zusammenhang mit der eigenen Fortbewegung in der Regel nicht bewusst werden, steht die Identifikation von Objekten und Personen oft im Vordergrund. Vermutlich hängt dies damit zusammen, dass willentliche Prozesse hier

eine größere Rolle spielen und dass wir im Rahmen zwischenmenschlicher Kommunikation oft nach Bezeichnungen für Objekte oder Personen suchen. Bei einer Reflexion über das Sehen findet sich daher eine Konzentration auf den Bereich der Objektidentifikation. Diese Schwerpunksetzung wurde von der Wahrnehmungsforschung lange Zeit geteilt. So gibt es deutlich mehr Forschung zu den einzelnen Phasen der Objekterkennung als etwa zur Navigation, zur Handlungsvorbereitung oder zu den Erkenntnisfunktionen des Sehens. Heute wissen wir, dass auch die Objektidentifikation überwiegend vorbewusst abläuft. Beispiele für Aufgaben in diesem Bereich sind etwa die Abgrenzung von Figur und Grund, die Bestimmung von Kanten- und Formen sowie der Vergleich verschiedener Ansichten.

(c) Vorbereitung und Kontrolle von Handlungen

Die Vorbereitung und die Kontrolle von Handlungen setzen bereits die Bestimmung der eigenen Position sowie eine korrekte Objektidentifikation und -lokalisierung voraus. Die vorbereitenden Aufgaben des visuellen Systems sind im Vergleich zu den Identifikationsaufgaben noch komplexer und anspruchsvoller. Kommen eigene Handlungen ins Spiel, so ist eine Eindeutigkeit der Interpretation, wie auch eine präzise zeitliche Koordination gefordert. Wichtige Aufgaben sind hier zum Beispiel: Bestimmung der Lage des Objektes im Raum relativ zur eigenen Position, Interpretation der Oberflächen hinsichtlich ihrer Krümmung und Festigkeit sowie Kontrolle von Greifbewegungen.

(d) Interpretation von Situationen

Im Zusammenhang mit der Entwicklung „sehender Maschinen“ ist deutlich geworden, dass eine Objektidentifikation alleine nicht ausreicht, um sich in unserer Welt zurechtzufinden. Vielmehr benötigen wir ein Situationsverständnis. In der Wahrnehmungsforschung sind deshalb

„visuelle Szenen“ ein wichtiges Thema. Sie bestehen aus mehreren Objekten sowie einer Unterscheidung zwischen Vorder- und Hintergrund. Eine Szene stellt einen Kontext her und die Beobachterin oder der Beobachter kann entscheiden, ob das Objekt in die Szene „passt“ oder nicht. Außerdem enthalten Szenen Informationen über die Relationen zwischen den Objekten, etwa ihre Größe, ihre Lage oder auch ihr semantisches Verhältnis. Zwischenmenschliche Kommunikation setzt ein gemeinsames Situationsverständnis voraus.

(e) Erkenntnisse

Der Bereich der Erkenntnisse umfasst alle Aufgaben des Sehens, die in Zusammenhang mit der Neugier des Menschen und seinen explorativen Tätigkeiten stehen. Er verweist darauf, dass das visuelle System nicht nur in die Zusammenhänge des Alltags eingebunden ist, sondern dass ihm eine wichtige Funktion beim Erwerb neuen Wissens zukommt. Dabei steht die Aufgabe des Beobachtens im Vordergrund. Eine systematische und kontrollierte Beobachtung ist grundlegend für die Entwicklung und Prüfung wissenschaftlicher Hypothesen. Sie erlaubt es, unsere Annahmen über die Welt zu verifizieren und weiterzuentwickeln. Für die Wahrnehmungsforschung ergibt sich hieraus eine Duplizität: Der Gegenstand ihrer Forschung ist gleichzeitig eine ihrer Methoden.

(f) Ästhetischer Genuss und ästhetische Erfahrung

Schließlich soll ein Aufgabenbereich angesprochen werden, bei dem das *Betrachten* im Mittelpunkt steht. Beispiele hierfür sind die Naturbetrachtung oder die Kunstbetrachtung. Typischerweise handelt es sich um Sehvorgänge, die keinem speziellen Handlungs- oder Erkenntnisinteresse folgen, sondern bei denen der Wahrnehmungsakt selbst im Mittelpunkt steht. Hier kommt es auch nicht auf Schnelligkeit an, vielmehr ist die Intensität des Wahrnehmungserlebnisses entscheidend. Es

sind Situationen, die wir als Genuss empfinden und in denen wir deshalb eine Zeit lang verweilen möchten. Die längere zeitliche Erstreckung führt dazu, dass andere kognitive Prozesse – wie die Imagination oder die Erinnerung – in stärkerem Maße an der Wahrnehmung beteiligt sind.

Für dieses weite Spektrum an Aufgaben benötigen wir ein System, das Flexibilität mit Schnelligkeit und Zuverlässigkeit verbindet – keine geringen Anforderungen. In biologischen Systemen können Mehrfachanforderungen durch Aufgabenteilung und den Einsatz von Lernprozessen gelöst werden. Im visuellen System finden wir eine Reihe von Teilsystemen, die eine deutliche Spezialisierung aufweisen. Hinzu kommt die Möglichkeit, neu eintreffende Informationen mit bereits bestehenden Gedächtnisinhalten zu vergleichen und auf Abweichungen zu prüfen.

3. Prinzipien

Die nachfolgenden Überlegungen beginnen mit der Frage nach dem „Wozu“, denn es ist leichter, das „Wie“ zu verstehen, wenn der Zweck bekannt ist. Beschrieben werden zunächst die Prinzipien, die den Sehprozess steuern. Das hat den Vorteil, dass bei der Analyse der einzelnen Prozesse eine Begrifflichkeit zur Verfügung steht, die es erlaubt, Zusammenhänge aufzuzeigen und theoretische Strukturierungen vorzunehmen.

Das Problem besteht darin, dass mehrere – zum Teil gegensätzliche – Anforderungen gleichzeitig zu erfüllen sind: Die phänomenale Welt muss die Komplexität der physikalischen Welt wiedergeben, in kürzester Zeit müssen Entscheidungen getroffen und Handlungen ausgeführt werden. Während die Stimulation sich permanent ändert, muss die phänomenale Welt stabil sein. Schließlich gilt es sicherzustellen, dass trotz

unterschiedlicher Erfahrung einheitliche Interpretationen einer Situation möglich sind.

Wie bereits angedeutet, handelt es sich bei den Prinzipien um zusammenfassende Beschreibungen auf einer hohen Abstraktionsstufe. Ihre Auswahl erfolgt primär unter dem Gesichtspunkt, das Verständnis für den Sehvorgang als einen einheitlichen Prozess zu fördern und eine Vorstellung von der komplexen Aufgabenstellung zu geben, vor der das visuelle System steht. Da die zugehörigen neuronalen Prozesse ihre Grundlagen in den anatomischen und physiologischen Gegebenheiten des Auges oder des Gehirns haben, der Mensch aber gleichzeitig immer auch als Organismus ein Bestandteil dieser Welt ist, lassen sich die Prinzipien noch einmal gruppieren. Vor allem die ersten beiden hängen eng mit dem Aufbau des Auges zusammen, die Prinzipien drei bis zehn reflektieren stärker die Arbeitsweise des Gehirns und elf bis vierzehn sind gleichzeitig Beschreibungen einer Existenz in Raum und Zeit. Das letzte Prinzip schließlich nimmt insofern einen Sonderstatus ein, als dass es versucht, die qualitativen Veränderungen, die sich im Laufe des Informationsverarbeitungsprozesses ergeben, zu charakterisieren. Es geht der Frage nach: Wie kann aus der Analyse einzelner Merkmale eine bedeutungsvolle Ganzheit entstehen?

(1) Retinotopie

Um uns in einer dreidimensionalen Welt zurechtzufinden, müssen die ursprünglichen Größen- und Entfernungsverhältnisse zwar nicht mit ihren absoluten, aber mit ihren relativen Werten über alle Stufen der visuellen Informationsverarbeitung hinweg erhalten bleiben. Für unsere Fortbewegung wie für unsere Handlungen benötigen wir eine Vorstellung der metrischen Verhältnisse der Außenwelt. Die Basis hierfür bildet eine räumliche Kartierung. Sie wird *retinotope Organisation* genannt und sichert die feste Zuordnung zwischen der Position eines Reizes im

Gesichtsfeld und dem Ort seiner neuronalen Verarbeitung. Die Bausteine hierfür bilden die *rezeptiven Felder* der Neurone (Bereich des Gesichtsfeldes, der bei einem Neuron eine Antwort hervorruft). Diese lassen sich ziemlich genau markieren. Zwar werden die rezeptiven Felder der Neurone, beginnend mit der Rezeptorebene in der Retina bis hin zu den visuellen Arealen des Cortex, zunehmend größer, aber die Neurone mit benachbarten oder überlappenden rezeptiven Feldern projizieren von Stufe zu Stufe jeweils wieder zu benachbarten Neuronen. Durch den Erhalt der Nachbarschaft wird eine Konstanz der räumlichen Anordnung erreicht.

(2) Fovealisierung

Für unser Überleben war es wichtig, einen Fernsinn zu entwickeln, der sowohl einen Überblick garantiert, als auch die Möglichkeit von Detailanalysen bietet. Dies wird erreicht durch die Kombination von zwei komplementären Vorgängen, dem *peripheren* und dem *fovealen Sehen*. Objekte treten in der Regel zunächst über die Peripherie in das Gesichtsfeld (horizontal 180°). Sie können hier zwar registriert, aber nicht analysiert werden. Hierfür ist eine Abbildung im Zentrum der Retina, der Fovea centralis, notwendig. Dies ist der Ort des schärfsten Sehens. Hier ist die Dichte der Photorezeptoren mit Abstand am größten und im Unterschied zu den anderen Bereichen der Netzhaut ist hier eine Eins-zu-eins-Verschaltung zu den nachfolgenden Ganglienzellen realisiert. Das bedeutet, dass die ursprüngliche Informationsdichte erhalten bleibt und keine Details verloren gehen. Von Beginn an gibt es eine Überrepräsentation der Fovea, eine *Fovealisierung* des Sehens, die während der weiteren Informationsverarbeitung eher noch verstärkt wird, denn die Fovea deckt nur etwa 2 % des Gesichtsfeldes ab und nimmt nur 0,01 % der gesamten Retina ein. Ihre Ganglienzellen beanspruchen aber ein Vielfaches an Cortexgewebe im Verhältnis zu einer Ganglienzelle der Netzhautperipherie. Die hohe Dichte der Fotorezeptoren und

Ganglienzellen in Verbindung mit den Charakteristika ihrer neuralen Verschaltung führt zu dieser Überrepräsentation. Die Konsequenz ist eine Fokussierung des Sehens auf einen kleinen Teil des Gesichtsfeldes, dessen Auswahl sich permanent ändert. Das Zentrum der Analyse wird neu bestimmt. Dieser ständige *Wechsel der Zentrierung* erlaubt bei begrenzten Ressourcen eine hochauflösende Analyse in einem weiten Gesichtsfeld.

(3) Kontrastverstärkung

Die proximalen Reize, die sich als Projektionen der distalen Reize auf den Retinae der beiden Augen ergeben, weisen viele kontinuierliche und nur wenige abrupte Übergänge auf. Das visuelle System verstärkt und akzentuiert die bestehenden Unterschiede. In der weiteren Reizverarbeitung werden die selbst erzeugten Kontraste zu Grundbausteinen für den Aufbau der visuellen Welt. Die Kontrastverstärkung wiederholt sich auf verschiedenen Stufen der Reizverarbeitung. Beginnend in der Retina, schützt sie das Sehen vor zu starken lokalen und zufälligen Einflüssen. Sie stellt sicher, dass sich Ereignisse nicht permanent ähneln und reduziert die Gefahr von Verwechslungen und Fehlinterpretationen. Auf der einen Seite führt die Kontrastbildung zu einem deutlichen Verlust an Detailinformation, auf der anderen Seite ist sie die Voraussetzung dafür, dass Oberflächen sicher und schnell voneinander abgegrenzt werden können: Ordnung und Struktur nehmen zu. Diese konstruktiven Prozesse beginnen bereits auf retinaler Ebene.

(4) Gruppierung

Das Perzept ist ein einheitliches, dreidimensionales Bild mit klaren Abgrenzungen zwischen Figur und Grund sowie deutlichen Entfernungsverhältnissen. Diese Ordnung ist so in der Stimulation nicht enthalten. Das visuelle System erzeugt sie erst durch Gliederungs- und Gruppierungsprozesse, bei denen die Inhomogenitäten der Stimulation

nach festen Regeln (z. B. Nähe und Gleichartigkeit) zusammengefasst werden. Diese Reorganisation wiederholt sich mit wechselnden Kriterien auf verschiedenen Stufen der Informationsverarbeitung. Ein strukturbildender Vorgang, der die Voraussetzungen schafft für eine Bewegung im Raum, das Erkennen von Objekten und die Interpretation von Situationen. In der Regel resultiert trotz der großen Zahl möglicher Gruppierungen für eine Stimulation nur eine Ordnung. Es entsteht ein stabiles Perzept mit eindeutiger Interpretation. Kommt es zu einer Konkurrenz von gleichwertigen Ordnungen, dann entsteht eine vorübergehende Mehrdeutigkeit (vgl. ausführlich Kapitel IV).

(5) Ambiguitätsreduktion

Die Mehrdeutigkeit der Stimulation wird uns nicht bewusst. Aber prinzipiell ist jeder proximale Reiz vieldeutig. Ein wesentlicher Aspekt dieser Mehrdeutigkeit ist geometrischer Natur. Es kann keine eineindeutige Abbildung eines dreidimensionalen Körpers auf einer zweidimensionalen Fläche geben. Die optische Projektion eines Gegenstandes auf die Retina ist zwar eindeutig, aber einem bestimmten Netzhautbild entsprechen theoretisch unendlich viele Objekte (Problem der inversen Projektion). Daher besteht eine permanente Aufgabe der Wahrnehmung in der Ambiguitätsreduktion. Es handelt sich um ein unlösbares Problem, für das es nur Annäherungen geben kann. Sowohl die eigene Fortbewegung als auch die Planung und Ausführung von Handlungen setzen Eindeutigkeit voraus. Um diese zu erreichen, wird notfalls die gesamte Stimulation neu interpretiert. Umgekehrt werden Analyseprozesse abgebrochen, wenn Eindeutigkeit erreicht ist.

(6) Ergänzung

Das retinale Bild ist aus verschiedenen Gründen unvollständig und diskret (Verteilung der Rezeptoren, Augenbewegungen, Lidschlussreflex). Das Perzept hingegen ist vollständig und kontinuierlich. Die visuelle Welt

erscheint der Beobachterin oder dem Beobachter wie ein farbiger Film. Die daraus resultierende Frage lautet: Welche Gesetze bestimmen diese räumlichen und zeitlichen Interpolationen? Es ist erstaunlich, auf welcher geringen Datengrundlage die visuelle Wahrnehmung noch ein vollständiges Perzept erstellt. Ergänzungen sind nicht die Ausnahme, sondern die Regel. Sie sind nicht willkürlich, vielmehr ein Spiegelbild der Verarbeitungsprozesse. Das visuelle System geht permanent über die bestehenden Informationen hinaus. Dabei nutzt es die Wahrscheinlichkeiten für das Auftreten visueller Ereignisse in der Außenwelt als Vorhersage. Ohne diese Ergänzungen wäre keine erfolgreiche Ambiguitätsreduktion möglich.

(7) Modularisierung

Vor allem in den frühen Phasen der visuellen Informationsverarbeitung werden die einzelnen Merkmale eines Reizes wie Form, Farbe, Orientierung und Bewegung getrennt verarbeitet. Deshalb lassen sie sich auch unterschiedlichen neuronalen Strukturen zuordnen. Die Modularisierung ist die Voraussetzung für eine schnelle, datengesteuerte und automatisierte Stimulusanalyse. Sie wird zunehmend ergänzt durch eine multidimensionale Vorgehensweise, welche die verschiedenen Merkmale eines distalen Reizes gleichzeitig berücksichtigt. Immer größere Teile des Gesichtsfeldes werden in die Analyse einbezogen und es verändert sich die Arbeitsweise der zugehörigen Neurone: Komplexere Merkmale können durch spezifischere Neurone untersucht werden. Da modulare und mehrdimensionale Prozesse sich überlappen, müssen die einzelnen Merkmale eines Objektes oder einer visuellen Szene später nicht neu zusammengesetzt werden.

(8) Bidirektionalität

Heute wissen wir, dass die Datenanalyse nicht abgeschlossen sein muss, bevor eine erste Interpretation einer komplexen Stimulation zur Verfügung steht. Parallel zu der von der Retina über das Geniculatum und den primären visuellen Cortex erfolgenden *bottom-up* Analyse finden erfahrungs- und hypothesengesteuerte Prozesse statt, die von kortikalen Zentren ausgehen. Je weiter die Informationsverarbeitung fortschreitet, umso größer wird der Einfluss dieser *top-down* Verbindungen. Schließlich stammt der überwiegende Teil der neuronalen Stimulation nicht mehr von den proximalen Reizen, sondern aus anderen Regionen des Gehirns. Daten- und hypothesengesteuerte Verarbeitung ergänzen sich in idealer Weise, um die Mehrdeutigkeiten des retinalen Bildes zu reduzieren. Die zunehmende Spezialisierung der beteiligten Neurone wird ergänzt durch eine zunehmende Konvergenz ihrer Verschaltung. Auf den höchsten Stufen schließlich verfügen wir über stark spezialisierte Einheiten, die mithilfe umfangreicher lateraler Verschaltungen effizient miteinander kommunizieren. Insgesamt steigt das Abstraktionsniveau: Unähnliches wird zunehmend vergessen.

(9) Perspektivität

Die visuelle Welt ist auf einen Standort bezogen, dieser bildet das Zentrum. Bewegt sich die Beobachterin oder der Beobachter, dann verlagert sich auch der Mittelpunkt der visuellen Welt. Die Perspektive definiert sowohl das Verhältnis der Beobachterin oder des Beobachters zu den Gegenständen wie die Relationen der Gegenstände zueinander. Perspektive bedeutet die Anpassung der Außenwelt an das Auge der Beobachterin oder des Beobachters. Abhängig vom Standort und den Kopf- und Augenbewegungen sehen wir die Welt immer aus einer bestimmten Perspektive. Diese *Perspektivenabhängigkeit* des Sehvorgangs ist uns während des Sehens nicht bewusst. Erst in der Reflexion über das Sehen werden mögliche Differenzen zwischen verschiedenen Ansichten

deutlich. Die Leistung des visuellen Systems besteht darin, Ansichten zu erstellen, die auch bei einer Veränderung des Standortes oder einer Bewegung des Objektes wiedererkannt werden können. Ein Objekt, das wir gewöhnlich von vorne sehen, sollten wir nicht deshalb für ein anderes Objekt halten, weil es um einige Grad gedreht wurde oder weil wir es aus einem anderen Blickwinkel sehen. Letztlich benötigen wir eine Repräsentation, die mehrere Ansichten verbindet, um uns in einer dreidimensionalen Welt zurechtzufinden. Gelingt die Identifikation trotz einer veränderten Perspektive, dann muss es verborgene Informationen geben, die das visuelle System nutzt. Denn das Netzhautbild, das den Ausgangspunkt bildet, ist immer eine zweidimensionale und perspektivenabhängige Abbildung. Wie erreicht das visuelle System Multiperspektivität? Nicht alle Perspektiven müssen gleichberechtigt sein. Vieles weist darauf hin, dass es eine Dominanz ausgewählter Ansichten gibt. Hinzu kommt, dass wir uns nicht mehrere Ansichten eines Objektes gleichzeitig vorstellen können. Streng genommen handelt es um eine Abfolge einzelner Ansichten, eine Perspektivensukzession bzw. einen sehr schnellen Perspektivenwechsel (zur Phänomenologie und Psychologie der Perspektivität vgl. Graumann, 1960; Kubovy, 1986).

(10) Antizipation

Was wir sehen, ist mehr als das, was uns die Projektion des Gegenstandes auf die Netzhaut zeigt. Phänomenal gegeben ist nicht nur die Ansicht eines Körpers, sondern der Körper selbst. Auch seine verdeckten (amodalen) Seiten sind präsent und wir haben eine Vorstellung von den Entfernungen zwischen den Gegenständen und vom Raum um uns herum. Der „Ausschnitt“ enthält immer auch den Verweis auf die „Ganzheit“. Dies führt dazu, dass der einzelne Sehvorgang bereits einen Bezug zu seinem Nachfolger enthält. In gewisser Weise bedingt der vorherige Blick bereits den nächsten. Dieser *Verlaufscharakter* wird besonders deutlich, wenn das Sehen in Handlungsabläufe eingebunden ist. Ausgehend von der ersten

Fixation lassen sich Übergangswahrscheinlichkeiten zu den nachfolgenden Fixationen berechnen. Die Gerichtetheit des Verlaufs ist Ausdruck einer Antizipation. Das visuelle System analysiert nicht nur die aktuelle Stimulation, sondern erstellt Vorhersagen: „Was könnte sein?“ und „Was ist am wahrscheinlichsten?“.

(11) Probabilität

Die Außenwelt weist einen hohen Grad an Regelmäßigkeiten auf. Die für die visuelle Wahrnehmung wichtigen Oberflächen zeigen charakteristische Texturen, die Helligkeitsveränderungen im Laufe des Tages erfolgen meist graduell und Objektgröße und Entfernung stehen in einem direkten Verhältnis zueinander. Das visuelle System geht bei seiner Analyse der Stimulation von diesen Regelmäßigkeiten aus. Es berücksichtigt Wahrscheinlichkeiten, mit denen Ereignisse in der Außenwelt auftreten. Es erstellt Vorhersagen, was zu sehen sein wird. Dabei geht es von den *erfahrungsgebundenen Wahrscheinlichkeiten* aus, die das Sinnessystem im Laufe seiner Phylogenese und Ontogenese als erfolgreich internalisiert hat. Diese Wahrscheinlichkeiten wiederum korrespondieren mit den Regelmäßigkeiten der Außenwelt. Der Begriff des *Ereignisses* wird hier nicht in seiner umgangssprachlichen Bedeutung verwendet. Ein Wahrnehmungsereignis kann so unterschiedliche Sachverhalte wie die Veränderung einer Beleuchtungssituation, die Bewegung eines Objektes oder die zunehmende Verdeckung eines Hintergrundes umfassen. Eine probabilistische Auffassung der Wahrnehmung berücksichtigt, dass es eine Entsprechung geben muss zwischen physikalischer und phänomenaler Welt. Sie lässt jedoch genügend Spielraum für die konstruktiven Aspekte der Wahrnehmung, für artspezifische und individuelle Lernprozesse sowie für Zufälle in der Stimulation und bei der Informationsverarbeitung.

(12) Stabilität

Durch permanente Eigen- und Objektbewegungen gibt es keine stabile Stimulation. Die aus didaktischen Gründen oft gewählte Situation der Einzelabbildung eines Objektes bei einem festen Augenpunkt ist eine Fiktion. Für diese Situation wurde das visuelle System nicht geschaffen. Hier arbeitet es ineffizient und fehlerhaft. Die Leistung des visuellen Systems besteht gerade darin, aus den Veränderungen der Stimulation die *Invarianten* zu ermitteln und auf der Basis dieser Invarianten eine stabile visuelle Welt zu erzeugen. So können wir zum Beispiel feststellen, ob sich ein Objekt im Gesichtsfeld bewegt oder ob wir uns selbst bewegen, ob sich lediglich die Beleuchtung eines Gegenstandes geändert hat oder ob der Gegenstand eine andere Farbe hat. Auf diese anspruchsvollen Aufgaben hin hat sich das Sehsystem im Laufe der Evolution entwickelt. Stabilität aber kann in einem nichtlinearen dynamischen System immer nur temporär sein. Der phänomenalen Beständigkeit der Welt entspricht auf der Ebene der neuronalen Prozesse lediglich ein kurzer stabiler Zustand zwischen periodischen Wechseln.

(13) Zirkularität

Typischerweise wird der Sehvorgang als ein einmaliger Vorgang mit einer festen Abfolge von Schritten beschrieben: von den ersten Stufen der Informationsverarbeitung in der Retina bis zu den kortikalen Zentren. Die Beziehung des Sehens zur Motorik und zu anderen Sinnessystemen bleibt außer Acht. In Wirklichkeit ist das Sehen eine sich permanent wiederholende *sensomotorische Leistung*. Das Auge ist Rezeptor und Effektor zugleich. Mithilfe der Okulomotorik wird die Stimulation ständig erneuert und verändert. Die Bedeutung dieser Zirkularität wird oft unterschätzt. Stabilität und Multiperspektivität der visuellen Welt werden erst durch die Zirkularität des Wahrnehmungsvorgangs erreicht. Beim Sehen wird das Blickfeld ständig abgetastet. Neue Objekte werden von der Peripherie ins Blickzentrum gerückt, um scharf und detailreich abgebildet

zu werden (Fovealisierung). Schließlich werden die Annahmen über die Objekte verifiziert oder der Vorgang wird unter veränderten Bedingungen wiederholt (z. B. mit einer Kopfbewegung, die den Aufsichtswinkel variiert).

(14) Komplementarität

Wie bereits betont, muss das visuelle System unter sehr unterschiedlichen und schnell wechselnden Bedingungen ein zuverlässiges Perzept erstellen. Es erreicht dies durch eine Kombination von zunächst gegensätzlich erscheinenden Prozessen. Bei näherem Hinsehen handelt es sich um eine Aufgabenteilung mit dem Ziel einer komplementären Ergänzung. Die Komplementarität von Teilprozessen erhöht die *Flexibilität* des Gesamtsystems. Beispiele hierfür sind etwa der Wechsel vom peripheren zum fovealen Sehen oder das Zusammenspiel von skotopischem und photopischem Sehen, um in Abhängigkeit von der aktuellen Beleuchtung eine hohe Sensitivität mit einer hohen Auflösung zu verbinden. Schließlich können auch der Wechsel in der Arbeitsweise zwischen „frühen“ und „späten“ Phasen sowie die Bidirektionalität des Sehvorgangs („bottom-up“ versus „top-down“) und der ständige Wechsel zwischen Segregations- und Integrationsprozessen unter diesem Aspekt betrachtet werden.

(15) Übersummativität

Da der Ausgangspunkt der visuellen Informationsverarbeitung instabil, unvollständig und mehrdeutig ist, die Perzepte aber stabil, vollständig und eindeutig sind, stellt sich die Frage, über welche Zwischenschritte im Gehirn diese Qualitäten erreicht werden. Unter der Annahme, dass das visuelle System ein *repräsentationales System* ist (ein System, das bedeutungshaltige Darstellungen der Außenwelt erstellt), müssen sich im Laufe des Informationsverarbeitungsprozesses die Eigenschaften dieser Repräsentationen auf der Basis ihrer neuronalen Korrelate ändern.

Beschreiben lassen sich diese qualitativen Veränderungen mithilfe des Prinzips der „Übersummativität“, das durch Christian von Ehrenfels (1890) zur Kennzeichnung von „Gestalten“ in die Wahrnehmungsforschung eingeführt wurde: Das Ganze ist etwas anderes als die Summe seiner Teile. Erläutert hat von Ehrenfels die Gestaltqualitäten „Übersummativität“ und „Transponierbarkeit“ an einem akustischen Beispiel, dem Verhältnis zwischen Einzeltönen und Melodie. Während der einzelne Ton durch seine Tonhöhe definiert wird, bestimmt sich die Melodie über die Abfolge der verwendeten Intervalle sowie die Zeitdauer der einzelnen Töne und Pausen (Notenwerte). Außerdem kommen auf der Ebene der Melodie Wahrnehmungsqualitäten hinzu, die auf der Ebene der einzelnen Töne nicht vorhanden sind (z. B. Dissonanz). Wird eine Melodie in unterschiedlichen Tonarten gesungen oder gespielt, wird immer dieselbe Melodie erkannt, obwohl die einzelnen Komponenten (Tonhöhen) verschieden sind. Die physiologische Basis für eine Übersummativität des Sehvorgangs könnten die Interaktionen eines nichtlinearen Systems bilden.

In der Regel führt die Anwendung dieser Prinzipien zu einem erfolgreichen Ergebnis: ein eindeutiges und stabiles Perzept. Allerdings gibt es eine Reihe von Situationen, bei denen es zu überraschenden Lösungen kommt. Betrachtet man diese Situationen genauer, fällt ein gemeinsames Merkmal auf: Viele sind reduktionistisch. Der Beobachterin oder dem Beobachter steht entweder nicht das gesamte Spektrum der Erkundungsmöglichkeiten zur Verfügung oder die Stimulation ist in ihrer Komplexität stark eingeschränkt. Oft werden reduktionistische Situationen im Labor hergestellt, um einzelne Aspekte der Arbeitsweise des visuellen Systems zu analysieren (z. B. werden Reize monokular dargeboten oder die Stimulusexposition ist extrem kurzzeitig). Hier kommt es dann vermehrt zu systematischen Abweichungen zwischen den physikalischen und den phänomenalen Verhältnissen. Die Ursache hierfür lässt sich als Konflikt zwischen Heuristiken beschreiben. Für die

Forschung sind diese Situationen besonders interessant, weil die konstruktive Arbeitsweise des Systems deutlich wird.

3. Optimale Wahrnehmung

Welche Bedingungen für das Sehen sind optimal? Intuitiv könnte angenommen werden, dass es für das visuelle System leichter ist, die Verhältnisse der Außenwelt adäquat wiederzugeben, wenn die Stimulation möglichst einfach und stabil ist. Dies ist nicht der Fall. Das visuelle System ist bei komplexen Reizbedingungen weniger fehleranfällig. Es benötigt eine ständige Veränderung der Stimulation. Natürlich kann es bei zu schnellem Wechsel oder zu hoher Komplexität zu einer Überforderung kommen. Grundsätzlich aber gilt: Unsere Wahrnehmung ist an komplexe und sich verändernde Stimulationen adaptiert. Unter diesen Bedingungen arbeitet sie effizient. Doch müssen hierzu weitere Voraussetzungen erfüllt sein. Insbesondere muss die Beobachterin oder der Beobachter die Möglichkeit haben, dieses breite Informationsangebot auch zu nutzen. Abstand und Perspektive müssen variiert werden können. Kopf- und Augenbewegungen müssen möglich sein. Sind diese Bedingungen sowohl auf Seiten der Reize als auch auf Seiten der Beobachterin oder des Beobachters erfüllt, kann von einer *optimalen Wahrnehmung* gesprochen werden. Sie liegt vor, wenn es zu komplexen dynamischen Stimulationen kommt und die Beobachterin oder der Beobachter gleichzeitig die Möglichkeit hat, über die Zirkularität die im proximalen Reiz enthaltenen Ambiguitäten zu reduzieren. Eine Situation, die diese Bedingungen nicht erfüllt, wird in der Wahrnehmungsforschung als *verarmte Wahrnehmung* bezeichnet.

Fünf Grundsituationen des Sehens werden im Folgenden näher in den Blick genommen: das homogene Ganzfeld, die Umgebungswahrnehmung, die Bildwahrnehmung, das Trompe-l'œil (die Verwechslung von Bild und Wirklichkeit) und die virtuelle Realität. Das homogene Ganzfeld

markiert den Endpunkt der verarmten Wahrnehmung. Die Umgebungswahrnehmung stellt die optimale Wahrnehmung dar. Die Bildwahrnehmung nimmt eine mittlere Position ein. Während der Bildwahrnehmung und der Umgebungswahrnehmung eine große praktische Bedeutung zukommt, weil sie unser alltägliches Leben bestimmen, gewinnt die virtuelle Realität erst zunehmend an Bedeutung. Das homogene Ganzfeld und das Trompe-l'œil sind vor allem von theoretischem Interesse. Ein Vergleich der fünf Grundsituationen kann den Stellenwert einzelner Prinzipien und die Rolle der Heuristiken verdeutlichen.

Homogenes Ganzfeld

Das *homogene Ganzfeld* ist die „Nullsituation“ der Wahrnehmung: Was passiert, wenn die Prinzipien nicht zur Anwendung kommen können, weil eine adäquate Stimulation fehlt? Was sehen wir, wenn es im gesamten Gesichtsfeld keine Inhomogenitäten gibt? Gemeint ist damit eine Situation homogener Stimulation, nicht vollständiger Dunkelheit. Hier ist weder eine Kontrastbildung noch eine Gruppierung möglich. Die Regelmäßigkeiten der physikalischen Welt spielen keine Rolle und die Fovealisierung des Sehens läuft ins Leere. Eine Situation, wie sie unter natürlichen Beobachtungsbedingungen auf der Erde kaum vorkommt. Stellen Sie sich vor, dass eine homogene graue Fläche Ihr gesamtes Gesichtsfeld ausfüllt. Dies entspricht etwa der Situation einer Beobachterin oder eines Beobachters, der auf einer weiträumigen und ebenen Schneefläche steht und gleichzeitig von dichtem Nebel umgeben ist. Unter Laborbedingungen lässt sich eine akzeptable Annäherung an ein homogenes Ganzfeld erzeugen, indem eine Beobachterin oder ein Beobachter in eine gleichmäßig ausgeleuchtete Kugel schaut oder in einem speziell ausgekleideten und beleuchteten Raum mit vorgegebener Position sitzt.

Ein frühes Experiment zum homogenen Ganzfeld führte Wolfgang Metzger (1930) durch. Der Versuchsaufbau war wie folgt (Abbildung 2): Eine ebene weiße Wand wurde durch weiße Schirmabdeckungen an der Decke und am Boden so ergänzt, dass die entstehende Fläche das gesamte Gesichtsfeld einer Person, die im Abstand von 1,5 m zu der Fläche saß, ausfüllte. Mit einem Projektionsapparat wurde gleichmäßiges, in der Intensität regulierbares Licht erzeugt. Lediglich bei voller Lichtintensität waren die Übergänge zwischen den Schirmen und der Wand als Konturen erkennbar. Wurde die Beleuchtungsstärke reduziert, entstand der Eindruck einer homogenen weißen Fläche. Die Aufgabe der Proband*innen bestand darin, eine phänomentreue Beschreibung zu geben, d. h. ihr Wahrnehmungserlebnis möglichst detailliert und unverfälscht zu beschreiben. Eines der wichtigsten Versuchsergebnisse war, dass die gegenüberliegende Wand bereits nach kurzer Zeit ihren Flächencharakter verlor und eine räumliche Wahrnehmung zu dominieren begann. Eine Variation der Beleuchtungsstärke wurde dann nicht als Modifikation der Helligkeit erlebt, sondern als Veränderung des wahrgenommenen Raumes („Erweiterung“ oder „Verengung“). Für die Proband*innen war es unmöglich, eine Oberfläche zu identifizieren. Diese Diskrepanz zwischen den physikalischen und phänomenalen Verhältnissen führte Metzger (1930, S. 13) zu der Schlussfolgerung, „dass die Wahrnehmung einer Oberfläche, und zwar auch einer völlig homogenen Oberfläche, das Bestehen objektiver Inhomogenitäten irgendwelcher Art voraussetzt [...]. Danach wäre die Wahrnehmung einer völlig homogenen Fläche von der Größe des Gesichtsfeldes unmöglich.“ Die Untersuchungen zum homogenen Ganzfeld führen zu folgenden Schlussfolgerungen:

1. Die Wahrnehmung reagiert nicht auf den Reiz, sondern auf Reizunterschiede oder Reizänderungen im Gesichtsfeld.
2. Wenn möglich, versucht das visuelle System eine dreidimensionale Interpretation herzustellen.

3. Nur, wenn Inhomogenitäten vorliegen, kann ein stabiles Perzept erzeugt werden.

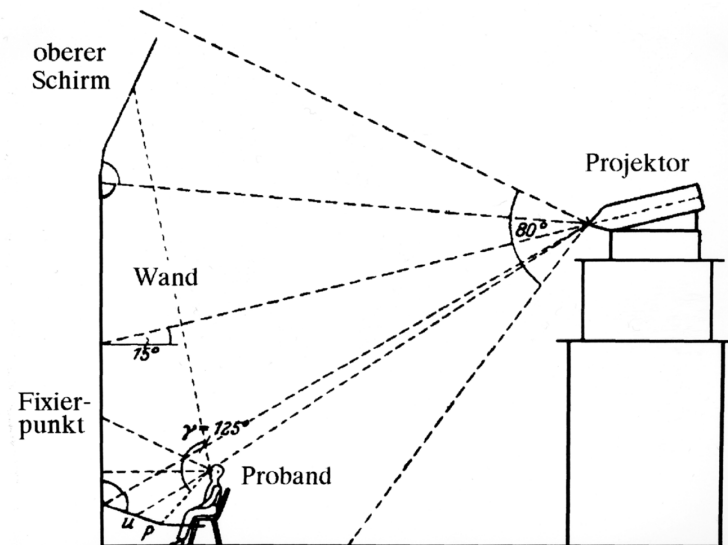


Abbildung 2: Versuchsanordnung zum homogenen Ganzfeld (Metzger, 1930).

Die Annahme, dass die Reizänderung und nicht der Reiz die entscheidende Bedingung für das Sehen bildet, lässt sich nur unter Laborbedingungen prüfen. Im Normalfall sichert das visuelle System mithilfe von Augenbewegungen, dass sich auch bei einem gleichbleibenden distalen Reiz der proximale Reiz permanent verändert. Die Reizänderung ist von so hoher Bedeutung, dass sie in jedem Fall sichergestellt wird.

Was aber sehen wir, wenn die Stimulation über einen gewissen Zeitraum identisch bleibt? Stellen Sie sich bitte folgendes Experiment vor. Sie betrachten eine statische Projektion einer komplexen Szene, etwa ein Foto Ihres Wohnzimmers. Mithilfe einer Kamera werden Ihre Augenbewegungen registriert. Schließlich wird die Projektion so gesteuert, dass sie mit den Augenbewegungen synchronisiert wird. Trotz Augenbewegungen werden von der Projektion stets dieselben Orte auf der Netzhaut gereizt. Es gibt eine Reizung, aber keine Reizänderung. Ein

stabilisiertes Netzhautbild entsteht. Der Effekt muss als dramatisch bezeichnet werden: Bereits nach einigen Sekunden wird kein Bild mehr gesehen. Die Wahrnehmung des Bildes ist wie ausgelöscht. Ähnliches geschieht im homogenen Ganzfeld bei längerer Exposition. Da die Rezeptoren der Retina auch durch Augenbewegungen nicht neu stimuliert werden können, also auch Zirkularität keine veränderte Stimulation schafft, geht ihre Erregung auf die Grundrate zurück. Eine strukturierte Wahrnehmung ist nicht möglich. Was auf der Ebene der Phänomene bleibt, ist lediglich eine undifferenzierte Grauempfindung, das *subjektive Augengrau*. Einigen Beobachter*innen wird schwarz vor den Augen.

Umgebungswahrnehmung

Der Gegensatz zwischen homogenem Ganzfeld und *natürlich gegliedertem Ganzfeld*, das die alltägliche Umgebungswahrnehmung bestimmt, könnte größer nicht sein (vgl. auch Landwehr 1986, 1988). Die Welt, in der sich der Mensch bewegt, ist eine gefüllte Welt, kein leerer Raum. Im Gesichtsfeld konkurrieren meist mehrere Gegenstände um die Aufmerksamkeit der Beobachterin oder des Beobachters. Das Anforderungsprofil für das Sehsystem ist unter diesen Bedingungen heterogen. Inhomogenitäten sind zu gruppieren, Oberflächen müssen voneinander abgegrenzt werden, Objekte sind zu identifizieren. Ebenso müssen bei Bewegung Objektidentitäten festgestellt oder zwischen Veränderungen der Beleuchtung und der Reflektanz unterschieden werden.

Die Regelmäßigkeiten der physikalischen Welt und die Fovealisierung helfen bei der Bewältigung der vielfältigen Aufgaben. Die vielleicht bedeutendste Regelmäßigkeit unter terrestrischen Bedingungen betrifft die Gliederung des *visuellen Feldes*. Das grundlegende Bezugssystem für die Raumwahrnehmung ist zweigeteilt (vgl. Abbildung 3): Der untere Teil wird mit dem Erdboden ausgefüllt, der obere mit dem Himmel. Beide werden durch den Horizont getrennt. Der Horizont liegt, je nach Standort

und Kopfneigung, höher oder tiefer im Gesichtsfeld und bildet eine waagerechte Linie. Der Boden als die fundamentale Oberfläche erstreckt sich kontinuierlich. Er ermöglicht die Bestimmung der relativen Entfernung von Objekten zur Beobachterin oder zum Beobachter.

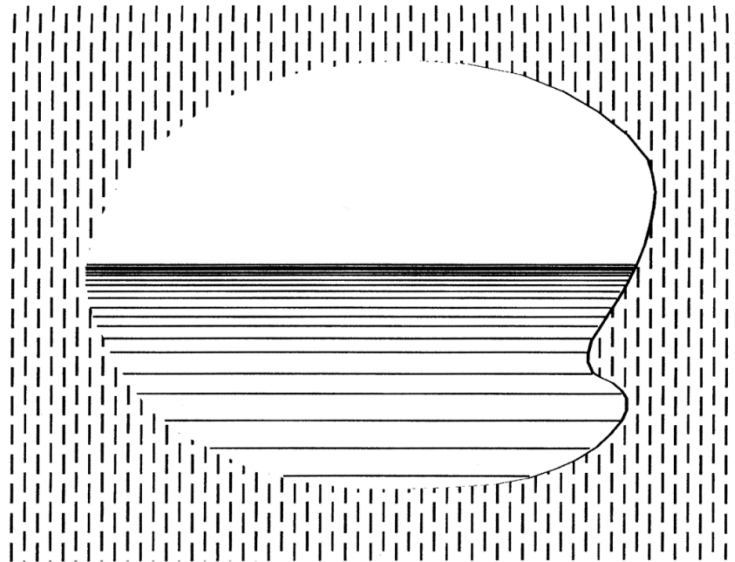


Abbildung 3: Typisches visuelles Feld für ein Auge mit einer Zweiteilung in „Himmel“ und „Erdboden“ durch die Horizontlinie (Gibson, 1950).

Wichtige Regelmäßigkeiten betreffen auch die zeitlichen Abläufe. Während des Tages wechselt sowohl die Lichtintensität als auch die spektrale Zusammensetzung. Durch die Erdrotation variiert der Beleuchtungswinkel mit Auswirkungen auf die Reflexionen an Oberflächen (z. B. Spiegelungen) als auch auf die Schattenbildung. Die meisten Veränderungen sind kontinuierlich und erstrecken sich über einen längeren Zeitraum. Diese Regelmäßigkeiten in den zeitlichen Abläufen von Ereignissen nutzt das visuelle System bei der Interpretation der Stimulationen. Lediglich einige Änderungen in der Lichtintensität, z. B. wenn eine Wolke sich vor die Sonne schiebt, führen zu abrupten oder sprunghaften Helligkeitsveränderungen.

Eine Beobachterin oder ein Beobachter bewegt sich. Der Standort verändert sich permanent und mit ihm variieren Abstand und

Aufsichtswinkel zu den Objekten. Zusätzlich zur Eigenbewegung gibt es oft Objektbewegungen. Sind Informationen nicht ausreichend, um Objekte zu identifizieren, kann die Beobachterin oder der Beobachter durch Lokomotion unterschiedliche Perspektiven auf Gegenstände einnehmen, gegebenenfalls sogar andere Sinnesmodalitäten einsetzen, um die notwendige Sicherheit zu gewinnen. Insgesamt hat die Situation einen explorativen Charakter. Erst durch die Suche nach Invarianten und die Zirkularität des Wahrnehmungsvorgangs wird eine stabile Interpretation der Situation erreicht. Kommt es zu einer Objektbewegung, kann mithilfe von Kopf- und Augenbewegungen eine Scharfstellung eingeleitet werden. Bewegen sich Objekte im Gesichtsfeld, resultieren lokale Veränderungen der Stimulation auf der Netzhaut. Bewegt sich dagegen die Beobachterin oder der Beobachter, verändert sich die gesamte Stimulation. Die Umgebungswahrnehmung ist ein ideales Beispiel für eine dynamische Stimulation.

Für diese Situationen der Umgebungswahrnehmung ist ein System notwendig, das gleichzeitig zwei zunächst kontrovers erscheinende Anforderungen erfüllt: hohe Zuverlässigkeit und hohe Flexibilität. Wir sind es gewohnt, Zuverlässigkeit dann zu erwarten, wenn etwas bei gleicher Ausgangssituation auf die gleiche Weise geschieht. Unsere Vorstellung von Zuverlässigkeit leitet sich von mechanischen Beispielen ab. Der Motor eines Autos ist zuverlässig, wenn er in Abhängigkeit vom Druck auf das Gaspedal stets dieselbe Leistung erbringt. Allerdings setzt diese Auffassung von Zuverlässigkeit eine identische Ausgangssituation voraus. Diese Bedingung ist beim Sehen – streng genommen – nie gegeben. Es gibt keine identischen Stimulationen. Hinzu kommt die Mehrdeutigkeit der Stimulation. Beide Bedingungen machen eine „mechanische Lösung“ ineffizient. In biologischen Systemen wird Zuverlässigkeit oft über Lernprozesse und Flexibilität erreicht. Das visuelle System hat auf verschiedenen Ebenen Mechanismen entwickelt, um bei schwierigen Ausgangsbedingungen effizient arbeiten zu können. Eine

optimale Wahrnehmung setzt, zusammenfassend formuliert, eine komplexe und dynamische Stimulation voraus.

Bildwahrnehmung

Die Wahrnehmung von Bildern ist ein Spezialfall des Sehens. Die Wahrnehmungsforschung beschreibt sie als eine *unterbestimmte Stimulation*. Kontrast und Gruppierung spielen eine bei der Umgebungswahrnehmung vergleichbare Rolle. Die Probleme der Stabilität, Ambiguitätsreduktion und Multiperspektivität stellen sich jedoch anders dar. Zur Vereinfachung soll für den Vergleich zur Umgebungswahrnehmung von zweidimensionalen, statischen Bildern und von konstanten Beleuchtungsbedingungen ausgegangen werden (etwa die Betrachtung eines Gemäldes in einem Museum). Durch die konstante Beleuchtung entfällt eine wichtige Aufgabe der Umgebungswahrnehmung. Denn um Stabilität zu erreichen, muss das visuelle System dort ständig zwischen den Veränderungen der Beleuchtung und den Veränderungen von Objekteigenschaften trennen. Bei konstanter Beleuchtung kann jede Veränderung der Reflexion einer Oberfläche dem Objekt zugeschrieben werden. Da das Objekt überdies stationär ist, kann eine veränderte Stimulation nur durch die Bewegungen der Betrachterin oder des Betrachters ausgelöst werden. Aber auch diese sind stark eingeschränkt. Idealerweise befinden sich Betrachter oder Betrachterin mittig vor dem Bild und haben einen festen Abstand zur Bildebene. Das Bild füllt den überwiegenden Teil des Gesichtsfeldes aus. Die Gliederung des visuellen Feldes in Himmel, Erde und Horizont entfällt. An ihre Stelle tritt die vom Bild vorgegebene Unterteilung (vgl. Abbildung 4). Alle Blickbewegungen beziehen sich auf andere Orte im Bild. Sie werden nicht durch andere Objekte im Gesichtsfeld ausgelöst. Den Strukturmerkmalen des Bildes kommt entscheidende Bedeutung zu. Die visuelle Suche und die Aufmerksamkeitssteuerung sind bildimmanent. Das visuelle Feld wird durch das Bild beherrscht. Das Bild gibt die Kartierung, die Perspektive und die Gruppierungen vor.



Abbildung 4: Piero della Francesca, Geißelung Christi, ca. 1465. Bildmerkmale wie Perspektive und Kontraste steuern die Augenbewegungen.

Da Bild und Betrachter*in gleichermaßen stationär sind, entfallen Bewegungshinweise. Das hat weitreichende Konsequenzen für das Eindeutigkeitsproblem. Ein Beispiel: Bewegungen von Objekten im Gesichtsfeld führen in der Umgebungswahrnehmung abwechselnd zu einer Verdeckung und Freilegung anderer Objekte. Der systematische Wechsel erleichtert die Objektidentifikation und die Bestimmung der Entfernungsverhältnisse zwischen den Gegenständen. Die Veränderung des Standortes in einer Umgebung mit unterschiedlich weit entfernten Objekten führt dazu, dass nahe Objekte im Gesichtsfeld stärker verschoben werden als weit entfernte (Bewegungsparallaxe). Da sich bei der Bildwahrnehmung im Gesichtsfeld nur das Bild ohne Tiefenausdehnung befindet, entfällt auch die Bewegungsparallaxe als Hinweisreiz. Das Fehlen sämtlicher Bewegungshinweise konstituiert einen der fundamentalen Unterschiede zwischen der Bild- und der Umgebungswahrnehmung.

Die meisten Situationen der Umgebungswahrnehmung sind komplexer als eine typische Bildwahrnehmungssituation. Aber das visuelle System ist darauf ausgerichtet, aus den Veränderungen der Stimulation die Invarianten zu ermitteln und über die Zirkularität ein eindeutiges Perzept zu erstellen. Bei der Bildwahrnehmung ist die Stimulation prinzipiell unterbestimmt. Diese Unterbestimmtheit kann nicht durch die Aktivität der Betrachterin oder des Betrachters ausgeglichen werden. Damit steigt die Wahrscheinlichkeit von Mehrdeutigkeiten. Die meisten visuellen Illusionen finden sich im Bereich der Bildwahrnehmung. Sie stellt zwar geringere Anforderungen an die Konstruktion einer stabilen Welt, aber sie stellt höhere Ansprüche an die Ambiguitätsreduktion. Letztlich ist bei Bildern keine Eindeutigkeit zu erreichen. Hierzu müsste die Betrachterin oder der Betrachter den Ausschnitt und die Perspektive wechseln können. Die Referenzsituation für die weiteren Überlegungen ist die Umgebungswahrnehmung. Auf den Spezialfall der Bildwahrnehmung wird vor allem bei der Beschreibung einzelner Heuristiken Bezug genommen. Denn viele Gesetzmäßigkeiten des Sehens treten in unterbestimmten Stimulationen deutlicher hervor. Gerade die Systematik der Fehler erlaubt einen Einblick in die produktive Arbeitsweise des Sehsystems.

Trompe-l'œil

Das Trompe-l'œil ist ein Spezialfall der Bildwahrnehmung. Ein Bild ist so täuschend echt, dass es mit der Wirklichkeit verwechselt wird. Obwohl es sich um eine extreme Situation handelt, ist sie von besonderer theoretischer Bedeutung. Die Bedingungen, unter denen es gelingt, bei der Betrachterin oder dem Betrachter eine vollständige Illusion von Wirklichkeit zu erzeugen, sind aufschlussreich für die Gesetze des Sehens. Illusionistische Malerei hat, ausgehend von einem reduktionistischen Modell der Wahrnehmung, bereits vor der experimentellen Forschung viele Gesetze des Sehens entdeckt (vgl. Spillmann 2007). Plinius (23-79)

berichtet vom Wettstreit der griechischen Maler Zeuxis und Parrhasios um die Frage, wer der bedeutendste Maler seiner Zeit sei. Zeuxis malte Weintrauben so täuschend echt, dass sie Vögel anlockten. Überzeugt, dass mit diesem Urteil der Natur auch der Wettstreit entschieden sei, ging er in das Atelier seines Kollegen Parrhasios und forderte ihn auf, den Vorhang, der das neue Bild bedeckte, zurückzuziehen. Der Vorhang aber war das Bild.

Um eine Illusion von Wirklichkeit zu erzeugen, muss eine Reihe von Bedingungen erfüllt sein. Nicht jedes illusionistische Werk ist ein Trompe-l'œil. Bei der perspektivischen Darstellung eines Gegenstandes besteht, insbesondere wenn sie virtuos gehandhabt wird, ein Dilemma: Die korrekte Perspektive ist die Voraussetzung für die Illusion. Je perfekter die Gestaltung des Bildraumes aber wird, umso größer wird auch die Abhängigkeit vom Standort. Die Täuschung einer sich bewegenden Betrachterin oder eines sich bewegenden Betrachters wird erschwert. Milman (1984, S. 36) fasst, nach einer Analyse zahlreicher Beispiele aus unterschiedlichen Epochen, die „Spielregeln“ des Trompe-l'œil zusammen: Das Trompe-l'œil muss naturgetreue Maße wiedergeben. Es muss sich in seine physische Umwelt vollkommen integrieren, Platzierung und Rahmung müssen einer einheitlichen Logik entsprechen. Keines der im Bild wiedergegebenen Elemente darf wegen der begrenzten Bildfläche unvollständig erscheinen. Die dritte Dimension sollte nicht durch eine zu tief geführte Perspektive oder ein zu stark hervortretendes Relief erreicht werden. Der Pinselstrich muss „unsichtbar“ sein, die Farben müssen weich ineinander übergehen. Schließlich ist die Darstellung menschlicher Figuren zu vermeiden, da, im Gegensatz zu dinglichen Objekten, ihre starre Erscheinung den Realitätscharakter des Trompe-l'œil beeinträchtigt (vgl. Abbildung 5).

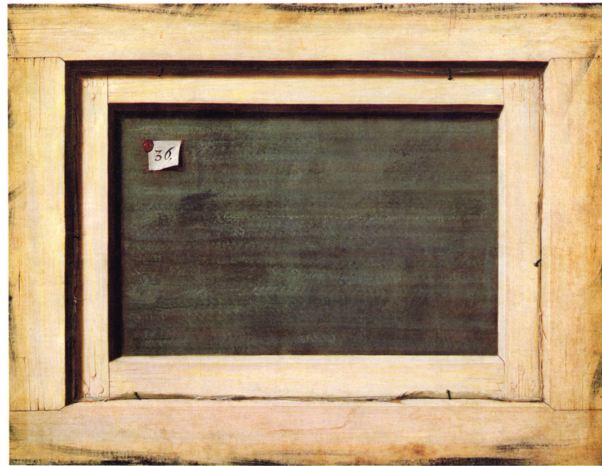


Abbildung 5: Beispiel für ein Trompe-l'œil, Gemälde von Cornelius Norbertus Gijysbrechts, um 1670, Öl auf Leinwand.

Alles spricht dafür, dass die Maler*innen des Trompe-l'œil sich Heuristiken der Umgebungswahrnehmung zunutze machen. Der Oberflächenaspekt des Bildes, seine Textur, wird negiert und der Repräsentationsaspekt betont. Mit der Übertragung der Regeln der Umgebungswahrnehmung auf die Bildwahrnehmung ist eine Aufhebung der Trennung von „Bild als Oberfläche“ und „Bild als Repräsentation“ intendiert. Der Bildcharakter soll nicht erkannt – oder zunächst nicht erkannt werden. Daher sind die Präsentationsbedingungen bedeutsam. Sie signalisieren Gegebenheiten der Umgebungswahrnehmung und die Übernahme ihrer Probabilitäten. Letztlich liegt die Attraktivität des Trompe-l'œil aber nicht in der Täuschung, sondern in ihrer Überwindung und der Desillusionierung der Betrachterin oder des Betrachters.

Virtuelle Realität

Wie beim Trompe-l'œil steht auch bei der virtuellen Realität die Verwechslung von Bild und Wirklichkeit im Mittelpunkt. Während beim Trompe-l'œil diese allerdings durch eine Veränderung des Standortes (z. B. eine Annäherung an die Bildoberfläche) aufgelöst werden kann, soll in der virtuellen Realität das Medium dauerhaft aus dem Bewusstsein verschwinden. Hierzu müssen die Gegebenheiten der

Umgebungswahrnehmung vollständig auf das Bild und seine Wahrnehmung übertragen werden. Das Bild muss nicht nur wie beim Trompe-l'œil eine Raum- und Lichtsimulation zeigen, sondern sich auch in Abhängigkeit von den Augen-, Kopf- und Rumpfbewegungen der Nutzer*innen in Echtzeit verändern. Ausschnitt, Perspektive und Auflösung müssen permanent an die Positionsveränderungen und Blickbewegungen angepasst werden. Daher ist ein Trompe-l'œil noch keine virtuelle Realität. Diese ist aus Sicht der Wahrnehmungsforschung erst gegeben, wenn die Nutzerin oder der Nutzer keine Möglichkeit mehr hat, das Bild als Bild zu identifizieren, die Bildebene zu bestimmen.

Neben der bewegungsunabhängigen Abbildungstreue werden als weitere Anforderungen an die virtuelle Realität häufig auch die Immersion und die Interaktivität genannt. Wie in der Umgebungswahrnehmung sollen Wahrnehmende sich nicht als Außenstehende, sondern als Bestandteil der Szene empfinden. Wird Interaktion verlangt, muss es Einflussmöglichkeiten auf die nachfolgenden Bilder geben. Als Akteurin oder Akteur muss eine Kausalität der Bilder zu vorausgegangenen Handlungen hergestellt werden können. Beide Bedingungen lassen sich in der Regel nur erfüllen, wenn der Handlungsspielraum eingeschränkt wird und die zuvor ausgewählten Handlungsalternativen plausibel sind. Plausibilität wiederum ist nicht nur an eine Widerspruchsfreiheit gebunden, sondern die Ereignisse müssen auch mit den aus der Umgebungswahrnehmung vertrauten Wahrscheinlichkeiten übereinstimmen.

Ein statisches Bild kann zwar zum Trompe-l'œil werden, nicht aber zur virtuellen Realität. Diese ist immer ein Film. Hinzu kommt die Binokularität: Dem rechten und dem linken Auge müssen zwei getrennte und leicht verschobene Bilder gezeigt werden. Aber auch ein „interaktiver binokularer Film“ ist noch keine virtuelle Realität. Erst wenn keine

Bildränder mehr sichtbar sind und der Film permanent das gesamte Gesichtsfeld ausfüllt, werden Bilder zur Realität.

Damit stellt sich ein neues Problem. In der Umgebungswahrnehmung ist die visuelle Wahrnehmung eng mit der Körpermotorik und den übrigen Sinnesmodalitäten verbunden. Wir erwarten kongruente Informationen von den verschiedenen Sinnessystemen (z. B. zwischen dem visuellen System und dem Gleichgewichtssinn oder dem Tastsinn). Kommt es zu Widersprüchen, können wir diese in der Umgebungswahrnehmung durch unsere Aktivität aufklären (z. B. indem wir die Dinge berühren). Bei Widersprüchen zwischen den Sinnesmodalitäten, die ihre Ursache darin haben, dass der eine Teil der Stimulation aus der Realität, der andere dagegen aus der virtuellen Realität stammt, ist das nicht möglich. Eine Kongruenz kann nur hergestellt werden, indem die virtuelle Realität verlassen wird. Bleiben Inkonsistenzen über längere Zeit bestehen, kommt es leicht zu Empfindungsstörungen. Die Folge ist z. B. die bekannte Bewegungsübelkeit (motion sickness) beim Tragen von VR-Brillen (zur Verwendung von VR-Methoden in den kognitiven Neurowissenschaften vgl. Tarr und Warren, 2002; Bohil, Bradly und Biocca, 2011).

4. Ideale Beobachterin und idealer Beobachter

Die Leistungen des visuellen Systems sind nur unter Berücksichtigung seiner Entstehungsgeschichte (Phylogenese) zu verstehen. Sie sind das Ergebnis eines langen, evolutionären Selektionsprozesses. Eine optimale Anpassung an die Gegebenheiten und Anforderungen der Lebensumwelt ist der Schlüssel für das Überleben einer Spezies. Doch wie sieht die optische Umgebung aus, an die sich das visuelle System des Menschen adaptiert hat? Was sind ihre wesentlichen Merkmale? Viele Reize, mit denen wir heute im Alltag konfrontiert sind, sind aus evolutionsbiologischer Perspektive neue Reize. Dies gilt nicht nur für Fernsehen, Internet und Kino, sondern auch für den Straßenverkehr und

viele Objekte des Alltags. Phylogenetisch gesehen sind sie wenig bedeutsam. Einflussreich waren vor allem charakteristische Landschaften (vgl. Abbildung 6). Die Analyse dieser *natürlichen Szenen* verspricht, einen wichtigen Beitrag zum Verständnis des Sehens zu leisten, denn das visuelle System des Menschen hat sich auf solche Szenen hin optimiert.



Abbildung 6: Beispiele für „natürliche Szenen“, auf die das visuelle System gut adaptiert ist.

Aufgrund ähnlicher Überlegungen haben Wahrnehmungsforscher ihr Arbeitsgebiet über die Beschreibung der Phänomene oder der physiologischen Prozesse hinaus auf die Analyse der Reize ausgedehnt. Als besonders fruchtbar hat sich eine stochastische Betrachtungsweise erwiesen. Wie erläutert, sind die proximalen Stimuli keinesfalls eindeutig, sondern lassen sich einer Vielzahl von distalen Stimuli zuordnen. Der Prozess der Informationsverarbeitung dient der Reduktion dieser Ambiguitäten. Zumindest Teile dieses Prozesses können als Inferenzprozess aufgefasst werden, als eine Entscheidung unter Unsicherheit: Welcher distale Reiz hat die aktuelle Stimulation am ehesten verursacht? Welcher hat die höchste Wahrscheinlichkeit? In diesem Sinne

ist die Wahrnehmung die Vorhersage dessen, was in der Außenwelt zu erwarten ist.

Auf das Problem der Ambiguitätsreduktion bezieht sich die *Theorie der idealen Beobachterin oder des idealen Beobachters*. Angelehnt an das Theorem des englischen Mathematikers Thomas Bayes (1702-1761), das die Berechnung *bedingter Wahrscheinlichkeiten* erlaubt, beschreibt diese Theorie, welche Leistung eine ideale Beobachterin oder ein idealer Beobachter vollbringen müsste, um die Mehrdeutigkeit des proximalen Reizes optimal zu reduzieren. Als bedingte Wahrscheinlichkeit ist die Wahrscheinlichkeit eines Ereignisses A definiert, unter der Bedingung, dass ein Ereignis B bereits eingetreten ist: $p(A|B)$. Der Satz von Bayes liefert eine Formel, mithilfe derer bedingte Wahrscheinlichkeiten ineinander überführt werden können, wie wir von $p(A|B)$ nach $p(B|A)$ kommen. Es gilt:

$$p(B|A) = \frac{p(A|B) \cdot p(B)}{p(A)}$$

Zur Erläuterung des Grundgedankens ein Beispiel: Wenn ein proximaler Stimulus sowohl eine essbare Frucht als auch das Blatt eines Baumes sein könnte, müssen Entscheidungsregeln gefunden werden, die eine möglichst zuverlässige Unterscheidung zwischen Früchten und Blättern erlaubt. Einschränkende Faktoren sind lediglich das Auflösungsvermögen der Retina oder die Anzahl der beteiligten Neuronen. Sofort wird die Verbindung zur evolutionsbiologischen Perspektive deutlich: Wenn das visuelle System des Menschen in der Folge von Selektionsprozessen tatsächlich die bestmögliche Anpassung an seine Umwelt darstellt, sollten die Resultate, die es erzielt – gerade bei der Betrachtung natürlicher Szenen – denen einer idealen Beobachterin oder eines idealen Beobachters ähneln. Umgekehrt: Lassen sich bei der idealen Beobachterin oder dem idealen Beobachter Algorithmen oder Heuristiken beschreiben, die unter Einbeziehung der bekannten biologischen und physikalischen Beschränkungen eine optimale Vorhersage erlauben, sollten diese

Algorithmen oder Heuristiken auch zum Verständnis der visuellen Informationsverarbeitung des Menschen beitragen. Chris Frith (2010, S. 175) hat es so formuliert: „Meine Wahrnehmung ist eine Vorhersage dessen, was in der Außenwelt sein sollte. Und diese Vorhersage wird ständig durch Handeln getestet.“

Auf diese Weise wird das Modell der visuellen Wahrnehmung auf ein Wahrscheinlichkeitsmodell mit der Annahme übertragen, dass beide Modelle in wichtigen Punkten vergleichbar sind. Sofern dies der Fall ist, können gefundene Gesetzmäßigkeiten im Wahrscheinlichkeitsmodell auf das visuelle System zurück übertragen werden. Beide Schritte ergeben nur Sinn, wenn das Wahrscheinlichkeitsmodell Vorteile bietet gegenüber dem ursprünglichen biologischen Modell, wenn es zum Beispiel präzisere Vorhersagen erlaubt, wie bei einer gegebenen Stimulation das Wahrnehmungserlebnis aussehen wird.

Ob der Beobachterin oder dem Beobachter ein korrekter Schluss gelingt, hängt in hohem Maße davon ab, inwiefern tatsächlich vorhandene Gegenstände auch richtig identifiziert werden können (Treffer). Die Theorie der idealen Beobachterin oder des idealen Beobachters beschreibt diesen Zusammenhang in Form von Wahrscheinlichkeitsverteilungen. Die Aufgabe wird als ein Vergleich bedingter Wahrscheinlichkeiten verstanden: Ist es wahrscheinlicher, dass der proximale Stimulus (s) eher dem distalen Reiz (ω) „Blatt“ oder dem distalen Reiz (ω) „Frucht“ entspricht. Also ist $\omega = \text{Blatt}$ oder ist $\omega = \text{Frucht}$ wahrscheinlicher?

Formalisiert handelt es sich um die Wahrscheinlichkeit $p(\omega|s)$, die zu ermitteln ist. Um zu entscheiden, ob auf einem Teil der Retina eher ein Blatt oder eine Frucht abgebildet ist, können eine Reihe anderer Wahrscheinlichkeiten herangezogen werden: Zum einen die Wahrscheinlichkeit $p(\omega)$, dass überhaupt ein Blatt oder eine Frucht mit objekttypischen Reflexionseigenschaften in dieser Szene vorkommt (bei

einer Meerlandschaft wäre beides extrem unwahrscheinlich). Zum anderen die bedingte Wahrscheinlichkeit $p(\mathbf{s}|\omega)$, dass eine vorhandene Frucht Licht mit einer dem Stimulus entsprechenden Wellenlänge auf die Retina reflektiert (z. B. gibt es viele reife Früchte im gelb-rot Bereich). Werden diese beiden Wahrscheinlichkeiten für alle möglichen Objektausprägungen (hier also für $\omega = \text{Blatt}$ und $\omega = \text{Frucht}$) empirisch bestimmt, ergibt sich eine gemeinsame Wahrscheinlichkeitsstatistik $p(\omega, \mathbf{s})$:

$$p(\omega, \mathbf{s}) = p(\mathbf{s}|\omega) \cdot p(\omega)$$

Teilen wir $p(\omega, \mathbf{s})$ nun durch $p(\mathbf{s})$, folgt hieraus nach dem Satz von Bayes die Wahrscheinlichkeit $p(\omega|\mathbf{s})$. Eine solche Statistik zeigt sich beispielsweise, wenn die Farbspektren verschiedener Blätter und Früchte, die die Retina erreichen, analysiert werden. Wir verdeutlichen dies an einem Beispiel: Bei einem Waldspaziergang sehen wir ein rotes Objekt bzw. eine rote Oberfläche. Angenommen, die Wahrscheinlichkeit, dass ein Objekt im Wald eine Frucht ist, sei 1 % ($p(\text{Frucht}) = 0,01$), die Wahrscheinlichkeit für ein Blatt läge bei 10 % ($p(\text{Blatt}) = 0,1$). Weiterhin reflektiert eine Frucht zu 80 % Licht mit einer dem Stimulus entsprechenden Wellenlänge auf die Retina, während dies bei einem Blatt nur zu 20 % der Fall ist (also $p(\mathbf{s}|\text{Frucht}) = 0,8$ und $p(\mathbf{s}|\text{Blatt}) = 0,2$). Außerdem kommt das rote Objekt zu 2.8 % in der Umgebung vor, sprich $p(\mathbf{s}) = 0,028$. Jetzt können wir mithilfe des Bayes-Theorems $p(\text{Frucht}|\mathbf{s})$ und $p(\text{Blatt}|\mathbf{s})$ berechnen. Es folgt $p(\text{Frucht}|\mathbf{s}) = 0,29$ und $p(\text{Blatt}|\mathbf{s}) = 0,71$. Somit ist es deutlich wahrscheinlicher, dass es sich bei dem roten Objekt um ein Blatt handelt. Allerdings ist dies nur ein sehr einfaches Beispiel. In der Realität ist die Berechnung in der Regel deutlich aufwändiger.

Um derartige Statistiken sinnvoll zu bestimmen, schlug Geisler (2008) vor, die unterschiedlichen Stufen der Wahrnehmung in vier verschiedene Ebenen oder Domänen zu unterteilen (vgl. Abbildung 7): den distalen Stimulus, den proximalen Stimulus, die neuronale Aktivität und die Reaktion bzw. das Verhalten, das aus der neuronalen Verarbeitung

resultiert. Für jede dieser Domänen lassen sich zunächst *domäneninterne Wahrscheinlichkeitsverteilungen* betrachten: Verteilungen, die von den anderen Domänen unabhängig sind. Für die Domäne der physikalischen Eigenschaften könnten sich solche Verteilungen $p(\omega)$ aus der Untersuchung der Reflexionseigenschaften verschiedener Oberflächen bei unterschiedlichen Lichtverhältnissen ergeben. Verteilungen von proximalen Stimuli $p(s)$ könnten darstellen, welche Bereiche der Retina Licht welcher Wellenlängen absorbieren. Ähnliche domäneninterne Statistiken über die Aktivitätsverteilungen von Neurone $p(z)$ oder über Fixationen im Gesichtsfeld $p(r)$ sind denkbar. Die für die Erforschung des Sehens aufschlussreicheren Statistiken sind *domänenübergreifende Wahrscheinlichkeitsverteilungen*. Sie beschreiben die Interaktionen zwischen verschiedenen Ebenen. Theoretisch sind Interaktionen zwischen allen vier Domänen denkbar.

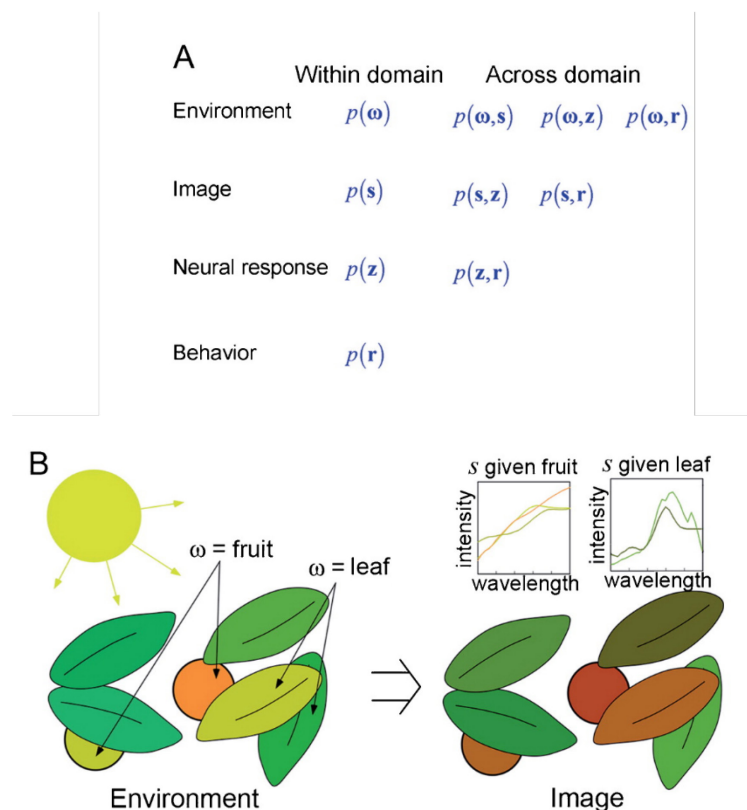


Abbildung 7: Statistiken natürlicher Szenen (Geisler, 2008). (A) Es gibt zwei grundlegende Arten von Statistiken: domäneninterne Statistiken und

domänenübergreifende Statistiken. (B) Hypothetisches Beispiel zur Unterscheidung zwischen einem „Blatt“ und einer „Frucht“. Erläuterung im Text.

Während die domäneninternen Statistiken eine Analyse der Eigenschaften natürlicher Szenen ermöglichen, lassen sich aus den domänenübergreifenden Statistiken in Kombination mit dem Theorem von Bayes Hypothesen über die Arbeitsweise des visuellen Systems ableiten. Dabei ist der bayes'sche Ansatz nur ein mögliches – wenn auch prominentes – Modell, um (statistische) Inferenz in den Wahrnehmungsprozess einzubeziehen. Eine Alternative wäre z. B. die Evidenztheorie von Dempster und Shafer (Dempster, 1967; Shafer, 1976), bei der unsichere Aussagen aus verschiedenen Informationsquellen zu einem Gesamtwert kombiniert werden können, ein Vorgehen, das z. B. bei der Entwicklung „sehender Maschinen“ oder bei der Integration von Informationen aus einzelnen Sinnessystemen an Bedeutung gewinnt.

Kapitel II: Visuelle Reize und neuronale Prozesse

Im Mittelpunkt der Überlegungen zum Sehen stehen meist die Gegenstände und die Wahrnehmenden. Die Rolle des Lichtes wird selten thematisiert. Allenfalls als Medium für die Übertragung der optischen Information anerkannt, wird wenig bedacht, inwieweit es die optischen Informationen selbst gestaltet. Ein Grund könnte darin liegen, dass Licht oft ausschließlich als physikalisches Phänomen betrachtet wird. Vernachlässigt werden seine Zusammenhänge zum Sinnessystem und zur Reizverarbeitung. Auch auf der Ebene der Phänomene ist das Licht kein Gegenstand der Aufmerksamkeit. Das ändert sich, wenn es als „zu grell“, „zu dunkel“ oder in seiner Farbe als „unnatürlich“ eingestuft und als störend empfunden wird.

Aus Sicht der physikalischen Optik besteht Licht aus kleinen Einheiten von Energie, die sich unter bestimmten Bedingungen wie eine Welle und unter anderen wie ein Teilchen (Photon) verhalten (Welle-Teilchen-Dualismus). In Bezug auf das menschliche Sehen ist – das Farbsehen ausgenommen – zumeist der Teilchencharakter interessant. Von einer Energiequelle (z. B. der Sonne) ausgehend breitet sich *Strahlung* radial und geradlinig mit einer Geschwindigkeit von annähernd 300.000 Kilometern pro Sekunde aus. Innerhalb des Spektrums der elektromagnetischen Wellen bildet der Teil, den wir das „sichtbare Licht“ nennen, nur einen kleinen Ausschnitt (ca. 400-700 Nanometer) und wird deshalb – metaphorisch – als das „Fenster zur Welt“ bezeichnet. Der überwiegende Teil der elektromagnetischen Wellen, wie Radiowellen oder Röntgenstrahlung, sind für das Auge unsichtbar, zum Teil sogar schädlich (vgl. Abbildung 8).

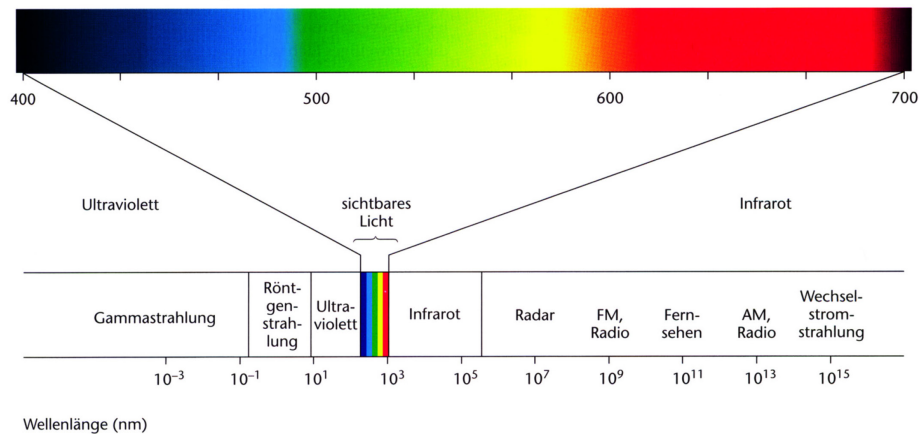


Abbildung 8: Elektromagnetisches Spektrum. Lediglich für einen kleinen Ausschnitt aus dem gesamten Wellenlängenbereich (ca. 400-700 Nanometer) sind die Fotorezeptoren der Retina sensitiv.

Treffen Photonen auf eine opake Oberfläche, wird ein Teil von ihnen absorbiert und vor allem in Wärmeenergie umgewandelt. Der übrige Teil wird reflektiert. Dieser reflektierte Anteil wird als *Leuchtdichte* (Luminanz) bezeichnet und ermöglicht die Wahrnehmung der Oberfläche. Bei nahezu vollständiger Absorption erscheint ein Körper schwarz, bei nahezu vollständiger Reflexion weiß. Durch selektive Reflexion einzelner Wellenlängenbereiche entsteht die Stimulation für einen spezifischen Farbeindruck. In Abhängigkeit von der Struktur der Oberfläche kann es zu einer diffusen Reflexion (unebene Oberflächen) oder zu einer spiegelnden Reflexion (glatte Oberflächen) kommen. Die tatsächlichen Verhältnisse von Brechung und Reflexion an natürlichen Oberflächen sind weitaus komplexer. Zum Beispiel entstehen, wenn (weißes) Licht auf eine geschichtete Oberfläche trifft, leicht Überlagerungen (Interferenzen). Sie sind in der Lage, den Eindruck des „farbigen Schillerns“ hervorzurufen (ein Beispiel hierfür ist die Oberfläche von Perlmutt). Helligkeit und Farbe sind demnach keine physikalischen Größen, sondern sie charakterisieren den phänomenalen Eindruck, den die Intensitäten bzw. Wellenlängen bei Wahrnehmenden hervorrufen. Es sind Eigenschaften des Perzeptes, nicht des Reizes.

Doch eine solche physikalische Optik reicht nicht aus. Sie beschreibt zwar die Funktionsweise eines Fotoapparates adäquat, nicht aber das menschliche Sehen. Zur Erklärung der Wahrnehmungsleistung von Organismen wird eine Optik benötigt, die die Bedeutung des Wahrnehmungsgegenstandes für den Organismus miteinbezieht. Gibson (1979) bezeichnete sie als „ökologische Optik“. Erst diese veränderte Art der Beschreibung macht verständlich, warum es für das visuelle System effektiv ist, mit Voreinstellungen und Probabilitäten zu arbeiten. Das visuelle System des Menschen ist auf ein terrestrisches Sehen bei aufrechtem Gang abgestimmt. Auf diese Bedingungen hin hat es sich im Laufe seiner Geschichte entwickelt.

1. Ökologische Optik

Die Erdoberfläche besteht aus Substanzen, meist undurchsichtige feste Stoffe oder Flüssigkeiten. Ihre Oberflächen haben eine charakteristische materielle und eine optische Struktur. Letztere ist gemeint, wenn im Folgenden von *Textur* die Rede ist: Die Rinde eines Baumes hat eine andere Textur als die Oberfläche eines Felsens. Doch auch die Rinde einer Buche hat eine andere Textur als die Rinde einer Eiche. Zentrales Beschreibungsmerkmal für Texturen ist die *Texturdichte*. Sie gibt an, wie viele unterscheidbare Elemente pro Flächeneinheit vorkommen. Blickt ein Beobachterin oder ein Beobachter auf eine texturierte Oberfläche, nimmt die Dichte in Abhängigkeit von der Entfernung stetig zu. Auch mit dem Aufsichtswinkel variiert die Texturdichte systematisch. Diese Beziehung wird als *Texturgradient* bezeichnet. Ist der Aufsichtswinkel auf eine Oberfläche genau 90° , so ist dieser Gradient radialsymmetrisch. Bei jedem anderen Aufsichtswinkel ergibt sich eine systematische Zu- oder Abnahme der Texturdichte. Mithilfe des Texturgradienten können so die Position und die Neigung von Oberflächen im Raum bestimmt werden (Abbildung 9).

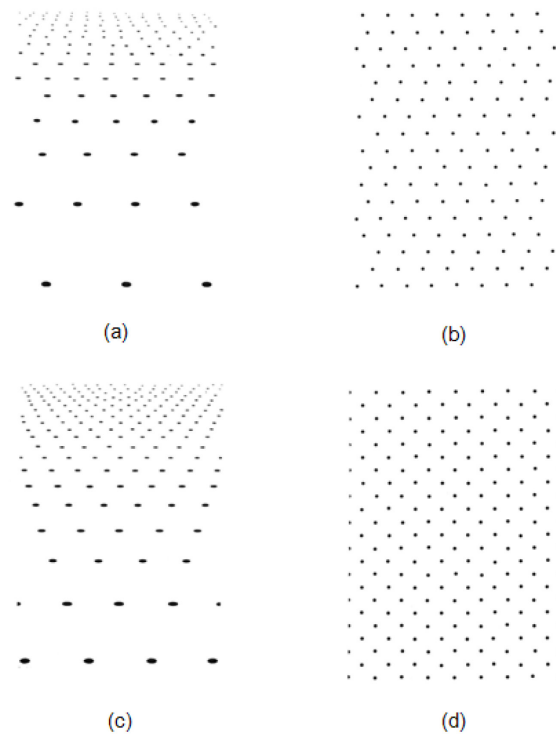


Abbildung 9: Punkt- bzw. Fleckenverteilungen, die den Eindruck einer longitudinalen (a und c) und einer frontalen Oberfläche (b und d) hervorrufen; (a) und (c) einerseits sowie (b) und (d) andererseits weisen jeweils denselben Texturgradienten auf, allerdings einmal mit einer unregelmäßigen (a und b) und einmal mit einer regelmäßigen Anordnung (c und d) der einzelnen Elemente (Gibson, 1950).

Das strukturierte Licht ist der Informationsträger. Es ist selbst nicht sichtbar. Eine Vorstellung von der Struktur des Umgebungslichts gewinnen Wahrnehmende am ehesten dadurch, dass sie den Blick nicht auf eine Szene als Ganzes richten, sondern auf ein Detail. Wenn sie sich in einem zweiten Schritt vom Gegenstandscharakter der Dinge lösen, und auf die Textur der Oberflächen konzentrieren, können sie in einem dritten Schritt ein Bild als aneinandergrenzende Flächen unterschiedlicher Textur zusammensetzen. Damit ergibt sich eine gute Annäherung an die

Grundlage, die dem visuellen System als Ausgangspunkt für seine Verarbeitung dient. Dieser Vorgang bedarf der Übung, weil die Wahrnehmung spontan eine Figur-Grund-Unterscheidung vornimmt und in der Folge die gegenstandsbezogene Wahrnehmung dominant ist. Hieran zeigt sich eindrucksvoll, wie schwer es ist, die im visuellen System angewandten Prinzipien (z. B. die Kontrastbildung oder die Gruppierung) bewusst in ihrem Wirken einzuschränken.

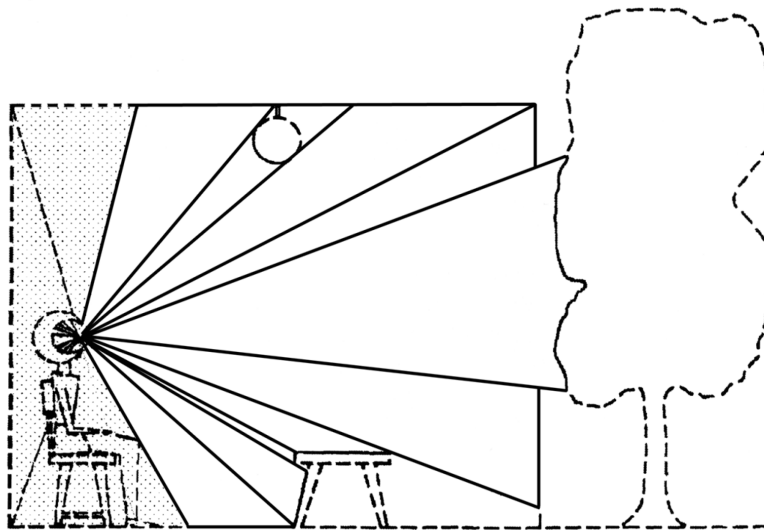


Abbildung 10: Optische Anordnung. Sie wird gebildet aus den Seh- oder Raumwinkeln und ist spezifisch für einen Standort (Gibson, 1979).

Typischerweise besteht eine Umwelt nicht aus einzelnen, parallel angeordneten Oberflächen, sondern aus untereinander verschachtelten Einheiten. Entsprechend sind auch die zugehörigen Raum- oder Sehwinkel untereinander gestuft. Alle Sehwinkel zusammen bilden die *optische Anordnung*. Sie ist definiert als das durch die Oberflächen strukturierte Umgebungslicht, welches auf den Ort im Raum hin konvergiert, an dem sich die Beobachterin oder der Beobachter befindet. Dort ist der Scheitelpunkt aller Sehwinkel (Abbildung 10). Bewegt sich die Beobachterin oder der Beobachter, verändert sich auch die optische Anordnung und in der Folge das Netzhautbild. Es entstehen *Fließmuster* (Abbildung 11), die Ausmaß und Richtung der Veränderung auf der

Netzhaut anzeigen. Gewisse Eigenschaften der Reizverhältnisse und damit auch der optischen Fließmuster bleiben auch bei einer Bewegung konstant. Diese *Invarianten* erlauben eine Orientierung im Raum oder auch die Identifikation von Objekten.

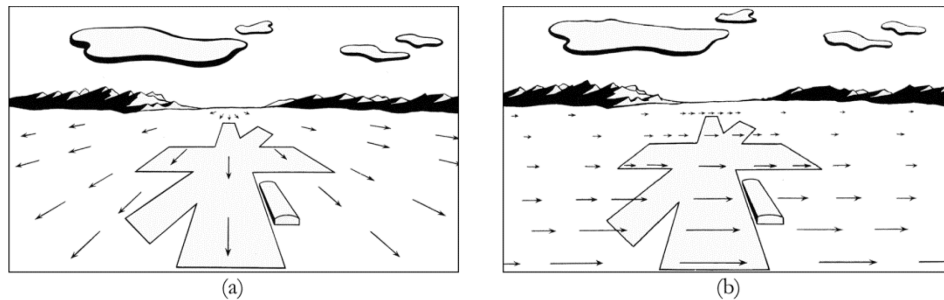


Abbildung 11: Optisches Fließmuster. Die Länge der Pfeile gibt die relative Geschwindigkeit an, mit der sich bei einer Bewegung der Beobachterin oder des Beobachters die Abbildung auf der Retina verschiebt. (a) Bewegung nach vorne, (b) Bewegung nach links (Gibson, 1950).

Nicht nur bei Oberflächen, sondern auch bei Beleuchtungssituationen dominieren Regelmäßigkeiten. Zwei Lichtarten sind zu unterscheiden: das *Strahlungslicht* und das *Umgebungslicht*. Ersteres bezeichnet das von einer Energiequelle ausgesendete Licht, letzteres ist das von Oberflächen reflektierte Licht. Für die menschliche Wahrnehmung ist das Umgebungslicht maßgebend. Die Struktur dieses Lichtes wird durch die Textur und die Materialeigenschaften der Oberflächen bestimmt, von denen es reflektiert wird. Die Regelmäßigkeiten dieser Oberflächen in Kombination mit den für unsere Lebensbedingungen typischen Beleuchtungssituationen bilden die wichtigsten Wahrscheinlichkeiten, mit denen das visuelle System eine Stimulation interpretiert. Sie markieren den Ausgangspunkt für die ständig stattfindenden Interpolationen. Unter den Bedingungen der Umgebungswahrnehmung ist eine zentrale Aufgabe die Entscheidung über Ursachen für eine veränderte Stimulation. Handelt es sich um einen Wechsel in der Beleuchtung oder um eine neue Oberfläche? Ist es ein Objekt, welches teilweise von der Sonne beschienen wird und teilweise im Schatten liegt, oder resultiert die Stimulation von zwei verschiedenen Objekten? Möglicherweise versucht sich ein Objekt zu

tarnen. Derselbe proximale Reiz kann sowohl von einer geringen Strahlung und einer wenig absorbierenden Oberfläche als auch von einer starken Strahlung und einer stark absorbierenden Oberfläche stammen. In beiden Fällen heben sich die physikalischen Größen Strahlung und Absorption gegenseitig auf. Mehrdeutigkeiten bei komplexen Reizverhältnissen schnell und zuverlässig aufzulösen ist eine anspruchsvolle Aufgabe, ohne die keine Stabilität der visuellen Welt zu erreichen ist.

Zur Charakteristik terrestrischer Beleuchtungssituationen gehört, dass sie unter natürlichen Bedingungen mit Sonne oder Mond nur eine Lichtquelle aufweisen und dass die Gegenstände bzw. die Oberflächen „von oben“ und nicht „von unten“ beleuchtet werden. Eine dritte Regelmäßigkeit betrifft die Art und Weise, wie sich die natürliche Beleuchtung verändert. Dies geschieht, wie bereits erwähnt, überwiegend graduell (z. B. Tag- und Nachtübergänge). Ausnahmen sind vergleichsweise selten. Die Regelmäßigkeiten werden als Voreinstellung vom visuellen System übernommen und sind eine wichtige Hilfe bei der Ambiguitätsreduktion. Die angewendeten leistungsstarken Heuristiken können jedoch – vor allem unter verarmten Wahrnehmungsbedingungen – zu Fehlinterpretationen führen. Flächen werden als unterschiedlich hell wahrgenommen, obwohl sie die gleiche Leuchtdichte haben (vgl. später die Beispiele Simultankontrast und amodale Konturen).

2. Optische Bilder

Das Auge hat zwei Aufgaben: Erstens muss es das Licht so bündeln, dass eine scharfe Abbildung entsteht und zweitens muss die Lichtenergie in neuronale Impulse umgewandelt werden. Für die Untersuchung des ersten Aspektes ist es üblich geworden, das Auge als einfaches optisches System mit nur einem Brennpunkt zu behandeln. Anatomisch betrachtet ist das Auge ein kugelförmiger, mit Flüssigkeit gefüllter Hohlkörper (Abbildung 12). Es hat einen Durchmesser von etwa 24 Millimetern und wird von

einer weißen Lederhaut (Sklera) umschlossen, die an der Vorderseite in eine durchsichtige Hornhaut (Cornea) übergeht. Wie viel Licht ins Augenninnere fällt, wird durch die Regenbogenhaut (Iris) reguliert. Sie ist ein ringförmiger Muskel, in dessen Mitte sich eine kreisförmige Öffnung (Pupille) befindet. Die Pupillengröße kann durch Kontraktion und Dilatation der Iris verändert werden. Allerdings kann diese Reaktion nicht willentlich, sondern nur durch einen Lichtreflex gesteuert werden. Fällt viel Licht auf das Auge, verengt sich die Pupille. Fällt wenig Licht darauf, vergrößert sich die Öffnung. Da der Durchmesser zwischen zwei und acht Millimetern variieren kann, ergibt sich eine breite Anpassungsfähigkeit an veränderte Beleuchtungsbedingungen. Kompromisse müssen gefunden werden zwischen der Sensitivität (Fähigkeit, auch schwache Reize zu erkennen) einerseits und der Sehschärfe (Auflösungsvermögen für kleine Details) sowie der Schärfentiefe (Entfernungsbereich, der gleichzeitig scharf abgebildet werden kann) andererseits. Es gilt: Je kleiner die Öffnung, umso schärfer das Bild. Bei guten Beleuchtungsverhältnissen ist dies kein Problem, weil auch bei einer Verengung der Pupille genügend Licht ins Auge fällt. Bei geringer Beleuchtungsintensität öffnet sich jedoch die Pupille, wodurch zwar mehr Licht ins Auge fällt, aber die Sehschärfe und die Schärfentiefe abnehmen.

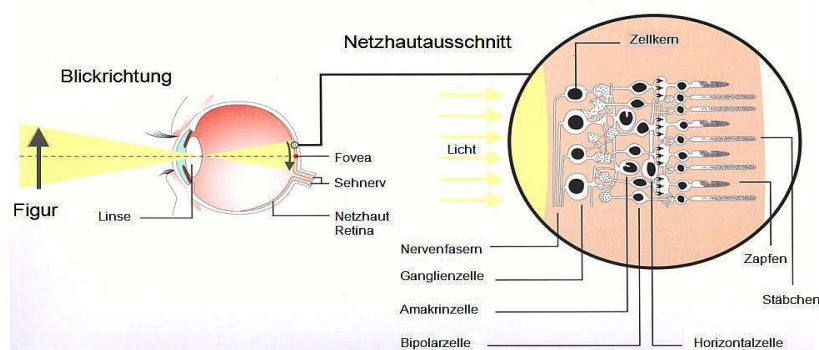


Abbildung 12: Lichteinfall und retinale Reizverarbeitung (Goldstein, 2008). Erläuterung im Text.

Hinter der Pupille befindet sich die Linse. Ihre Funktion ist es, das eintretende Licht so zu bündeln, dass das Licht, das von einem bestimmten

Punkt aus reflektiert wird, auch auf einen einzelnen Punkt projiziert wird. Das optische Bild, das schließlich auf der Netzhaut entsteht, ist invertiert und spiegelverkehrt. Die Linse ist durch Veränderung ihrer Form (gewölbter oder flacher) in der Lage, ihre Refraktion (Brechkraft) zu verändern und Strahlen von unterschiedlich weit entfernten Gegenständen so zu bündeln, dass sie auf der hinteren Wand des Auges scharf abgebildet werden. Die Anpassung des Strahlengangs an die Entfernung des scharf abzubildenden Objektes wird als *Akkommodation* bezeichnet. Sie wird erreicht durch die Kontraktion und Dilatation der seitlich ansetzenden Ciliarmuskeln und einer daraus resultierenden Veränderung der Zugkraft der an der Linse befestigten Bänder. Bei einer Fokussierung auf nahe Gegenstände werden die Ciliarmuskeln kontrahiert. Die Zugkraft der Bänder wird verringert und durch ihre natürliche Elastizität wird die Linse dicker. Die Brechkraft der Linse wird erhöht. Entsprechend ist bei entspannten Ciliarmuskeln der optische Apparat des Auges auf größere Entfernungen eingestellt. Wenn ein Objekt sich zu nahe am Auge befindet, stößt die Akkomodationsfähigkeit der Linse an ihre Grenzen. Der Nahpunkt ist altersabhängig und beträgt beim gesunden Auge eines Zehnjährigen unter 10 Zentimeter, bei über Sechzigjährigen meistens mehr als 100 Zentimeter. Ohne den hier beschriebenen *dioptrischen Apparat* (zusammenfassende Bezeichnung für alle lichtbrechenden Elemente) des Auges könnten wir zwar Helligkeiten und Farben wahrnehmen, aber keine Konturen.

Die Größe der Abbildung eines Objektes auf der Retina wird durch den *Sehwinkel* angegeben. Er wird gebildet durch die geradlinige Verbindung der Eckpunkte des Objektes mit dem Zentralpunkt des Auges. Bei gleichbleibender Größe, aber zunehmender Entfernung des Objektes wird der Sehwinkel kleiner (Minifikation). Umgekehrt wird bei einer Annäherung des Objektes an die Beobachterin oder den Beobachter ein immer größerer Teil des Gesichtsfeldes ausgefüllt (Magnifikation). Bei einer Projektion auf eine zweidimensionale Fläche lässt sich in Bezug auf ein einzelnes Bild nicht entscheiden, ob sich die Größe des Objektes

verändert hat, oder ob es sich um dasselbe Objekt in einer anderen Entfernung handelt. Mechanismen, die diese Ambiguität zu reduzieren versuchen, werden später untersucht. Wichtig für das Verständnis dieses Problems ist, dass der Sehwinkel nicht die räumlichen Dimensionen des Objektes (distaler Stimulus) beschreibt, sondern die Größe des optischen Bildes auf der Retina (proximaler Stimulus).

Die Veränderungen der optischen Anordnung (durch Eigen- oder Objektbewegungen) müssen so abgebildet werden, dass die räumlich und zeitlich invarianten Komponenten erhalten bleiben, denn aus ihnen wird über die Zirkularität des Sehvorgangs die Stabilität der visuellen Welt gewonnen. Hier setzen viele Heuristiken des visuellen Systems an. In der uns vertrauten euklidischen Geometrie werden Körper als gleich angesehen, wenn sie metrisch äquivalent sind. Zwei Quadrate sind dann gleich, wenn sie gleich groß sind. Nur eine bestimmte Klasse von Transformationen (starre Bewegungen) lassen die Objekte unverändert. Hierzu gehören Schiebungen, Drehungen und Spiegelungen. Die entsprechenden Projektionen werden als Parallelprojektionen bezeichnet. Sie kommen in der Natur selten vor und sind wenig geeignet, um die Veränderungen auf der Netzhaut einer sich bewegenden Beobachterin oder eines sich bewegenden Beobachters zu beschreiben. Eine angemessenere Methode sind projektive oder darstellende Geometrien wie z. B. die Linearperspektive. Sie beinhalten Axiome, die es erlauben, die Maße von Längen und Winkeln sowie die Proportionen von Strecken bei einem Wechsel des Standortes unverändert zu lassen. Allerdings können auch projektive Geometrien nicht die tatsächlichen Verhältnisse von Reflexion, Absorption und Brechung adäquat darstellen. Ein Beispiel: Gibt es eine Reihe von Objekten, die sich von einem Standort aus gegenseitig verdecken, wird nur das Licht, das von der nächsten Oberfläche reflektiert wird, ins Auge gelangen. Die Photonen der übrigen Oberflächen werden entweder von anderen Oberflächen absorbiert oder reflektiert. Solche komplexen Interaktionsverhältnisse kann auch eine darstellende Geometrie nicht wiedergeben.

In der Geschichte der Wahrnehmungsforschung ist wiederholt diskutiert worden, ob die Zentralperspektive ein universelles Modell für das menschliche Sehen sein könnte. Wie wir heute wissen, beschreibt dieses Modell zwar eine fotografische Abbildung adäquat, nicht aber das Sehen. Zunächst ist eine zentralperspektivische Darstellung nur für einen Spezialfall wirklich korrekt: monokulare Beobachtung und fester Augenpunkt. In der Umgebungswahrnehmung kommt eine solche Situation praktisch nicht vor. Vor allem ist sie als Modell für die weiteren Verarbeitungsschritte irreführend. Die phänomenale Welt ist nicht lediglich eine projektive Abbildung der Außenwelt, selbst wenn sie so erscheint. Vor allem Kontrastbildung, Gruppierung und Ergänzung spielen eine wichtige Rolle, um die im retinalen Bild enthaltenen Ambiguitäten schnell und effektiv zu reduzieren.

3. Transduktion und Transformation

Die Retina entsteht in der Embryonalentwicklung aus einer Ausstülpung des Zwischenhirnbodens. Sie ist ein Stück des Gehirns. Bestehend aus mehr als 120 Millionen lichtempfindlichen Photorezeptoren bildet sie die Projektionsfläche des Auges. Als dünne Schicht bedeckt sie den hinteren Teil der Augenhaut. Eine Ausnahme bildet lediglich die Stelle, an der der Sehnerv das Auge verlässt: Hier finden sich keine Rezeptoren. Im Gesichtsfeld resultiert ein blinder Fleck. In der Retina wird die Lichtenergie in Nervenimpulse umgewandelt, die als Information über die visuelle Umgebung durch den Sehnerv weitergeleitet wird.

Die Photonen gehen ungehindert durch die ersten beiden (durchsichtigen) Schichten der Retina und treffen in der hintersten Schicht auf Photorezeptoren (vgl. Abbildung 12). Ihre Sehpigmente können Licht absorbieren und zerfallen bei dieser Absorption in chemische Substanzen. Es kommt zu einer Veränderung des elektrischen Potentials der Zelle. Der Prozess der Umwandlung des äußeren Reizes in ein elektrophysiologisches

Signal wird als *Transduktionsprozess* bezeichnet. Er ist für alle Sinnessysteme grundlegend, nicht nur für die visuelle Wahrnehmung.

Die graduellen Signale (Rezeptorpotentiale) werden an ein Netzwerk aus verschiedenen Zellen (horizontale, bipolare und amakrine Zellen) in der zweiten Schicht weitergeleitet. Diese Zellen sind zu einem großen Teil auch untereinander (horizontal) verschaltet. Dadurch kann eine aktivierte Zelle nicht nur nachfolgende Zellen erregen, sondern auch benachbarte in ihrer elektrischen Aktivität hemmen (*laterale Hemmung*). Schließlich werden die Ganglienzellen (oberste Schicht) aktiviert. Ihr besonderes Merkmal für die Reizfortleitung ist, dass sie das graduelle Signal über eine Summation in ein Aktionspotential übersetzen. Es geht damit in ein „Alles-oder-Nichts-Gesetz“ über. Dieser Prozess wird als *Transformation* bezeichnet. Der Mechanismus der lateralen Hemmung führt dazu, dass die Signale von Reizsprüngen verstärkt und kontinuierliche Intensitätsverläufe verwischt werden. Es erfolgt eine Übersetzung mit Akzentuierung. Sie ist wichtig, weil eine statistische Analyse von Bildern natürlicher Szenen zeigt, dass die visuelle Umwelt zum überwiegenden Teil aus kontinuierlichen Übergängen besteht. Ohne eine laterale Hemmung würden benachbarte Ganglienzellen fast immer die gleichen Signale weitergeben. Das wäre höchst ineffizient und würde bei geringen Veränderungen der Beleuchtungs- und Umgebungsbedingungen leicht zu Fehlinterpretationen führen. Die laterale Hemmung ist ein wichtiges Beispiel für das Prinzip der Kontrastbildung.

Bisher wurde mithilfe der Prozesse Transduktion und Transformation erklärt, wie eine bestimmte Lichtverteilung in neuronale Impulse umgewandelt wird. Auf dieser Basis lässt sich die Fähigkeit des visuellen Systems, die Lichtverteilung in Mustern zu strukturieren und zwischen einzelnen Elementen des visuellen Feldes zu unterscheiden, noch nicht verstehen. Notwendig ist eine Analyse der neuronalen Verschaltung. Insgesamt weist die Vernetzung in der Retina eine starke *Konvergenz* auf: Den etwa 120 Millionen Rezeptorzellen stehen nur etwa 1 Million

Ganglienzellen gegenüber. Die Zuordnung von Photorezeptoren zu einer Ganglienzelle ist fest. Da gleichzeitig eine feste Beziehung zwischen der Position eines Lichtreizes im Gesichtsfeld und dem erregten Photorezeptor besteht, spricht man von einem *rezeptiven Feld*.

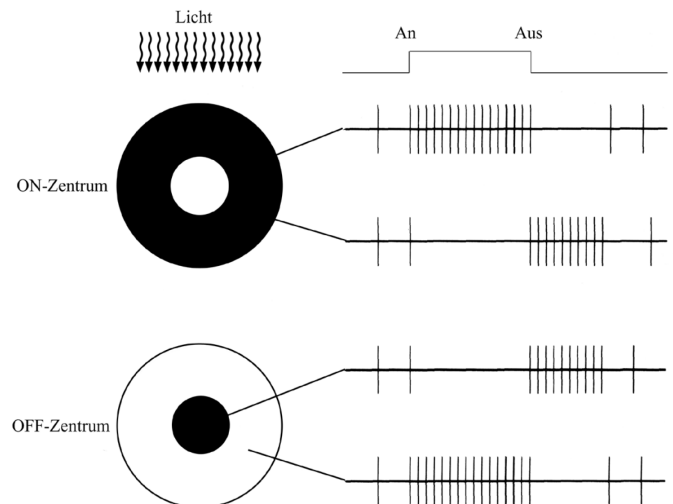


Abbildung 13: Funktionsweise von rezeptiven Feldern. Die Striche im rechten Teil der Abbildung symbolisieren Aktionspotentiale und damit die Aktivität der Zellen. Weitere Erläuterungen im Text.

Jedes dieser kreisförmigen und sich wechselseitig überlappenden Felder besitzt ein Zentrum und eine Peripherie, die im Allgemeinen antagonistisch reagieren (Abbildung 13). Bei einem *ON-Zentrum-Feld* führt eine Lichtreizung im Zentrum des Feldes zu einer Aktivierung der Ganglienzelle (exzitatorische Reaktion), während eine Stimulation im peripheren Bereich eine Hemmung bewirkt (inhibitorische Reaktion). Im Gegensatz hierzu führt bei einem *OFF-Zentrum-Feld* eine zentrale Reizung zu einer Hemmung und eine periphere Reizung zu einer Aktivierung des Neurons. Die resultierende Aktivität einer Ganglienzelle hängt davon ab, wo und in welchem Umfang Photonen auf ihr rezeptives Feld treffen. Bei einer gleichmäßigen Reizung des gesamten rezeptiven Feldes heben sich nachfolgende Aktivierung und Hemmung auf, die Aktivität bleibt unverändert. Je asymmetrischer die Verteilung zugunsten der exzitatorischen Komponente ist, umso wahrscheinlicher wird eine

Weitergabe des Signals durch das Neuron. Die Kontrastbildung ist damit nicht auf die Netzhaut beschränkt. Sie findet sich auch auf den höheren Stufen der Verarbeitung. Dabei werden die rezeptiven Felder von Stufe zu Stufe größer. Gleichzeitig wächst die Spezialisierung der Neurone und die Abstraktheit der Repräsentation nimmt zu. Dies ist die Voraussetzung dafür, dass die Inhomogenitäten je nach Verarbeitungsstufe immer wieder neu gruppiert und zusammengefasst werden können. Bereits auf Ebene der Retina werden die Grundsteine für viele spätere Verarbeitungsschritte gelegt. Bereits hier beginnen erste konstruktive Prozesse.

4. Skotopisches und photopisches Sehen

Das visuelle System muss in unterschiedlichen Beleuchtungssituationen zuverlässige Informationen liefern. Die Beleuchtung bei Sonnenlicht ist etwa 800.000-mal stärker als in einer Vollmondnacht. Die Unterschiede im Umgebungslicht, das von den Oberflächen reflektiert wird, sind bei einer Szene, die von Tageslicht beleuchtet wird, etwa um den Faktor 1000 größer. Es erreicht diese Anpassung zum einen über die Steuerung der Pupillenweite, zum anderen über unterschiedliche Typen von Photorezeptoren (Stäbchen und Zapfen, benannt nach ihrer Form) sowie über eine zeitweise getrennte Verarbeitung dieser Informationen (skotopisches und photopisches System). Stäbchen sind sehr lichtempfindlich und weisen dieselben spektralen Absorptionseigenschaften auf. Sie können nicht zwischen den Wellenlängenunterschieden diskriminieren. Farbunterscheidungen sind mit diesem System nicht möglich. In Experimenten konnte gezeigt werden, dass unter bestimmten Voraussetzungen bereits ein bis drei Photonen ausreichen, um eine Veränderung der Rezeptoraktivität zu bewirken. Aufgrund dieser hohen Lichtempfindlichkeit ermöglicht das skotopische System ein Sehen bei Dämmerung und bei Dunkelheit. Unter Tageslichtbedingungen sind die Stäbchen gesättigt, das System ist weitgehend inaktiv. Geeignet für das Tagessehen sind die weniger

lichtempfindlichen Zapfen. Die Komplementarität beider Systeme ermöglicht die Anpassung an extreme Beleuchtungsbedingungen. Der Vorgang wird als Hell- bzw. Dunkeladaptation bezeichnet. Beide lassen sich bei einem Wechsel von hellem Sonnenlicht in einen stark abgedunkelten Raum und umgekehrt gut beobachten. Die vollständige Dunkeladaptation (maximale Empfindlichkeit der Stäbchen) dauert etwa 20 Minuten.

In den Systemen ist die Konvergenz unterschiedlich. Während bei den Stäbchen zum Teil mehrere hundert Einzelrezeptoren auf eine Ganglienzelle konvergieren, sind es im photopischen System oft nur einige wenige Zapfen oder eine Eins-zu-eins-Zuordnung. Die hohe Konvergenz der Stäbchen führt bei einem schwachen Lichtreiz, der viele Stäbchen stimuliert, zu einer Aufsummierung der Wirkung. Auch ein schwacher Reiz ist in der Lage, bei der Ganglienzelle ein Aktionspotential auszulösen und die Alles-oder-Nichts-Schranke zu überwinden. Allerdings hat das Gehirn dann keine Möglichkeit, das Eingangssignal präzise zu orten. Das führt zur geringeren Detailgenauigkeit des skotopischen Sehens. Bei einer von Zapfen innervierten Ganglienzelle ist eine höhere Lichtintensität notwendig, um die Impulsrate der Ganglienzelle zu verändern. Aufgrund der geringeren Konvergenz lässt sich der Ort der Reizung genauer lokalisieren.

Die drei Zapfenarten werden nach dem Spektralbereich ihrer höchsten Empfindlichkeit häufig als Rotzapfen (langwellenlängen-sensitiv), Grünzapfen (mittelwellenlängen-sensitiv) und Blauzapfen (kurzwellenlängen-sensitiv) bezeichnet. Ein Vergleich der Absorptionsspektren zeigt Abbildung 14: 1. Alle drei Zapfenarten absorbieren über einen großen Wellenlängenbereich hinweg Licht. 2. Die Kurzbezeichnungen (Rot, Grün, Blau) sind irreführend, weil das Maximum für die Rotzapfen in einem Wellenlängenbereich liegt, der eher als Gelb wahrgenommen wird und das Maximum bei den Blauzapfen als Violett gesehen wird. 3. Die Absorptionskurven der Rot- und Grünzapfen

verlaufen sehr ähnlich. Wie man heute weiß, sind die drei Zapfenarten entwicklungsgeschichtlich erst vor kurzer Zeit aus einem gemeinsamen Urzapfen entstanden. Ihre Absorption ist abhängig vom Sehpigment. Eine genetische Analyse der Proteine dieser Pigmente hat gezeigt, dass die Aminosäuresequenzen für das Rot- und das Grünpigment sich in weniger als zwei Prozent unterscheiden.

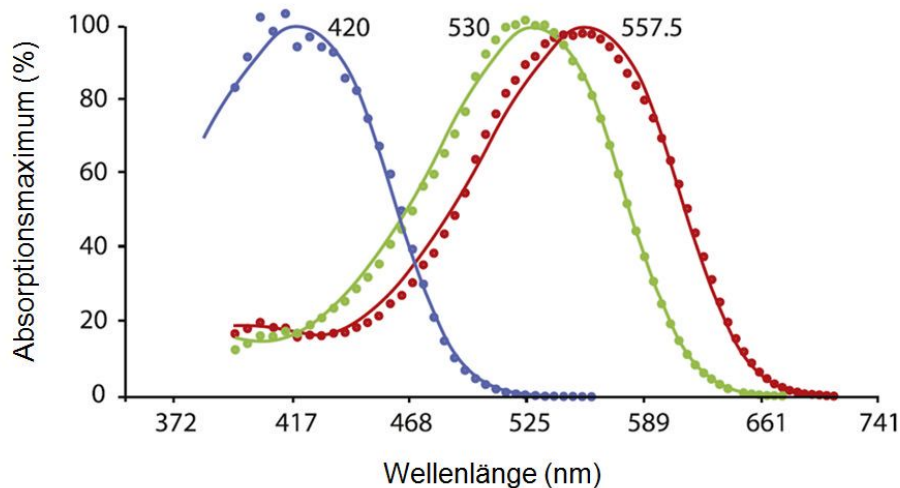


Abbildung 14: Absorptionsspektren. Die Absorptionskurven für Rot und Grün verlaufen ähnlich.

Die Signale der Zapfen werden in den drei sogenannten Gegenfarbenkanälen weiterverarbeitet. Im Helligkeitskanal wird die Summe der Signale aus den Rot- und Grünzapfen (R+G) gebildet, im Rot-Grün-Kanal die Differenz der beiden (R-G) und im Blau-Gelb-Kanal die Differenz aus dem Signal der Blauzapfen und der Summe der Rot- und Grünzapfen (B-(R+G)). Damit lassen sich alle Farben des Spektrums darstellen. Gleichzeitig findet eine weitere Kontrastbildung statt.

Die Verteilung der Stäbchen und Zapfen auf der Netzhaut ist unterschiedlich. Im Bereich der Fovea centralis (dem Ort des schärfsten Sehens) finden sich nur Zapfen. Ihre Anzahl nimmt zur Peripherie hin deutlich ab. Die Dichte der Stäbchen ist bei einem Abstand von etwa 20°

zur Fovea am größten und wird zur Peripherie ebenfalls geringer. Die Zapfen sind wie ein unregelmäßiges Mosaik angeordnet. In der Fovea befinden sich wesentlich mehr langwellige Zapfen. Die Rezeptordichte definiert das Auflösungsvermögen und damit die Sehschärfe. Sie nimmt insgesamt zur Peripherie hin stark ab und weist dort auch eine wesentlich größere Konvergenz auf. Oft sind mehrere hundert Rezeptorzellen mit einer Ganglienzelle verschaltet. Die Auffassung, dass im peripheren Gesichtsfeld keine Farben gesehen werden können, hat sich als nicht richtig erwiesen. Allerdings ist hier die räumliche Auflösung beim Farbsehensystem geringer als beim Helligkeitssystem. Sind Objekte hinreichend groß, kann auch deren Farbe erkannt werden.

Unsere Augen sind in ständiger Bewegung und tasten das Sehfeld permanent ab, auch wenn es uns nicht bewusst ist. Einen festen Augenpunkt gibt es nur theoretisch. Das Ziel der Augenbewegungen (Okulomotorik) ist es, die jeweils wichtigen Informationen scharf und detailgenau abzubilden. Die Augen werden so ausgerichtet, dass die Projektion der zugehörigen Punkte in den Bereich der Fovea centralis fällt. Innerhalb einer Sekunde finden etwa drei Fixationen an unterschiedliche Stellen des Blickfeldes statt. Verbunden werden sie durch sehr schnelle Augenbewegungen (Sakkaden). Die Bilder mehrerer Fixationen werden zu einem weitwinkligen, farbigen und scharfen Bild integriert (Prinzip der Ergänzung). Diese Integration führt dazu, dass die Welt nicht jedes Mal verschwindet, wenn etwa 15- bis 30-mal pro Minute ein Lidschlussreflex erfolgt. Die Komplementarität von fovealem und peripherem Sehen ermöglicht die Registrierung von Reizen in einem weiten Gesichtsfeld (horizontal ca. 180°) und die präzise Abbildung von Details; zwei Anforderungen, die in optischen Systemen nur sehr schwer gleichzeitig zu optimieren sind.

Fast alles, was bisher über die Arbeitsweise der Augen gesagt wurde, bezog sich auf das einzelne Auge (monokulares Sehen). Es stellt sich die Frage nach der Funktion des beidäugigen oder binokularen Sehens. Neben der

Vergrößerung des Gesichtsfeldes dient es vor allem der Räumlichkeit der Wahrnehmung. Erst beim binokularen Sehen entsteht die volle Intensität des Tiefeneindrucks. Die *Konvergenz* der beiden Augen stellt zunächst sicher, dass bei einer Fixation eines Gegenstandes die Abbildung auf *korrespondierende Netzhautstellen* beider Augen erfolgt, sodass eine eindeutige Zuordnung möglich wird. Der Grund für die besondere Qualität des binokularen Sehens liegt darin, dass die meisten Gegenstände auf den beiden Netzhäuten seitlich verschoben abgebildet werden. Durch den unterschiedlichen Winkel der beiden Augen zum Objekt (der durchschnittliche Augenabstand bei einem Erwachsenen ist 6,4 Zentimeter) entstehen zwei disparate Bilder. Das Ausmaß dieser *Querdisparation* ist abhängig von der Entfernung fixierter Objekte und damit ein wichtiger Tiefenhinweis, der zur Ambiguitätsreduktion einer Szene beiträgt.

5. Retinotopie Organisation

Einen Überblick, wie die Informationen von den beiden Netzhäuten weitergeleitet werden, gibt Abbildung 15. Alle Neurone einer Retina werden im Sehnerv (Nervus opticus) zusammengefasst. Er besteht aus etwa einer Million Nervenfasern und zieht vom Auge zum Corpus geniculatum laterale (CGL) im Thalamus. Dort erfolgt eine Umschaltung und Verarbeitung, bevor eine Weiterleitung als Sehstrahlung (Radiatio optica) zum primären visuellen Cortex im Hinterhauptslappen (Occipitallappen) erfolgt. Auf dem Weg zum CGL kreuzt etwa die Hälfte der Nervenfasern im Chiasma opticum. Die Informationen aus den innenliegenden (nasalen) Netzhautbereichen und damit aus den äußeren Hälften des Gesichtsfeldes werden an die jeweils gegenüberliegende Hirnhälfte weitergeleitet (kontralateral), während die Neurone der außenliegenden (temporalen) Gebiete ungekreuzt weiterlaufen (ipsilateral). Das CGL ist die wichtigste Schaltstation zwischen Retina und primärem visuellen Cortex. Es besteht aus sechs übereinander liegenden

Schichten. Die Schichten eins und zwei enthalten (große) Magno-Neurone (M-Neurone), die Schichten vier bis sechs (kleinere) Parvo-Neurone (P-Neurone). Die Neurone aus der rechten Retinahälfte jedes Auges verlaufen zum rechten CGL und liefern eine vollständige Repräsentation der linken Gesichtshälfte. Entsprechend erreichen die Fasern aus den linken Retinahälften den linken Teil des CGL. Die Neurone der kontralateralen nasalen Hälfte enden in den Schichten eins, vier und sechs, die der ipsilateralen temporalen Retinahälfte in den Schichten zwei, drei und fünf. Die Schichten sind so angeordnet, dass die räumliche Organisation im CGL der räumlichen Organisation auf der Retina entspricht (Retinotopie) und folgen damit dem Prinzip der Kartierung. Die Informationen beider Augen werden auf dieser Stufe getrennt verarbeitet und erst im primären visuellen Cortex zusammengeschaltet.

Die überragende Bedeutung der visuellen Informationsverarbeitung im Vergleich zu anderen Sinnessystemen lässt sich an der Anzahl der beteiligten neuronalen Zentren und deren Gehirnmasse ablesen. Mehr als die Hälfte des Cortex und mehr als 30 verschiedene Areale sind an der Wahrnehmung und der Reaktion auf visuelle Reize beteiligt. Der primäre visuelle Cortex (V1) nimmt etwa 15 Prozent der Großhirnrinde ein. Diesen Nervenzellen ist gemeinsam, dass sie ihr Erregungsmaximum nicht bei punktförmigen Lichtreizen zeigen (wie die Neurone der Retina oder des CGL), sondern bei kurzen Balken oder Lichtstreifen. Diese Eigenschaft ist wichtig für Textur- und Mustererkennung.

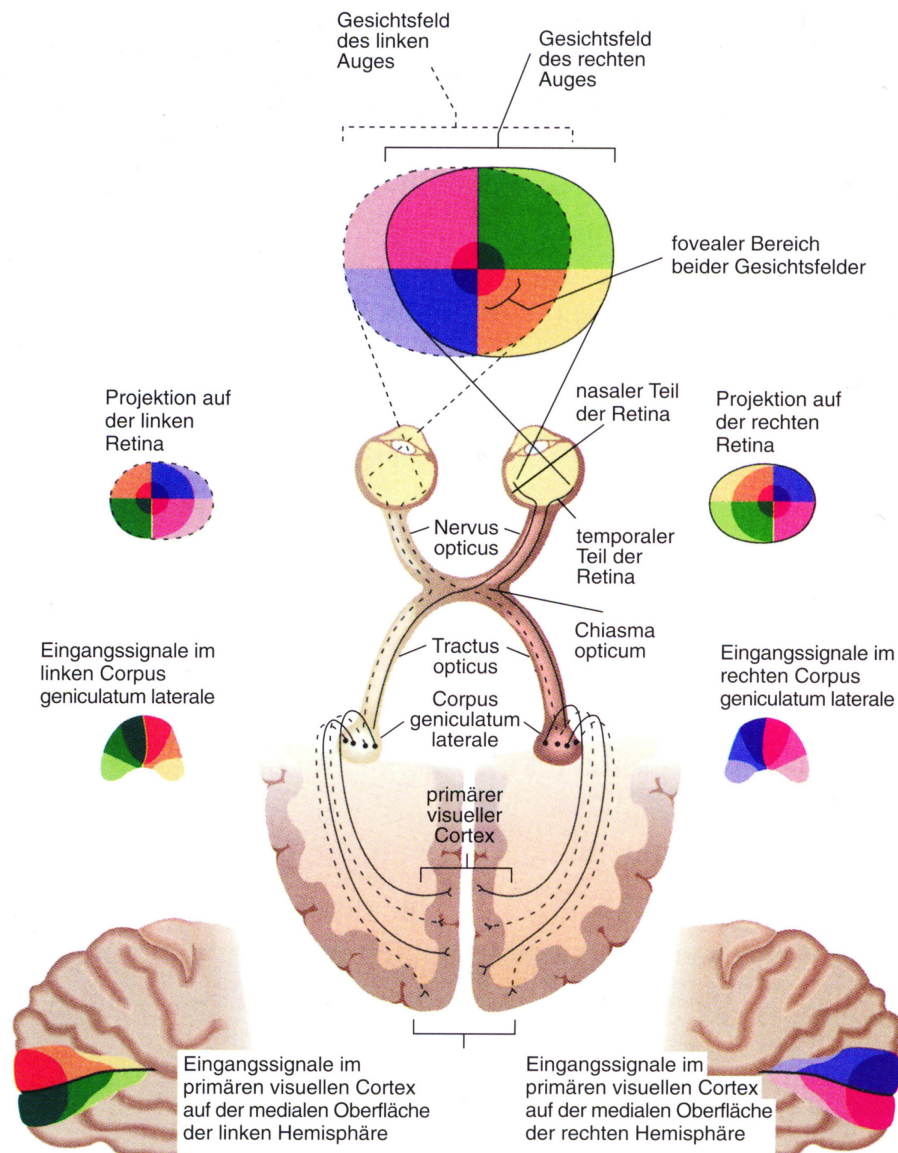


Abbildung 15: Verlauf der Sehbahn. Überblick mit den beteiligten Strukturen und eine Zuordnung zum Gesichtsfeld der beiden Augen (Pinel, 2001).

Drei Neuronenarten oder Merkmalsdetektoren sind zu unterscheiden: 1. Einfache cortikale Zellen haben längliche rezeptive Felder mit einer Unterteilung in eine erregende und eine hemmende Zone. Sie reagieren am stärksten auf Streifen (Linien) einer bestimmten Orientierung. 2. Komplexe cortikale Zellen reagieren optimal auf die Bewegung eines richtig orientierten Streifens oder Winkels (Ecken) innerhalb ihres

rezeptiven Feldes. 3. Hyperkomplexe oder endinhibierte Zellen antworten auf Streifen oder Winkel einer bestimmten Länge, wenn sie sich gleichzeitig in einer festgelegten Richtung über das rezeptive Feld bewegen. Es wird deutlich, dass die Zellen in V1 zu deutlich globaleren Prozessen der Gruppierung und Kontrastbildung fähig sind als die Zellen der Retina.

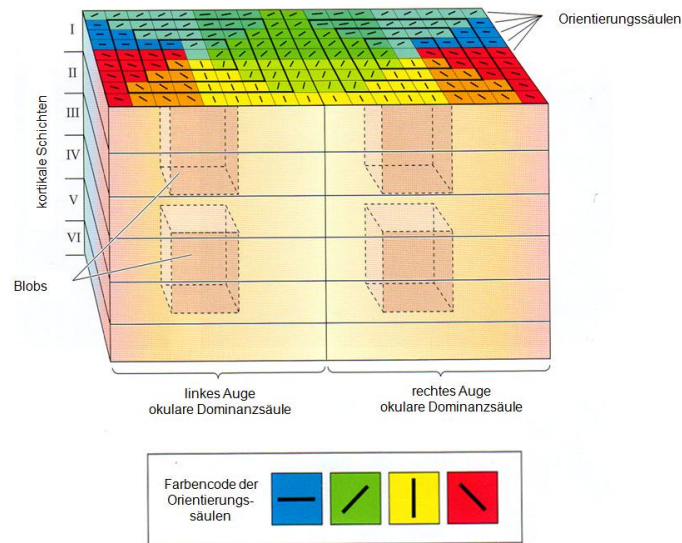


Abbildung 16: Schichtenaufbau des primären visuellen Cortex (V1) am Beispiel einer Hypersäule (Wolfe, Kluender und Levi, 2009).

Der primäre visuelle Cortex ist streng geordnet (Abbildung 16). Es können mit Karten und Säulen zwei Ordnungsprinzipien unterschieden werden. Wie auch das CGL ist V1 retinotop organisiert: Jedem Punkt auf der Retina entspricht genau ein Ort im visuellen Cortex. Die Repräsentation der Fovea nimmt einen überproportionalen Raum ein (cortikaler Vergrößerungsfaktor). Anatomisch gesehen sind die Neurone in regelmäßigen Kolumnen oder Säulen angeordnet, die senkrecht zur Oberfläche des Cortex stehen. Drei Arten von Säulen lassen sich als Zusammenfassung von Neuronen gleicher Spezialisierung unterscheiden: Orientierungssäulen, Augendominanzsäulen und Hypersäulen. Während die beiden ersten Säulenarten Aufgaben gemäß ihrer Bezeichnung erfüllen, können die Hypersäulen als elementare Verarbeitungsmodule

aufgefasst werden. Sie stellen alle Informationen zur Verfügung, die für die Bestimmung der Länge und der Orientierung einer Linie oder Kante notwendig sind. Eine Hypersäule besteht aus zwei Augendominanzsäulen, die jeweils primär mit Informationen aus dem rechten bzw. linken Auge versorgt werden.

6. Modularität und Mehrdimensionalität

Zu den höheren Verarbeitungsstufen hat es vor allem in den letzten 20 Jahren eine Fülle interessanter und aufschlussreicher Experimente gegeben. Leider haben sie bisher nicht zu einem übergreifenden Modell geführt. Ein einflussreicher Vorschlag, in den sich viele Einzelbefunde gut einfügen lassen, beschreibt die Verarbeitung der Informationen in einzelnen Modulen und die Weiterleitung eines großen Teils der Informationen in zwei getrennten Bahnen (Abbildung 17): ein *dorsaler Pfad*, der über das Mediotemporale Areal (MT) zum Scheitellappen (Parietaler Cortex) führt und ein *ventraler Pfad*, der über V4 und IT (Inferotemporaler Cortex) zum temporalen Cortex führt. Die dorsale Bahn wird auch als *Magno-System* bezeichnet, weil sie in den magnozellulären Schichten des CGL beginnt. Die ventrale Bahn, die in den parvozellulären Schichten ihren Ausgangspunkt hat, wird auch *Parvo-System* genannt. Eine Veranschaulichung der Aufgabenteilung beider Pfade ist das Ergreifen eines Gegenstandes mit der Hand. Beide Bahnen sind beteiligt. Im Magno-System werden vor allem Informationen übermittelt, die für die Lokalisation des Objektes und die Steuerung der eigenen Bewegung zuständig sind. Ausgangspunkt sind die Helligkeitskontraste. Die Auflösung ist relativ gering. Die Reaktionen erfolgen gewöhnlich sehr schnell und dauern weniger lange an. Das ist vor allem für die Analyse von Bewegungen bedeutsam. Das Parvo-System ermöglicht mittels Muster- und Formanalysen die Objektidentifikation. Es hat eine zwei- bis dreimal höhere Auflösung. Die Reaktionen sind langsamer und dauern länger an. Die getrennt arbeitenden Systeme

ergänzen sich ausgezeichnet, um den breiten Anforderungskatalog einer Wahrnehmung unter sich verändernden Umgebungsbedingungen zu erfüllen.

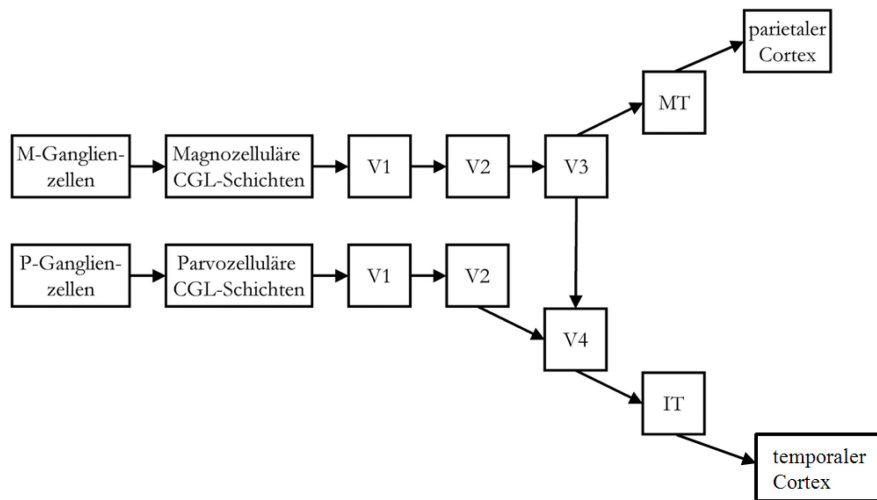


Abbildung 17: Schematische Darstellung des Parvo-Systems und des Magno-Systems (Goldstein, 2008). Erläuterung im Text.

Die Annahme, dass auch den höheren Zentren unterschiedliche Aufgaben bei der visuellen Analyse zukommen, führte zur Suche nach grundlegenden Modulen. Wo wird primär Form verarbeitet, wo Farbe, wo Bewegung? Doch ist zu bedenken, dass fast alle Zentren Informationen von anderen Teilen des Gehirns erhalten und dass ihr Einfluss umso größer wird, je höher die Verarbeitungsstufe ist. Neuere physiologische und psychophysische Untersuchungen bestätigen, dass Zentren, in denen eine sehr spezifische Sensitivität vermutet wurde (z. B. V4 für Farbe, zum Parvo-System gehörend), auch eine Sensitivität für andere Merkmale aufweisen. Daher ist gegenwärtig eine Sichtweise angemessen, die neben einer *Modularität* auch eine *Mehrdimensionalität* berücksichtigt (vgl. Gegenfurtner, 2001; Gegenfurtner und Kiper, 2003). Die verschiedenen Reizmerkmale (z. B. Orientierung, Farbe, Form, Bewegung) werden nicht ausschließlich in weitgehend getrennten Kanälen verarbeitet, sondern parallel einer mehrdimensionalen Analyse unterzogen. Sie müssen später

nicht wieder neu zusammengesetzt werden. Diese Annahme widerspricht nicht der Vermutung, dass es in höheren Zentren eine starke Spezialisierung einzelner Neuronenpopulationen für komplexe Reizmerkmale geben kann. Ein Beispiel hierfür ist die Gesichtererkennung. Warum glauben wir, auch in Vorlagen, die nur aus einer Ansammlung von Flecken bestehen, noch ein Gesicht zu sehen? Eine Erklärung könnte sein, dass wir auf einer späten Verarbeitungsstufe über sehr spezifische Neurone verfügen, die in der Lage sind, einen großen Teil des visuellen Feldes unter einer Interpretation zusammenzufassen und ihm eine einheitliche Bedeutung zu geben. Je weiter die Verarbeitung im Gehirn voranschreitet, desto größer werden die rezeptiven Felder einzelner Neurone.

Fasst man die Gesetzmäßigkeiten der visuellen Reizverarbeitung zusammen, zeigt sich ein System, das unter schwierigsten Bedingungen ein adäquates Bild der Wirklichkeit erzeugt. Es wird nicht über ein mechanisches Abbild gewonnen, sondern durch eine Analyse der invarianten Reizmerkmale in Kombination mit einer gedächtnisgesteuerten Interpretation und einer engen Verbindung zur Motorik.

Bei der Darstellung der physiologischen Grundlagen wurde eine sequentielle Strategie verfolgt. Beginnend mit den ersten Analyseprozessen in der Retina wurde die Reizverarbeitung bis zu den kortikalen Zentren aufgezeigt. Diese Richtung der Informationsverarbeitung wird in der Forschung als „aufwärtsgerichtet“, „bottom-up“, „datengesteuert“ oder „induktiv“ bezeichnet. Sie wird ergänzt durch „abwärtsgerichtete“, „top-down“, „hypothesengesteuerte“ bzw. „deduktive“ Prozesse. Sie bezeichnen den Einfluss höherer Verarbeitungsstufen auf niedrigere und zeigen damit die Bedeutung von Lernen und Gedächtnis. Viele Forscher sind der Auffassung, dass beide Prozesse zum Teil parallel ablaufen und dass die hypothesengesteuerten Prozesse auf die datengesteuerte Stimulusanalyse einwirken können

(Bidirektionalität). Unklar ist, ob dieser Einfluss auch bei sehr frühen Verarbeitungsstufen nachzuweisen ist (z. B. bei Helligkeitsanalysen). Viele Ergebnisse der letzten Jahre sprechen für eine neuronale Plastizität und perzeptuelles Lernen. Nur der kleinere Teil der im Rahmen der visuellen Informationsverarbeitung verwendeten Signale stammt direkt aus der retinalen Stimulation. Der überwiegende Teil kommt aus verschiedenen Zentren des Gehirns. Je höher die Verarbeitungsstufe, umso größer wird der Einfluss dieser internen Stimulation.

Die „aufwärtsgerichtete“ Darstellung wurde aus didaktischen Gründen gewählt. Sie wird auch bei der Beschreibung des Aufbaus der visuellen Welt durch Heuristiken beibehalten. Den tatsächlichen Abläufen wird sie nicht gerecht. Sie unterschätzt die Zirkularität und betont die analytischen Aspekte des Sehvorgangs gegenüber den konstruktiven. Zusammenfassend die wichtigsten Erkenntnisse:

1. Das visuelle System hat eine hohe Sensitivität und ist effektiv in der Anpassung an unterschiedliche Beleuchtungsverhältnisse.
2. Innerhalb gewisser Grenzen kann es mithilfe von Blickbewegungen, Akkommodation und Konvergenz jederzeit sicherstellen, dass ein Objekt im Gesichtsfeld scharf und detailgetreu abgebildet wird, auch wenn dieses sich bewegt.
3. Die retinotop Organisation schafft eine räumliche Ordnung, die von den Photorezeptoren der Retina bis zu den höheren Zentren beibehalten wird, auch wenn die Kartierung schrittweise gröber wird.
4. Über die neuronale Konvergenz – verbunden mit einem Wechselspiel aus Hemmung und Aktivierung – werden mithilfe der Kontrastbildung eine Kompensation von Detailverlusten und eine effiziente Auswahl der Information erreicht. Ergänzungen mit dem Ziel der Ambiguitätsreduktion sind nicht die Ausnahme, sondern die Regel.
5. Tendenziell nimmt die Spezialisierung und damit die Möglichkeit, komplexe Aspekte der Stimulation mit wenigen Neuronenverbänden zu analysieren, mit fortschreitender Verarbeitung zu. Dadurch kommt es immer wieder zu neuen Zusammenfassungen und Gruppierungen.
6. Auf höheren Verarbeitungsebenen nimmt der Einfluss zunächst unbeteiligter Gehirnstrukturen auf die aktuelle Reizverarbeitung zu.

Die Konsequenz ist, dass Aufmerksamkeit, Gedächtnis und schlussfolgernde Prozesse an Bedeutung gewinnen.

7. Die Stabilität der visuellen Welt wird letztlich über die Zirkularität und damit über die Sensomotorik erreicht. Können die Ambiguitäten des retinalen Bildes nicht aufgelöst werden, erfolgen neue Augen-, Kopf- oder Rumpfbewegungen.

Das Sehen lässt sich als ein Prozess beschreiben, bei dem, ausgehend von der Stimulation, die wahrscheinlichste und stabilste Lösung präferiert wird. Das Ziel ist, mit möglichst wenig Aufwand einen möglichst hohen Grad an Eindeutigkeit zu erreichen, damit das Perzept als Grundlage für Entscheidungen und Handlungen dienen kann. Bedenkt man, dass selbst bei komplexen Ausgangsdaten der gesamte Prozess nur zwischen 100 und 300 Millisekunden dauert, dann dürfte die visuelle Wahrnehmung des Menschen eines der effizientesten Systeme sein, das die Natur entwickelt hat. Die Unterschätzung des Sehens, etwa gegenüber dem Denken und Sprechen, resultiert vor allem daher, dass mit dem Sehen oft nur das Auge, nicht aber das Zusammenspiel von Auge, Gehirn und Motorik gemeint ist. Die produktiven Aspekte des Sehens werden zu wenig beachtet.

Kapitel III: Aufbau der visuellen Welt

Mit dem Wissen über die neurophysiologischen Prozesse als Basis stehen jetzt die Analyse- und Konstruktionsregeln im Blickpunkt. Die visuelle Welt ist, wie mehrfach betont, eine gefüllte Welt. Gegenstände bilden die grundlegenden Einheiten. Sie wechseln mit der Umgebung und sind jeweils charakteristisch. Eine moderne Großstadt, ein Wald oder auch ein Museum haben unterschiedliche, jeweils typische Wahrnehmungsgegenstände. Während phänomenal die Objekte im Mittelpunkt stehen, sind sie als Ausgangspunkt für die Analyse des visuellen Systems nahezu unbedeutend. Bei der Besprechung des Umgebungslichtes und der Reflexion wurde erläutert, dass lediglich die Oberflächen kritisch sind. Präziser: Es sind die Texturen der Oberflächen und die Struktur des von den Oberflächen reflektierten Lichtes, die die Wahrnehmung entscheidend beeinflussen. Aus Sicht der Wahrnehmungsforschung sind die Verteilungen der Oberflächen und ihrer Texturen typisch für eine Reizumgebung.

Damit Oberflächen identifiziert und verglichen werden können, sind Inhomogenitäten notwendig. Die Differenzen sind häufig minimal und fallen kaum auf. Der weiße Fensterrahmen könnte nicht von der weißen Wand unterschieden werden, wenn es keine materialspezifischen Reflexionsunterschiede gäbe. Eine wesentliche Leistung des visuellen Systems besteht in der sicheren Identifikation der Texturen bei unterschiedlichen Beleuchtungsbedingungen. Aus didaktischen Gründen wird auch hier der aufwärts gerichteten Informationsverarbeitung gefolgt: Helligkeits- und Kontrastwahrnehmung, das Entstehen von Texturen, die Figur- und die Objektwahrnehmung und schließlich die Wahrnehmung komplexer visueller Szenen. Doch stehen jetzt nicht die physiologischen Prozesse im Vordergrund, sondern ihre Beschreibung durch Heuristiken.

1. Kontraste

Nachdem am homogenen Ganzfeld deutlich geworden ist, dass Inhomogenitäten und Reizänderungen die Voraussetzungen für das Sehen sind, können die einzelnen Schritte beim Aufbau der visuellen Welt systematisch als Abfolge unterschiedlicher Gruppierungen der Inhomogenitäten verfolgt werden. Ausgangssituation soll eine homogene Fläche mittleren Graus sein. Sie wird durch ein einfaches, regelmäßiges Streifenmuster modifiziert. Ein wichtiges Merkmal von Streifenmustern ist ihre *Ortsfrequenz*. Hiermit wird die Anzahl von Perioden (ein „heller“ und ein „dunkler“ Streifen“) pro Sehinkel auf der Netzhaut bezeichnet. In einem nächsten Schritt wird das Streifenmuster so variiert, dass entweder die Amplitude, die Frequenz oder die Form der zugehörigen Funktion geändert wird. Je nach Manipulation entstehen Inhomogenitäten, die als kontinuierliche oder diskrete Übergänge gesehen werden und die verschiedene Bildbereiche voneinander trennen (Abbildung 18). Im Vergleich der Verlaufsformen ist zu erkennen, dass kontinuierliche Übergänge (Sinusfunktion) leichter zu einem räumlichen Eindruck führen als diskrete. Diskrete oder abrupte Übergänge bilden die Grundlage für die Wahrnehmung von Kanten und Konturen (Rechteckfunktion). Die Sägezahnfunktion schließlich zeigt eine Kombination von beiden. Zwei einfache, leistungsstarke Heuristiken, die auf dem Prinzip der Kontrastbildung aufbauen, lauten: 1. Interpretiere abrupte Intensitätsübergänge als Grenze zwischen zwei Oberflächen. 2. Interpretiere kontinuierliche Übergänge als unterschiedliche Beleuchtung ein und derselben Fläche.

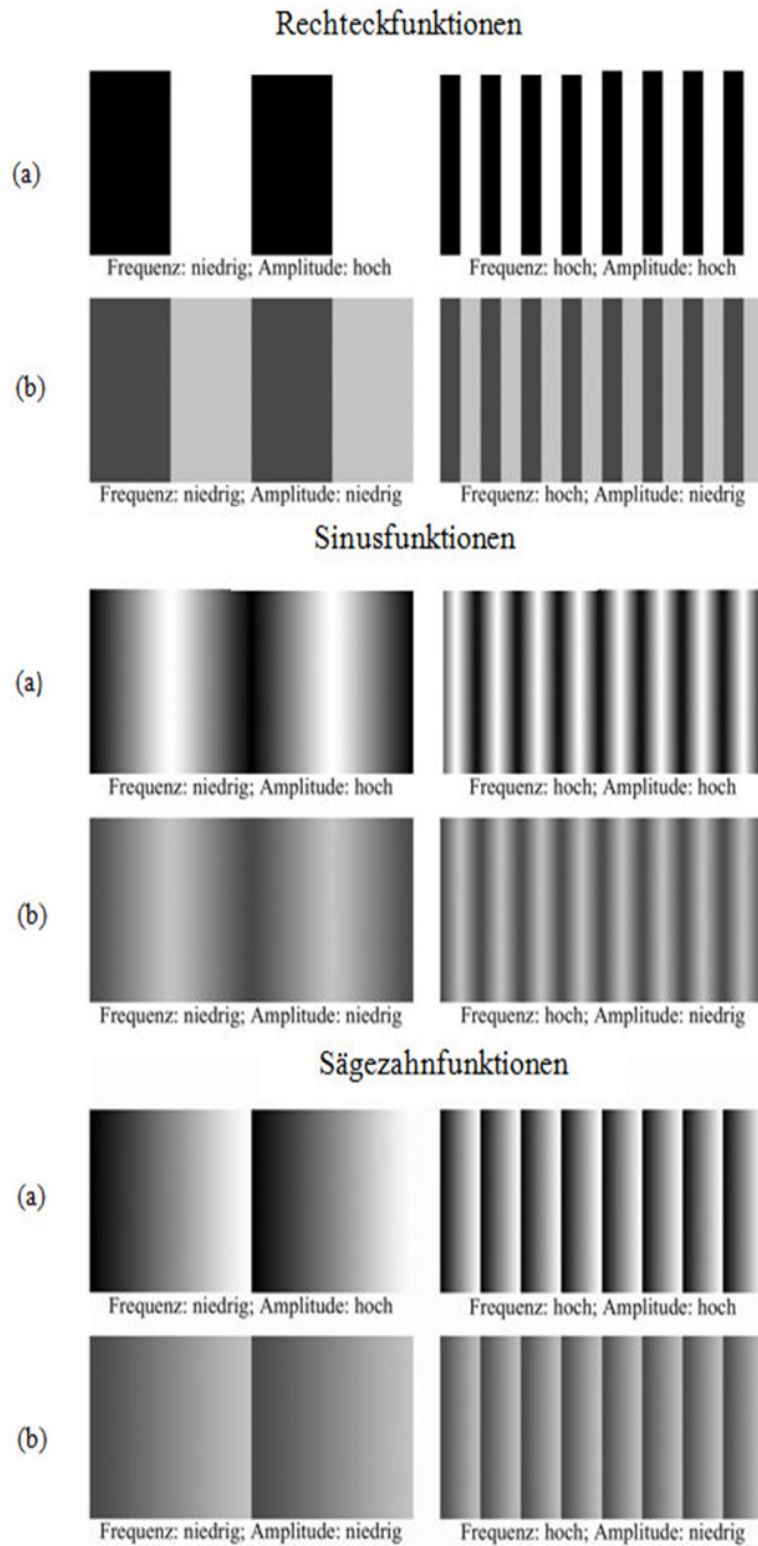


Abbildung 18: Diskrete und kontinuierliche Übergänge. Variationen eines Streifenmusters in Abhängigkeit von Amplitude, Frequenz und Form der Funktion. Der Kontrast ist bei (a) jeweils hoch und bei (b) jeweils niedrig.

Jede Stimulation einer komplexen dreidimensionalen Szene lässt sich als ein Intensitätsmuster aus Ortsfrequenzen beschreiben. Sie können mithilfe der *Fourieranalyse*, einem mathematischen Verfahren, bei dem das Intensitätsmuster in seine Sinuswellen-Komponenten zerlegt wird, ermittelt werden. Diese Beschreibung hat den Vorteil, dass sie Parallelen zur Arbeitsweise des visuellen Systems auf den ersten Stufen der Informationsverarbeitung aufweist.

2. Texturen

Die Abbildung 19 zeigt drei Oberflächen. Bitte versuchen Sie zu erraten, um welche Oberfläche bzw. welches Material es sich handelt. Lässt sich die Aufgabe anhand einer einfachen und qualitativ schlechten Schwarzweißfotografie entscheiden? Zumindest eine Zuordnung sollte gelingen. Offensichtlich gibt es visuelle Merkmale, die eine Identifikation erlauben. Wir müssen die Oberflächen nicht berühren. Die Textur ist das zentrale optische Merkmal. Ähnlichkeit von Oberflächen bedeutet immer Ähnlichkeit von Texturen. Unter den Bedingungen der Umgebungswahrnehmung geht es nicht um die Ähnlichkeit der physikalischen, sondern der visuellen Textur. Wie an verschiedenen Beispielen zu sehen sein wird, ist eine Ähnlichkeit im physikalischen Sinne weder eine notwendige noch eine hinreichende Bedingung. Visuelle Ähnlichkeit ist ausschließlich davon abhängig, ob durch die Stimulation vergleichbare Prozesse im Gehirn aktiviert werden.

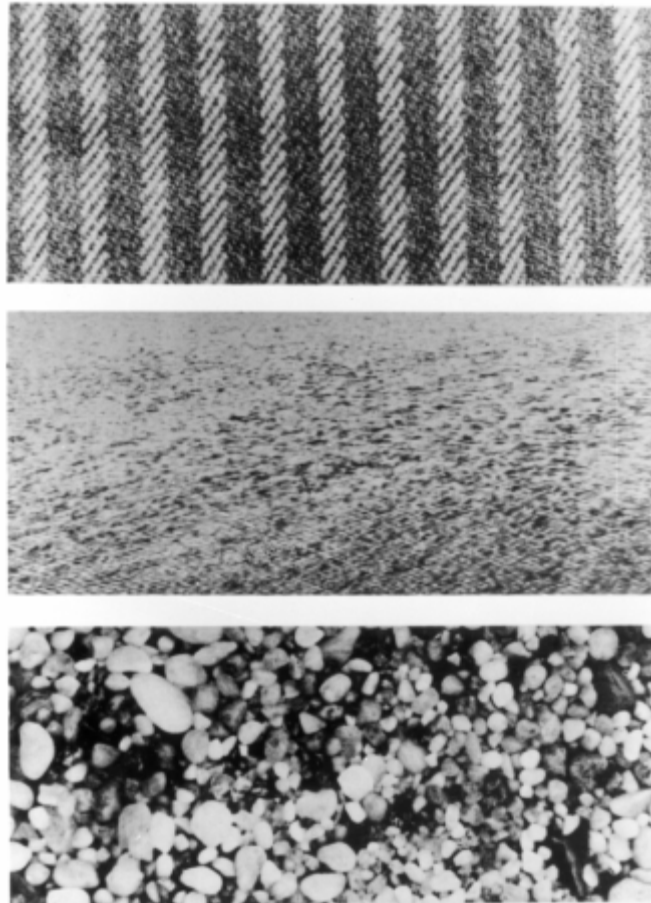


Abbildung 19: Charakteristische Texturen von Oberflächen. Die gesuchten Oberflächen bzw. Materialien sind Stoff, Wasser und Kiesel (Gibson, 1979).

Die Diskrepanz zwischen physikalischer und visueller Struktur fällt der Beobachterin oder dem Beobachter in der Regel nicht auf, weil sie oder er auf den *Repräsentationsaspekt* der Gegenstände fokussiert ist. Angenommen, es klingelt. Man sucht nach dem Telefon, dem Objekt, welches das Klingeln erzeugt. Das von der Oberfläche des Objektes reflektierte Licht liefert zwar die entscheidende Information für die Objektidentifikation, aber die Oberfläche als Oberfläche interessiert nicht. Sie erhält keine Aufmerksamkeit. Gewöhnlich wird das „Was“, nicht aber das „Wie“ gesehen. Es bedarf eines Vorsatzes, um auf den *Oberflächenaspekt* zu achten. Etwa die Frage: Aus welchem Material ist das Telefon? Interessanterweise kann sich der Mensch trotz Anstrengung nicht auf beide Aspekte gleichzeitig fokussieren. Allerdings kann durch

eine Variation des Abstandes – und damit der Größe der retinalen Abbildung – der Bereich bestimmt werden, in dem der Wechsel erfolgt (vgl. auch Pelli, 1999). Objektwahrnehmung und Texturwahrnehmung haben ihren eigenen *kritischen Abstand*. Die Texturidentifikation setzt wegen des begrenzten Auflösungsvermögens des Auges meist einen geringeren Betrachtungsabstand voraus. Bei ihm besteht die Chance, die Materialität mit visuellen Mitteln zuverlässig zu bestimmen. Reichen optische Merkmale nicht aus, wird der Gegenstand berührt, der Tastsinn zu Hilfe genommen. Eine vorläufige Hypothese soll damit verifiziert werden (Prinzip der Zirkularität). Fehlt eine solche Bestätigung, wird die Oberfläche als potentiell gefährlich eingeschätzt, der Umgang mit ihr ist entsprechend vorsichtig.

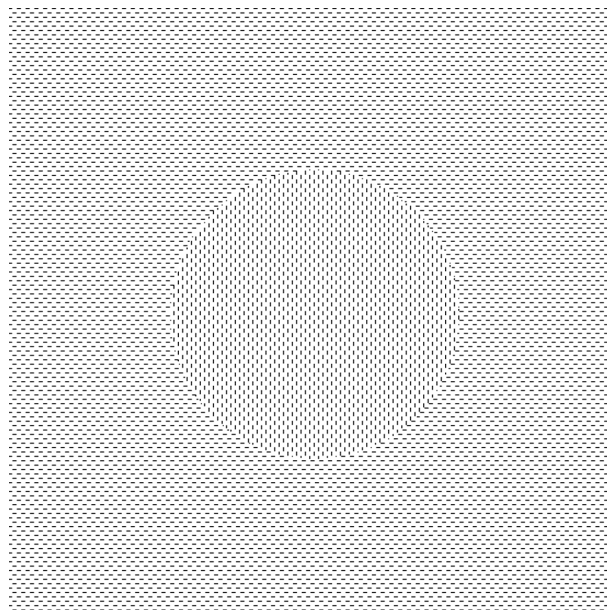


Abbildung 20: Zwei Texturfelder mit sichtbarer Grenze, die durch einen Wechsel der Orientierung erreicht wird (außen Quadrat und innen Kreis).

Eine der zentralen Voraussetzungen für das Erkennen von Figuren und Objekten besteht in der Abgrenzung von Oberflächen. Wo beginnt die Oberfläche, wo endet sie? Eine Hilfe hierbei ist, dass sich die Texturen nach ihrer Ordnung unterscheiden lassen. Stoßen Texturen unterschiedlicher Ordnung aneinander, sehen wir eine Grenze zwischen zwei Flächen (Abbildung 20). Die Texturelemente unterscheiden sich nur durch ihre

Orientierung (horizontal versus vertikal). So lassen sich anschauliche Grenzen und mit ihnen Figuren ohne durchgängige Reizsprünge bilden. Auf der Basis des Prinzips der Gruppierung lauten die zugehörigen Heuristiken: 1. Interpretiere Texturelemente gleicher Orientierung als zu einer Oberfläche gehörig. 2. Interpretiere einen Wechsel in der Orientierung der Elemente einer Textur als Grenze zwischen zwei Oberflächen.

3. Figuren

Jede visuelle Szene ist ein permanenter Wechsel zwischen diskreten und kontinuierlichen Reizübergängen. Damit einer Stimulation Bedeutung zugewiesen werden kann, reicht es nicht, Oberflächen voneinander abzugrenzen und ihre Texturen zu identifizieren. Darüber hinaus ist zu beurteilen, welche Oberfläche zu einer Figur gehört und welche den Hintergrund bildet. Erst diese Trennung schafft die Voraussetzung, um Objekte zu erkennen und Szenen zu interpretieren. Die entscheidende Frage lautet: Wann wird eine Fläche zur Figur? Nach Edgar Rubin (1921) unterscheiden sich Figur und Grund auf der phänomenalen Ebene durch drei Merkmale:

1. Die Figur ist begrenzt und erscheint geformt, der Grund ist unbegrenzt und erscheint ungeformt.
2. Die Figur erscheint fest und hat Gegenstandscharakter, der Grund erscheint eher locker und unbestimmt.
3. Die Figur ist hervortretend und auffallend, der Grund eher zurückweichend und unscheinbar.

Der Konturlinie kommt eine besondere Bedeutung zu. Sie umschließt die Fläche und macht sie zur Figur. Sie trennt Figur und Grund, gehört aber phänomenal zur Figur. Diese *Einseitigkeit der Konturlinie* ist die Voraussetzung für das sichere und schnelle Erkennen von Figuren. Entsprechende Heuristiken gewährleisten die Figur-Grund-Differenzierung. Sie gehen von Wahrscheinlichkeiten für bestimmte

Ereignisse aus: Figuren sind gegenüber dem Grund in der Regel die kleineren Flächen; häufig weisen sie eine stärkere Binnendifferenzierung auf. Innerhalb von Millisekunden werden die Teile einer optischen Anordnung als Figuren identifiziert, die eine reichhaltigere und geordnetere Struktur aufweisen. Der Hintergrund ist meist homogen oder er hat eine Textur, die sich einer Zufallsanordnung nähert. Unter den Bedingungen der Umgebungswahrnehmung erhält er kaum Aufmerksamkeit. Handlungsrelevant sind in den meisten Situationen nur Figuren. Die Natur nutzt dieses Prinzip in seiner Umkehrung zur Tarnung (Camouflage). Drei für die Figur-Grund-Unterscheidung wichtige Heuristiken lauten: 1. Interpretiere die kleinere Fläche als Figur. 2. Ordne die Begrenzung einer Fläche durch eine Konturlinie der Figur zu. 3. Interpretiere die Figur als vor dem Grund liegend.

Von der Wahrnehmung der Helligkeitsverteilungen zur Wahrnehmung von Figuren besteht ein qualitativer Sprung. Nach Auffassung der Gestalttheoretiker Wolfgang Köhler, Max Wertheimer und Kurt Koffka, die sich in der ersten Hälfte des 20. Jahrhunderts intensiv mit der Erforschung der grundlegenden Gruppierungs- und Gestaltgesetze beschäftigten, kommen Gestalten – und damit auch Figuren – phänomenale Eigenschaften zu, die die einzelnen Elemente nicht aufweisen (Prinzip der Übersummativität). Nach Auffassung der Gestalttheoretiker gliedert sich das visuelle Feld durch autochthone (ursprüngliche) Gesetzmäßigkeiten nicht nur in Texturen oder Helligkeitsverteilungen, sondern in bedeutungshaltige Ganze, die wir identifizieren und wiedererkennen können (vgl. Kapitel IV).

4. Objekte

Im Verständnis der Wahrnehmungsforschung sind Figuren die Darstellung von Objekten in einer Ansicht. Deshalb sollte in Bezug auf die Bildwahrnehmung von Figuren und nicht von Gegenständen oder Objekten gesprochen werden. Für die Umgebungswahrnehmung und für

einen Vergleich der Figuren mit den Gedächtnisbildern ist jedoch eine Objekterkennung notwendig. Sie setzt eine Beschreibung der Gegenstände voraus, die unabhängig von der aktuellen Perspektive ist. Denn das Objekt muss auch dann erkannt werden, wenn es aus einem anderen Blickwinkel gesehen wird oder wenn es in der Zwischenzeit seine Position verändert hat. Wie erreicht das visuelle System diese Multiperspektivität, wenn doch das Netzhautbild immer perspektivenabhängig ist?

Eine Reihe von Theorien ist zur Beantwortung dieser Frage aufgestellt worden. Sie lassen sich in zwei Gruppen zusammenfassen. Die erste geht davon aus, dass *strukturelle Beschreibungen* der Objekte notwendig sind. Jedes Objekt lässt sich als eine einzigartige Kombination von elementaren Teilkörpern auffassen. Werden diese Teilkörper vom visuellen System erkannt, kann auch das Objekt identifiziert werden. Die zweite Gruppe postuliert, dass ausgewählte *Ansichten* eines Gegenstandes gespeichert werden und die aktuelle Ansicht hiermit verglichen wird. Stellvertretend für die Positionen sollen das Geonen-Modell (Biederman, 1985, 1987; Hummel und Biederman, 1992) und das Modell der kanonischen Ansichten (Palmer, Rosch und Chase, 1981) dargestellt werden.

Das *Geonen-Modell* nimmt an, dass Objekte in elementare Teilkörper (geometrische Ione) zerlegt werden. Sie sind die Grundbausteine der Objektwahrnehmung. Alle Gegenstände lassen sich aus nur 36 Elementen zusammensetzen; die Vielfalt ergibt sich aus deren Kombination. Abbildung 21 zeigt das Vorgehen. Bereits eine geringfügig veränderte Anordnung derselben Geone führt zu einer Vielzahl von Objekten. Wie schnell ein Gegenstand erkannt wird, hängt nach dieser Theorie davon ab, wie leicht die Geone identifiziert werden können. Nach Biederman (1987) geht es bei der Objektwahrnehmung um die Rückgewinnung elementarer Teilkörper. Können die Geone identifiziert werden, kann auch der Gegenstand bestimmt werden. Deshalb ist es, wenn auch mit etwas Aufwand, möglich, in der Abbildung 22 bei (a) eine Taschenlampe zu erkennen, während das bei (b) ausgeschlossen ist. Biederman konnte

darüber hinaus zeigen, dass bei einer Darbietung von unvollständigen Umrisszeichnungen der Gegenstände sowohl die Zeitdauer für die richtige Erkennung als auch die Anzahl der Fehler direkt von der Menge der gezeigten Geone abhängt.

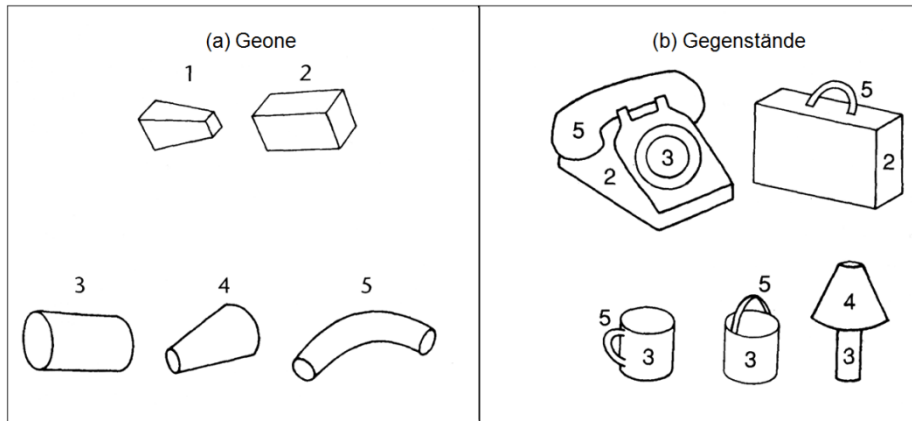


Abbildung 21: Geone als Grundbausteine der Objektwahrnehmung (Biederman, 1987).

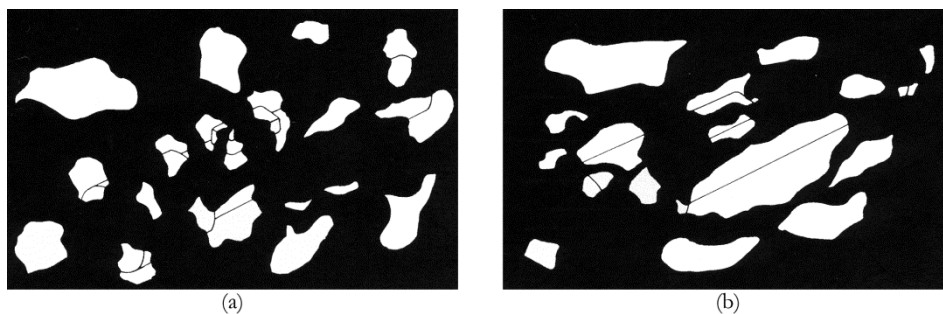


Abbildung 22: Versteckte (a) und unauffindbare (b) Taschenlampe (Biederman, 1987).

Das *Modell der kanonischen Ansichten* hingegen geht von Probabilitäten aus. Untersucht man statistisch, wie häufig ein bestimmter Gegenstand (z. B. eine Teekanne oder ein Stuhl) in einer bestimmten Ansicht gesehen wird und welche Ansicht bevorzugt wird, dann zeigt sich, dass es eindeutige Präferenzen und einen korrelativen Zusammenhang zwischen Präferenz und Häufigkeit der Ansicht gibt. Gegenstand und Ansicht sind nicht zufällig. Gewählt wird zumeist eine besonders charakteristische

Perspektive (Abbildung 23). Die resultierende Figur enthält spezifische Informationen, die es erlauben, den Gegenstand ohne einen Wechsel der Ansicht zu identifizieren. Diese Perspektiven werden in der Wahrnehmungsforschung als *kanonische Ansichten* bezeichnet (Palmer, Rosch und Chase, 1981). Das Modell der kanonischen Ansichten geht im Gegensatz zum Geonen-Modell davon aus, dass charakteristische Ansichten gespeichert werden und dass die neu eintreffenden Informationen mit diesen Ansichten verglichen werden. Die Arbeiten von Bülthoff, Tarr und Edelman (1994) zeigen, dass solche Ansichten nicht nur bevorzugt werden, sondern dass sie auch kürzere Erkennungszeiten aufweisen. Offensichtlich gibt es mehr als eine herausragende Ansicht; in der Regel sind es zwei bis drei.

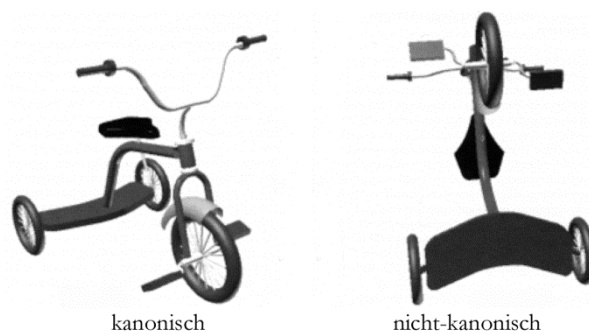


Abbildung 23: Kanonische (links) und nicht-kanonische (rechts) Ansicht eines Dreirads (Bülthoff, Edelman und Tarr, 1994).

Der Gesamtvorgang der visuellen Informationsverarbeitung, von der Analyse der ersten Inhomogenitäten und Helligkeitskontraste bis zur Objektidentifikation, kann in frühe und späte Phasen gegliedert werden. Zwischen beiden gibt es deutliche Unterschiede. Die Informationsverarbeitung in den frühen Phasen ist stärker datengesteuert. Sie ist die Analyse des retinalen Bildes. Lokale Zusammenfassungen und Interpretationen stehen im Vordergrund. Die Mechanismen, die hier wirksam sind, scheinen weitgehend unabhängig von Erfahrung und individuellen Lernprozessen zu sein. Sie können besser durch Algorithmen beschrieben werden. Je weiter die Informationsverarbeitung

fortschreitet, umso größer sind die Regionen des visuellen Feldes, die in die Interpretation einbezogen werden. Prozesse des induktiven Schließens und damit Heuristiken werden bedeutsamer. Die konstruktiven Aspekte des Sehens treten in den Vordergrund. Das Sehen wird hypothesengeleitet und weist nun stärkere Bezüge zum Gedächtnis auf. „Aufwärts“ und „abwärts“ gerichtete Prozesse ergänzen sich in idealer Weise.

5. Szenen

Dieses Wechselspiel wird bei der Wahrnehmung von Szenen noch deutlicher. Oft bestimmt der Kontext die Bedeutung eines Objektes und Mehrdeutigkeiten lassen sich erst durch die Einbindung in eine Szene auflösen. Der Einfluss von Augenbewegungen, Aufmerksamkeit und die Zirkularität des Sehvorgangs sind bei Szenen evident. Während auf niedrigeren Verarbeitungsebenen die Hauptaufgabe des visuellen Systems in der Extraktion von physikalischen Reizeigenschaften gesehen wird, steht jetzt die semantische Verarbeitung im Vordergrund. Als *Szene* wird folgerichtig eine semantisch kohärente Ansicht einer realistischen Umgebung bestehend aus Hintergrundelementen und einer Vielzahl unterscheidbarer Objekte verstanden (Henderson und Hollingworth, 1999, S. 244). Die Abgrenzung von Hintergrund und Objekten ist in hohem Maße abhängig vom Maßstab der Gesamtszene: Stellt ein Schreibtisch in einer Büroszene ein einzelnes Objekt dar, wird er bei einer Vergrößerung des Maßstabes zum Hintergrund. Jetzt bilden die auf dem Schreibtisch platzierten Gegenstände eine eigenständige Szene.

Typisch für die Wahrnehmung von Szenen ist, dass nur kleine Ausschnitte foveal und damit in voller Schärfe abgebildet werden. Für die Beobachterin oder den Beobachter ist es unerlässlich, einzelne Regionen in aufeinanderfolgenden Fixationen zu fokussieren. Die Registrierung von Augenbewegungen spielt bei den Untersuchungen zur Szenenwahrnehmung deshalb eine hervorgehobene Rolle (zu den einzelnen Methoden vgl. Duchowski, 2007). Auffälligerweise arbeiten fast

alle bisherigen Untersuchungen mit Bildmaterial. Streng genommen sind es Untersuchungen zur Bildwahrnehmung, auch wenn die Bilder Ausschnitte der Umgebung zeigen. Schon frühe Analysen von Blickbewegungen zeigten, dass die Abfolge der Fixationen vor allem von zwei Faktoren beeinflusst wird: 1. den strukturellen Merkmalen des Bildes und 2. den Intentionen der Betrachterin oder des Betrachters.

In einer klassischen Untersuchung berichtete Yarbus (1967) von Augenbewegungsmustern bei unterschiedlichen Instruktionen der Proband*innen am selben Bild (Abbildung 24): zunächst die Situation ohne spezifische Information (B), dann die Muster für die Beurteilung des Vermögens der Familie (C), das Alter der Personen (D), die Tätigkeit der Familie unmittelbar vor dieser Situation (E) und die Kleider der Personen (F).

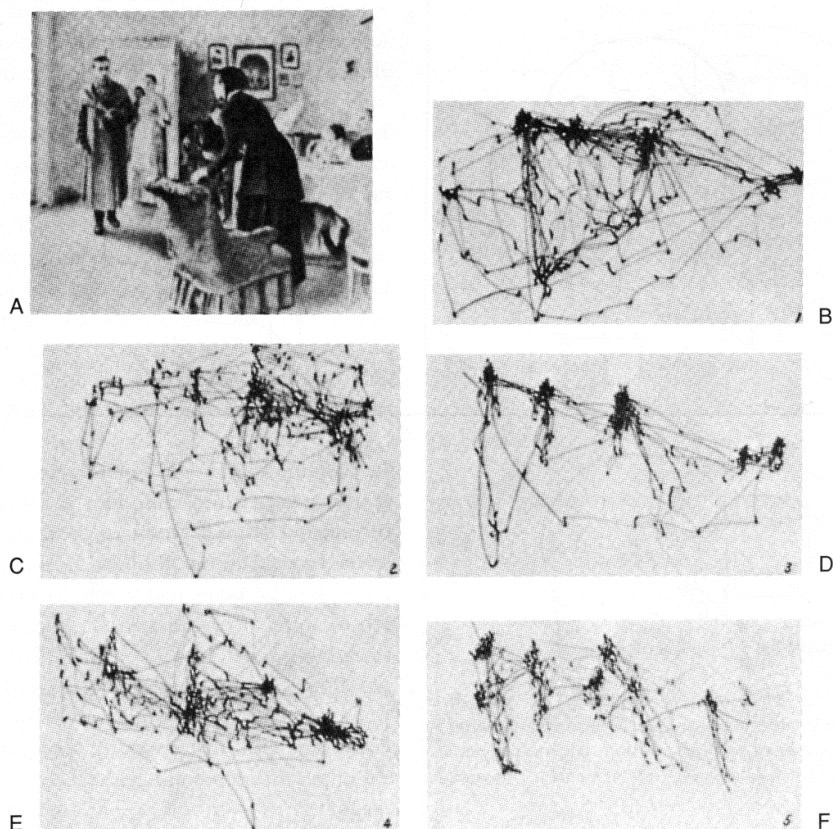


Abbildung 24: Augenbewegungspfade: B ohne spezifische Aufgabe, C-F mit unterschiedlichen Aufgaben (Yarbus, 1967).

Der Vergleich der Aufzeichnungen verdeutlicht die Abhängigkeit der Muster vom Ziel der Informationssuche. Doch sprechen miteinander vergleichbare Teilmuster auch für eine reizabhängige Steuerung. Das wird deutlicher, wenn für die fixierten Bereiche eine höhere Auflösung gewählt wird. Dann zeigt sich, dass „top-down“ nur die Bildregion festgelegt wird, die Feinsteuerung aber durch die Strukturmerkmale des Bildes erfolgt. Die Augenbewegungen werden in einem gewissen Ausmaß durch Eigenschaften der Szene bzw. des Bildes bestimmt. Das Bild leitet die Betrachterin oder den Betrachter. Zahlreiche neuere Untersuchungen haben gezeigt, dass für zentral angeordnete Bildteile mit einer gewissen Größe und starken Helligkeits- oder Farbunterschieden die Wahrscheinlichkeit steigt, früh entdeckt zu werden (für einen Überblick solcher Merkmale vgl. Henderson und Hollingworth, 1999). Da die Suche foveal erfolgt, geht es um die Figur-Grund-Trennung und die Abgrenzung zwischen Figuren; der effektivste Weg, um mit möglichst wenigen Zwischenschritten die Mehrdeutigkeiten zu reduzieren. Globale Lösungen haben Priorität vor lokalen. Schließlich wird eine Gruppierung gesucht, die es erlaubt, die gesamte Stimulation als eine in sich konsistente Szene zu interpretieren.

Insgesamt werden Bildregionen bevorzugt, die sowohl visuell als auch semantisch besonders informativ sind; sie zeigen eine höhere Fixationsdichte. Die ersten Fixationen scheinen weitgehend unter der Kontrolle optischer Merkmale des Bildes zu stehen (wichtig sind z. B. starke Helligkeits- oder Farbkontraste). Sind semantische Aspekte relevant, dann ist es am ehesten die globale Klassifikation der Szene (z. B. als „Landschaft“ oder als „Gespräch“). Mit fortschreitender Analyse werden lokale Bereiche inspiziert, wobei es zu einem engen Zusammenspiel von visuellen und semantischen Merkmalen mit dem Ziel einer konsistenten Interpretation der Szene kommt. Gibt es Teilbereiche, die mit dieser Interpretation nicht vereinbar sind (beispielsweise ein Küchenherd in einem Kinderzimmer), erhält dieser Bereich des Bildes häufigere und längere Fixationen, um die Widersprüche aufzulösen.

Der Mensch erlebt seine Umgebung als ein einziges, farbiges Bild, das über den gesamten Bereich eine hohe Auflösung aufweist. Die Basis für dieses farbfotoähnliche Perzept kann nicht das Resultat einer einzelnen Fixation sein. Viele diskrete Einzelbilder, die von blinden Phasen (sakkadische Suppression) unterbrochen sind, müssen zusammengefasst werden. Wie gelingt die Integration? Wie erreichen wir Stabilität? Die Vorstellungen hierzu sind noch vage. Sicher ist, dass bei komplexen Szenen nur ein kleiner Teil der Information aus den Einzelbildern über die Sakkaden hinweg erhalten bleibt. Unklar ist die Speicherung der Information. Wenn sie, was die überwiegende Anzahl der Arbeiten vermuten lässt, in einem abstrakten und nicht perzeptuellen Format aufbewahrt wird, muss erklärt werden, wie aus diesem Format mithilfe schematischer und aktueller Information ein kohärentes Bild entsteht.

Kapitel IV: Gliederung und Gruppierung

„Ich stehe am Fenster und sehe ein Haus, Bäume, Himmel. Und könnte nun, aus theoretischen Gründen, abzuzählen versuchen und sagen: da sind ... 327 Helligkeiten (und Farbtöne). (Habe ich „327“? Nein; Himmel, Haus, Bäume; und das Haben der „327“ als solcher kann keiner realisieren.) Und seien in dieser sonderbaren Rechnung etwa Haus 120 und Bäume 90 und Himmel 117, so habe ich jedenfalls *dieses* Zusammen, dieses Getrenntsein, und nicht etwa 127 und 100 und 100: oder 150 und 177. In dem bestimmten Zusammen, der bestimmten Getrenntheit sehe ich es; und in welcher Art des Zusammen, der Getrenntheit ich es sehe, das steht nicht einfach in meinem Belieben: ich kann durchaus nicht etwa nach Belieben jede andere gewünschte Art der Zusammengefaßtheit einfach realisieren.“ Mit diesen Worten beginnt einer der bedeutendsten Texte der Wahrnehmungsforschung: Max Wertheimers „Untersuchungen zur Lehre von der Gestalt II“ (Wertheimer, 1923, S. 301).

1. Gestaltgesetze

Max Wertheimers Text erschien 1923 und stützt sich auf Untersuchungen aus den Jahren 1911-14. Wie gliedert sich das Sehfeld in bedeutungshaltige Ganze? Abgrenzung der Wahrnehmungserlebnisse und ihre begriffliche Fassung sind die Ziele dieser phänomenologisch-experimentellen Forschung. Für seine Untersuchungen verwendete Wertheimer einfaches und künstliches Material: Punkte und Punktmuster, Linien und Linienverläufe, geometrische Figuren. Dahinter steht die Überzeugung, dass die Gesetzmäßigkeiten des Psychischen deutlicher hervortreten, wenn das Material aus dem Kontext der Alltagerfahrungen gelöst wird und die Einflussfaktoren isoliert auftreten (ein schönes Beispiel für Wertheimers phänomenologisch-experimentelles Vorgehen ist auch seine 1912 erschienene Arbeit zu Scheinbewegungen und zum Phi-Phänomen).

Sieben Gruppierungsgesetze, die gleichzeitig Gestaltgesetze sind, bilden die Quintessenz seiner Forschung: Nähe, Gleichheit und Ähnlichkeit, gemeinsames Schicksal oder übereinstimmendes Verhalten, objektive Einstellung, Aufgehen ohne Rest, durchgehende Kurve oder glatter Verlauf, Geschlossenheit (vgl. Abbildung 25).

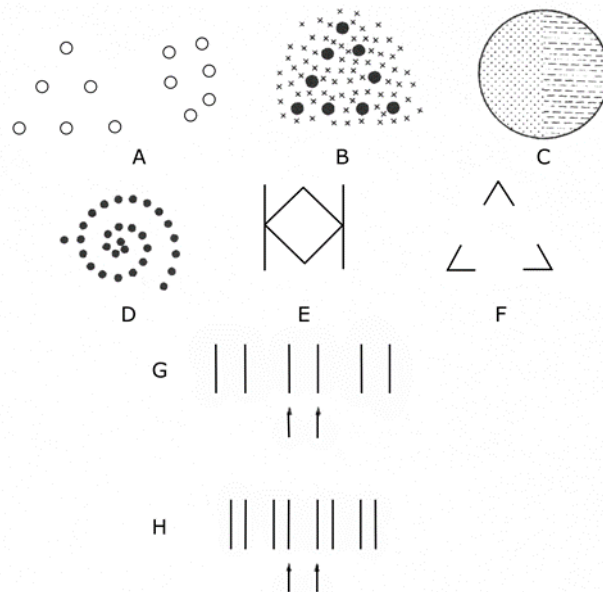


Abbildung 25: Auswahl von Gruppierungs- und Gestaltgesetzen. A, G, und H: Gesetz der Nähe; B und C: Gesetz der Gleichartigkeit; D: Gesetz des glatten Verlaufs; E und F: Gesetz der Geschlossenheit (Coren, Porac und Ward, 1984).

Diese „Einzelgesetze“ versteht Wertheimer als Auswirkungen eines noch allgemeineren Gesetzes, der „Prägnanztendenz“ oder dem „Gesetz der guten Gestalt“. Ihm gesteht er eine übergeordnete Rolle zu. Koffka (1935, S. 110) formuliert den Grundgedanken so: „Psychological organization will always be as ‘good’ as the prevailing conditions allow.“ Historisch gesehen war Max Wertheimer nicht der Erste, der sich mit Gruppierungs- oder Gestaltgesetzen beschäftigte. George Berkley (1709) stellte in seiner „New Theory of Vision“ bereits die Frage, wieso die menschliche Gestalt trotz ihrer Umkehrung im Netzhautbild problemlos erkannt wird und

welche figuralen Eigenschaften verantwortlich sind. Der entscheidende Schritt hin zur Formulierung von Gesetzmäßigkeiten erfolgte durch Christian von Ehrenfels (1890) in seiner Schrift „Über Gestaltqualitäten“. Erstmals wird das Konzept der *Gestalt* in seiner heutigen Bedeutung für die Wahrnehmungsforschung verwendet: ein „Ganzes“, das über die Eigenschaften „Übersummativität“ und „Transponierbarkeit“ verfügt (vgl. Kapitel I). Elias Müller (1903) schließlich erläutert in seiner Komplextheorie der Wahrnehmung erste „Kohärenz-Faktoren“ wie Ähnlichkeit, Nachbarschaft, symmetrische Lage und Eingeschlossenheit.

Die Forschung in der Nachfolge von Wertheimer bestätigte die ursprüngliche Aufzählung und benannte weitere Kandidaten. Insgesamt wurden Vorschläge für über 100 Gestaltgesetze gemacht! Vorrangig handelt es sich um Ableitungen oder Spezialfälle. Wolfgang Metzger (1967) bemühte sich um eine erneute Systematisierung. Dabei lehnte er sich eng an die Begrifflichkeit von Wertheimer an. Dem Faktor „Gleichartigkeit“ (entspricht „Gleichheit und Ähnlichkeit“ bei Wertheimer) allerdings gesteht er den Status einer notwendigen Bedingung zu: „Der Faktor der Gleichartigkeit hat unter sämtlichen Faktoren eine Sonderstellung. Sämtliche übrigen – gegenständlichen und subjektiven – Faktoren setzen eine vorgängige Wirkung dieses Grundfaktors voraus. Sie können nur wirksam werden, nachdem durch Qualitätssprünge das Sehfeld „zerrissen“ ist und irgendwelche – jeweils in sich verhältnismäßig homogene, aber durch Verschiedenheit der Qualität von ihrer Umgebung abgesetzte – Teilgebiete (Punkte, Striche, Flecken) darin ausgesondert sind.“ (Metzger, 1967, S. 700 f).

2. Transformationsgüte

Nach dem 2. Weltkrieg setzte die Forschung neue Akzente. Seit den 1950er Jahren finden sich vermehrt informationstheoretische Arbeiten (vgl. z. B. Hochberg und McAlister, 1953; Attneave, 1954, 1955; Attneave und

Arnoult, 1956; Hochberg und Brooks, 1960; Garner und Clement, 1963; Garner, 1974; Leeuwenberg 1971, 1978; Palmer, 1982). Gemeinsam ist ihnen der Versuch, die gestalttheoretischen Vorschläge messbar und vorhersagbar zu machen. Unter Anwendung des Redundanzkonzeptes wird angenommen, dass „gute Gestalten“ weniger Information enthalten. Prägnanz oder „figural goodness“ wird mathematisch als Invarianz gegenüber Transformationen beschrieben (vgl. Abbildung 26). Ein Beispiel: Wird ein griechisches Kreuz, bei dem alle vier Arme gleich lang sind, um 90°, 180° und 270° rotiert, so wird es jeweils auf sich selbst abgebildet; wir sehen eine identische Figur. Wird dagegen ein lateinisches Kreuz, bei dem der Längsbalken länger ist als der Querbalken und dieser den Längsbalken oberhalb dessen Mitte kreuzt, um dieselben Winkel rotiert, entstehen zusätzlich zur Ausgangsfigur drei unterscheidbare Figuren. Das griechische Kreuz bleibt gegenüber den Rotationen aufgrund seiner Symmetrieeigenschaften invariant. Es hat die „bessere Gestalt“.

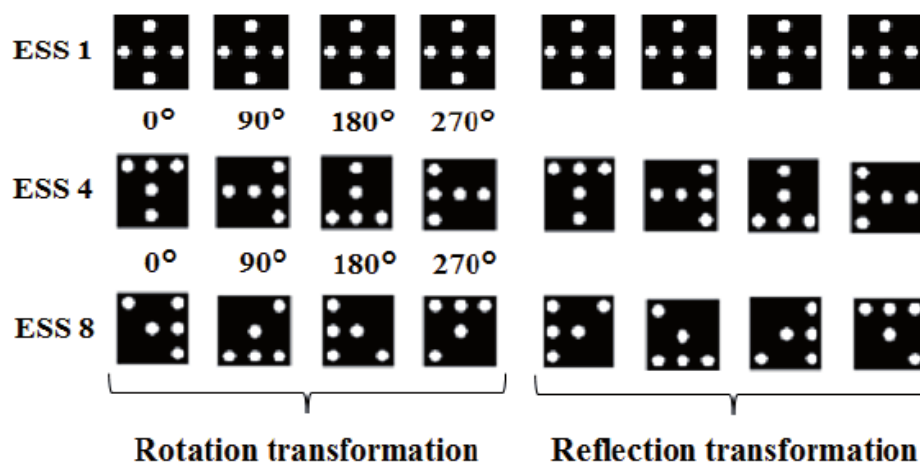


Abbildung 26: Invarianz gegenüber Spiegelungen und Rotationen nach Garner und Clement (1963); EES = equivalent set seize; die Ursprungsfigur (0°) in der oberen Zeile links hat die „beste Gestalt“, da sie gegenüber allen Transformationen invariant bleibt.

Seit den 1990er Jahren rückt zunehmend die Frage nach den neuroanatomischen und neurophysiologischen Grundlagen der Gruppierungs- und Gliederungsprozesse in den Vordergrund (vgl. z. B. Singer und Gray, 1995; Castello-Branco, Goebel, Neuenschwander und

Singer, 2000; Han, Song, Ding, Yund und Woods, 2001; Rolfsema, Lamme und Spekrejse, 2004; Rolfsema, 2006; Seymour, Karanath und Himmelbach, 2008; Casco, Campana, Han und Guzzon, 2009). Fragen der Gruppierung werden heute häufig im Zusammenhang mit Theorien der Figur- und der Objektwahrnehmung diskutiert (vgl. Kapitel III). Derzeit bemühen sich kognitionswissenschaftliche Arbeiten, gestalttheoretische, informationstheoretische und neurowissenschaftliche Perspektiven zusammenzuführen (für einen Überblick zum aktuellen Forschungsstand vgl. Wagemans, 2015).

Es zeigt sich, dass die von Wertheimer beschriebenen Gesetzmäßigkeiten auch heute noch Bestand haben. Zunehmend wird aber auch deutlich, dass weitere „ernsthafte Kandidaten“ in die Diskussion einzubeziehen sind (vgl. z. B. Brooks, 2015 sowie die „emergent features“ bei Pommerantz und Cragin, 2015). Die Gründe liegen sowohl auf der methodischen wie auf der inhaltlichen Ebene. Zum einen können wir komplexes und alltagsnahes Material mithilfe von Computern heute systematisch variieren. Die Einschränkung auf einfaches, figürliches Material ist nicht mehr zwingend. Zum anderen erweist sich, dass Wertheimer seine Untersuchungen an einem Spezialfall durchgeführt hat: der Bildwahrnehmung. Verglichen mit der alltäglichen Umgebungswahrnehmung werden die Anpassungsleistungen des Sehsystems an die Gegebenheiten und Wahrscheinlichkeiten der Außenwelt unterschätzt (Gibson 1950, 1966, 1979).

Für das Verständnis der „klassischen“ Gruppierungsgesetze ist entscheidend, dass es nach Auffassung der Gestalttheoretiker in der Wahrnehmung keine vorgefundenen Einheiten gibt. Vielmehr ist die Gliederung des Sehfeldes stets von den aktuellen Gegebenheiten abhängig. Wie das Zitat von Wertheimer erläutert, werden die Gruppierungen zum überwiegenden Teil nicht durch unseren Willen bestimmt, sondern folgen automatischen und vorbewussten Prozessen. Die Gesetze dienen sowohl

der Bildung von Gruppen und Zusammenfassungen, wie auch der Gliederung und Abgrenzung. Die Segmentierung des Wahrnehmungsfeldes in Einheiten und die Zusammenfassung von Einheiten zu „Ganzen“ oder „Ganzheiten“ sind inverse Prozesse. Die Gruppierungsgesetze sind gleichzeitig Gliederungsgesetze. Zwar sind die Aufgaben für das Sehsystem unterschiedlich, aber die Regeln sind identisch.

Das Zusammenspiel der Gruppierungsgesetze wird als „ceteris paribus“ bezeichnet: wenn sonst alle Bedingungen gleich bleiben. Folglich gelten sie nicht absolut, sondern relativ. Im Regelfall wirken mehrere von ihnen gleichzeitig, und es entstehen sowohl ein Zusammenspiel wie ein Wettstreit (zahlreiche Beispiele finden sich bei Metzger, 1975). Primäres Ziel dieser ordnungsstiftenden Prozesse ist die Reduktion der im retinalen Bild enthaltenen Mehrdeutigkeiten. Bewegungen wie Handlungen setzen eine eindeutige Interpretation der Stimulation voraus. Das visuelle System akzeptiert keine dauerhafte Mehrdeutigkeit. Kann zunächst keine Eindeutigkeit hergestellt werden, dann wird das Wahrnehmungsfeld neu organisiert. Nach heutigem Verständnis sind die Gliederungs- und Gruppierungsgesetze leistungsstarke Heuristiken, die das visuelle System zur Ambiguitätsreduktion einsetzt. Sie haben ihre Entsprechung in den Probabilitäten der Außenwelt (vgl. auch Elder und Goldberg, 2001).

3. Gliederung des Ganzfeldes

Bevor einzelne Heuristiken vorgestellt werden, sind einige Begriffe zu erläutern. Das *Ganzfeld* bezeichnet die Stimulation im Gesichtsfeld der Beobachterin oder des Beobachters zu einem Zeitpunkt und ist damit nicht auf das homogene Ganzfeld beschränkt. Voraussetzung für Zusammenfassungen wie für Gliederungen sind Inhomogenitäten oder Veränderungen. Ihre Verteilung bzw. ihre Abfolge kann zufällig oder regelmäßig sein. Die Regelmäßigkeiten bilden den Ausgangspunkt für die

Heuristiken zur Ambiguitätsreduktion. Je nach Betrachtungsweise führen sie zur Zusammenfassung von *Elementen* zu *Gruppen* oder *Teilen* zu *Ganzen* sowie zur Untergliederungen des Ganzfeldes in *Regionen*. Dabei lässt sich die Frage, was ein „Element“ ist bzw. was „Teile“ oder „Ganze“ sind, nicht a priori entscheiden, sondern ist wiederum abhängig von der Analyse des Ganzfeldes: ein rekursiver Prozess.

Die Überlegungen zum Verhältnis von verarmter und optimaler Wahrnehmung legen nahe, von der Gliederung des Wahrnehmungsfeldes unter natürlichen Umgebungsbedingungen auszugehen. Nicht mit den kleinsten Einheiten und ihren Zusammenschlüssen, sondern mit dem Ganzfeld zu beginnen und die *Prozesse der Segmentierung* stärker zu betonen. Wertheimer legte den Fokus auf das Entstehen von zweidimensionalen Figuren. An dieser Stelle endet der Wahrnehmungsprozess jedoch nicht. Unter den Bedingungen der Umgebungswahrnehmung gliedert sich das Gesichtsfeld in Objekte und diese wiederum sind Bestandteile von Szenen. Die vorgeschlagenen Ergänzungen behandeln verstärkt Fragen der globalen räumlichen Organisation sowie der zeitlichen Veränderung der Stimulation. Da der Ausgangspunkt die Umgebungswahrnehmung ist, ergeben sich weitere Akzentverschiebungen. Zum einen rücken neben den sachbezogenen auch die beobachtungsbezogenen Faktoren stärker in den Vordergrund, zum anderen gewinnen die höheren Stufen der visuellen Informationsverarbeitung und erfahrungsabhängige Prozesse an Bedeutung. Implizit ging Wertheimer wohl vom Ideal eines homogenen Ganzfeldes aus, in das sukzessiv einzelne Elemente eingefügt werden. Im Folgenden wird das natürlich gegliederte Ganzfeld als Ausgangspunkt dienen.

Einschließlich der bei Wertheimer (1923) genannten Kandidaten werden nachfolgend 20 Gliederungs- und Gruppierungsheuristiken vorgestellt. Die relativ hohe Anzahl legt die Bildung von Untergruppen nahe. Sie sind

lediglich als Akzentuierung zu verstehen: Die Heuristiken 1-7 gehen stärker vom natürlichen Ganzfeld und seiner Untergliederung aus, während 8-14 die Gruppierung von Elementen zu größeren Einheiten betonen. Bei den Heuristiken 15-17 kommen Veränderungen der Stimulation über die Zeit hinzu und bei 18-20 steht die Aktivität der Beobachterin oder des Beobachters im Vordergrund.

(1) Perspektive

Gehe bei der Gliederung des Ganzfeldes von nur einer Perspektive zu einem Zeitpunkt aus.

In Bezug auf Gliederungs- und Gruppierungsprozesse bildet die Perspektive die oberste Analyseebene. Der Ausgangspunkt der visuellen Informationsverarbeitung sind die projektiven Abbildungen der Außenwelt auf den Retinae der beiden Augen. Deshalb haben wir – streng genommen – zwei leicht unterschiedliche Perspektiven, die im Laufe des Informationsverarbeitungsprozesses zu einer Perspektive integriert werden (Abbildung 27). Alle Informationen, die in der Stimulation enthalten sind, werden bei einer Gruppierung unter den Aspekt der Perspektive untergeordnet. Deshalb wird die Perspektive uns nicht bewusst. Wir erwarten zu Recht Widerspruchsfreiheit. Werden, wie in einem kubistischen Bild oder bei einer Anamorphose, verschiedene Ansichten gleichzeitig gezeigt, dann fällt es uns schwer, den Gegenstand zu erkennen. Treten aufgrund einer neurologischen Störung Doppelbilder auf, so empfinden wir diese als störend.

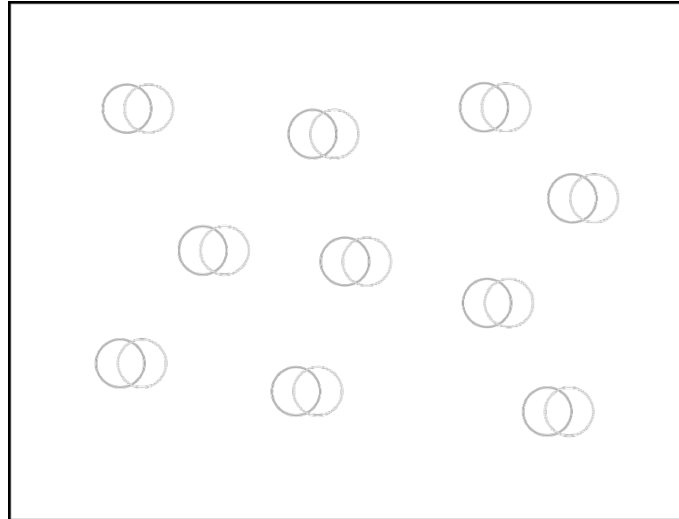


Abbildung 27: Perspektive. Das visuelle System hat aufgrund der Binokularität zunächst zwei leicht unterschiedliche Perspektiven auf ein und denselben Gegenstand. Die Informationen aus beiden Augen werden im Laufe des Informationsverarbeitungsprozesses zu einer einzigen Perspektive integriert.

(2) Horizont

Gliedere das Ganzfeld in zwei Regionen, die durch die Horizontlinie getrennt werden. Gruppierere Elemente, die unter- oder oberhalb der Horizontlinie liegen, jeweils als zusammengehörig. Nutze für diese Zuordnung den untersten Punkt des Elementes.

Der Horizont ist eine evolutionäre Konstante. Er bildet die Grenze des jeweils Sichtbaren und gleichzeitig den Übergang zum aktuell nicht Sichtbaren. Bewegt sich die Beobachterin oder der Beobachter, verschiebt sich der Horizont. Er gibt eine Richtung vor und bleibt doch unerreichbar. Je nach Standort und Kopfneigung liegt er höher oder tiefer im Gesichtsfeld und erzeugt eine waagerechte Linie, welche die beiden Regionen Himmel und Erdboden voneinander trennt (vgl. Kapitel I). Er führt zu einer globalen Zweiteilung des natürlichen Ganzfeldes (Gibson, 1950). Die Beobachterin oder der Beobachter verortet sich relativ zum Horizont sowohl in Bezug auf die Distanz als auch auf den Neigungswinkel

(vgl. Abbildung 28). Der Horizont ermöglicht die Bestimmung der relativen Entfernung von Objekten zueinander und zum Standort. Für die Beobachterin oder den Beobachter dient die Horizontlinie als Orientierung und zur Kontrolle der eigenen Bewegung. Auch wenn der Horizont von der aktuellen Position aus nicht sichtbar ist, bleibt er präsent. Die Bedeutung des Horizonts fällt uns im Alltag kaum auf, Pilot*innen ist sie vertraut. Eines der wichtigsten Instrumente im Cockpit eines Flugzeugs ist der künstliche Horizont. Mit ihm kann nach jeder Steig- und Sinkbewegung beurteilt werden, ob die Ausrichtung wieder horizontal ist.

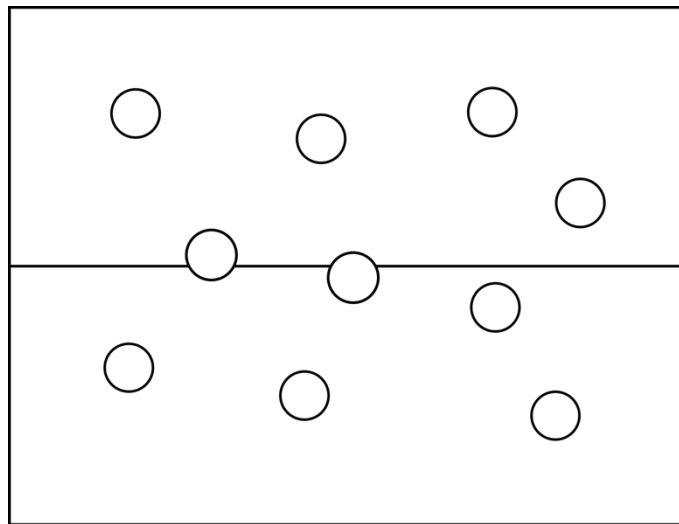


Abbildung 28: Horizont. Elemente, die oberhalb der Horizontlinie liegen, werden als zusammengehörig gesehen.

(3) Beleuchtung

Gliedere das Ganzfeld in beleuchtete und unbeleuchtete Regionen. Gruppieren Elemente, deren Beleuchtung sich derselben Lichtquelle zuordnen lassen, als zusammengehörig.

Wie beim Horizont hat sich unser Sehen auch an die natürliche Beleuchtung auf der Erde angepasst. Deshalb geht das visuelle System zunächst davon aus, dass das Gesichtsfeld nur durch eine einzige Strahlungsquelle beleuchtet wird – die Sonne oder den Mond – und sich

die Strahlungsquelle oberhalb der Beobachterin oder des Beobachters befindet (Abbildung 29). Dieses *natürliche Lichtschema* dient als Gliederungsheuristik. In der Stimulation sind die beiden Einflussfaktoren Beleuchtung und Reflexion noch konfundiert, erst im Laufe der visuellen Informationsverarbeitung werden sie mithilfe von Konstanzmechanismen getrennt (vgl. Kapitel VI). Diese machen das Sehen unabhängig von lokalen Lichtintensitäten und Wellenlängen, indem sie nicht die absoluten, sondern die relativen Werte im Gesichtsfeld berücksichtigen. Sie haben die Funktion, die Beleuchtungsveränderungen von den Objektveränderungen zu trennen.

Die Wahrnehmung geht zunächst davon aus, dass die Reflexionseigenschaften der Oberflächen über längere Zeiträume konstant bleiben, während sich die Beleuchtungsverhältnisse ändern. Diese Relation hat sich auch in den letzten 100 Jahren, in denen künstliche Lichtsituationen zunehmend die Alltagswirklichkeit bestimmen, nicht verschoben. Gemeinsam ist natürlichen wie künstlichen Beleuchtungssituationen eine Teilung des Gesichtsfeldes in beleuchtete und unbeleuchtete Regionen. Sie bilden jeweils eigenständige Regionen. Dem Schatten kommt für die Interpretation einer Lichtsituation eine besondere Bedeutung zu. Von einer gemeinsamen Lichtquelle angestrahlt zu werden oder in einem gemeinsamen Schatten zu liegen, kann Elemente auch bei einem größeren räumlichen Abstand und bei einer Verschiedenheit in anderen Merkmalen wie Form, Farbe oder Größe zusammenschließen (für Beispiele vgl. Casati, 2001).

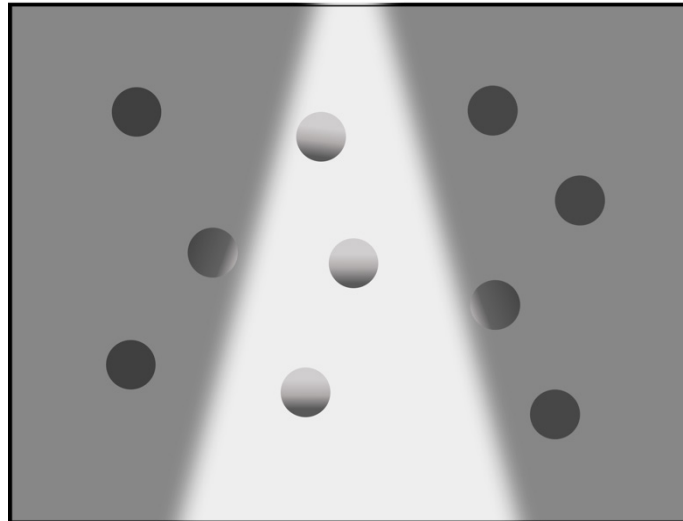


Abbildung 29: Beleuchtung. Objekte, die von einer gemeinsamen Lichtquelle angestrahlt werden, werden als zusammengehörig gesehen.

(4) Planität

Gliedere das Ganzfeld in Ebenen räumlicher Tiefe. Gruppiere Elemente, die auf einer gemeinsamen Tiefenebene liegen, als zusammengehörig.

Sowohl für die eigene Fortbewegung wie auch für die Planung von Handlungen ist es erforderlich, die Entfernung der Objekte zur Beobachterin oder zum Beobachter zu bestimmen. Das visuelle System betreibt erheblichen Aufwand, um die Entfernung von Objekten im Gesichtsfeld unabhängig von ihrer Größe korrekt einzuschätzen. Das Problem ist nicht leicht zu lösen, da im Netzhautbild die Informationen zu Entfernung und Größe nicht getrennt sind (vgl. Kapitel II). Objekte, die derselben *Tiefenebene* angehören, können oft mit einem vergleichbaren Aufwand und einem ähnlichen Bewegungsmuster erreicht werden. So ist zu entscheiden, ob ein Gegenstand sich in Reichweite unserer Arme befindet oder ob ein Abstand gering genug ist, um ein Hindernis zu überspringen. Planität ist vor allem im Zusammenhang mit der Bewegung von Objekten und der Eigenbewegung der Beobachterin oder des Beobachters ein wichtiger Gliederungsgesichtspunkt.

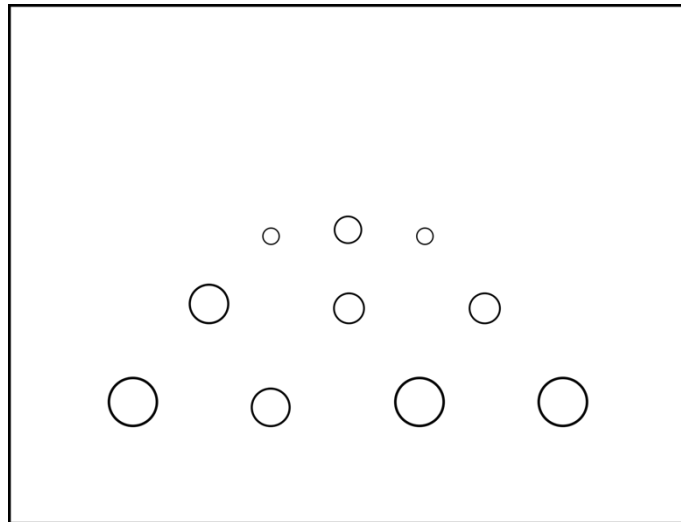


Abbildung 30: Planität. Elemente, die auf einer Tiefenebene liegen und damit anschaulich gleich weit von der Beobachterin oder dem Beobachter entfernt sind, werden als zusammengehörig gesehen.

(5) Topologie

Gliedere das Ganzfeld in eigenständige Regionen, die durch Grenzen markiert sind. Gruppiere Elemente, die einer gemeinsamen Region angehören, als zusammengehörig.

Im Gegensatz zur Zweiteilung des visuellen Feldes durch die Horizontlinie meint Topologie oder *gemeinsame Region* die Abgrenzung von lokalen Bereichen innerhalb des Gesichtsfeldes im Sinne eines Umschlossenseins (Abbildung 31). Die Grenzziehung kann durch eine eigenständige Kontur, aber auch durch einen Wechsel in der Textur oder in der Farbe einer sonst homogenen Region erfolgen (Palmer, 1992). In jedem Fall gehört die Grenze zur Region. Die Topologie ist eine Gliederungsebene zwischen den Objekten und dem Ganzfeld. Ein Beispiel: Im Gesichtsfeld befinden sich zwei Tische A und B, auf denen jeweils Objekte stehen. Die Objekte werden vorzugsweise so gruppiert, dass ihre Zugehörigkeit zu einer abgegrenzten Region (A oder B) erhalten bleibt. Der Tisch wird zum gemeinsamen Hintergrund für alle jene dreidimensionalen Figuren, die sich auf ihm befinden. Der Tisch selbst ist noch einmal Figur in Bezug auf

den Fußboden oder die Wand. Solche Hierarchisierungen des Figur-Grund-Verhältnisses können sich im Gesichtsfeld mehrfach wiederholen. Regionen können Bestandteil umfassenderer Regionen sein.

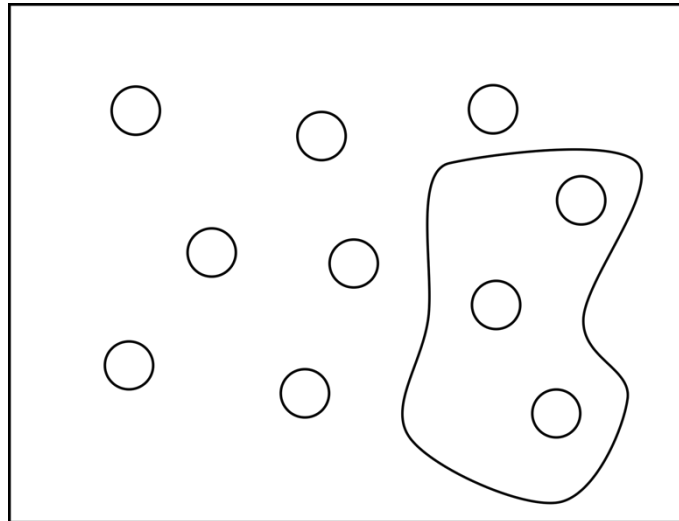


Abbildung 31: Topologie. Elemente, die von der Grenzlinie umschlossen sind, werden als zusammengehörig gesehen.

(6) Textur

Gliedere das Ganzfeld in Oberflächen mit unterschiedlicher visueller Textur. Gruppieren Oberflächen, die einen gemeinsamen Texturgradienten aufweisen, als zusammengehörig.

Zentrale Beschreibungsmerkmale für visuelle Texturen sind die Größe, Dichte und Orientierung ihrer Elemente (vgl. Kapitel II). Bei den meisten natürlichen Oberflächen ist die Gleichartigkeit der Textur Elemente nicht vollständig, sondern statistisch. Blickt eine Beobachterin oder ein Beobachter auf eine texturierte Oberfläche, verändern sich einerseits Größe und Dichte in Abhängigkeit von der Entfernung, andererseits Dichte und Orientierung in Bezug auf den Aufsichtswinkel. Die Größe der Elemente nimmt mit der Entfernung stetig ab, ihre Dichte nimmt zu. Auf diese Weise enthält die Textur Hinweise über die Distanz zu Oberflächen. Der Verlauf der Texturdichte und die Orientierung der Elemente

wiederum geben Auskunft über die Perspektive. Die Texturgradienten bestimmen die Neigung und Entfernung von Oberflächen und helfen bei der Lokalisation des Standortes. Texturgradienten grenzen Oberflächen voneinander ab und führen – auch bei einer Identität ihrer Elemente – zu einer Gliederung des Ganzfeldes (Abbildung 32). Die Bedeutung dieser Gliederung besteht darin, dass sie unmittelbare Korrespondenzen zur Abgrenzung von Oberflächen in der Außenwelt erlaubt (für einen aktuellen Überblick vgl. Rosenholtz, 2015).

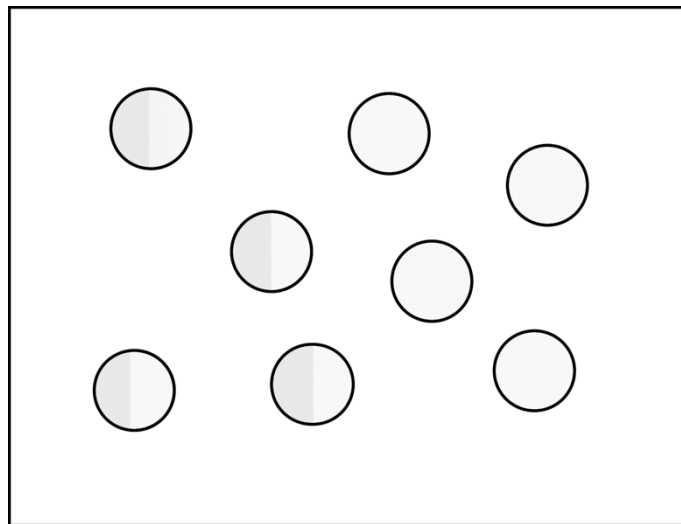


Abbildung 32: Textur. Dichteunterschiede in den Punktmustern führen zur Abgrenzung von jeweils zwei Flächen innerhalb der Kreise im linken Teil der Abbildung.

(7) Einfachheit und Prägnanz

Gliedere das Ganzfeld möglichst einfach. Gruppiere Teile so zu einem Ganzen, dass möglichst gute oder prägnante Gestalten entstehen.

Neben den einzelnen Gruppierungsheuristiken gibt es die weitergehende Vorstellung, dass die Gesamtorganisation des visuellen Feldes nach einem Maximum an Ordnung oder Prägnanz strebt. Gleichzeitig beschreibt die *Prägnanztendenz* oder das *Gesetz der guten Gestalt* auch die Zusammenfassung von Elementen zu einer Ganzheit (Abbildung 33).

Ausgangspunkt war Wertheimers Beobachtung, dass Proband*innen in Wahrnehmungsexperimenten bei bestimmten Reizen systematisch andere Phänomene berichteten, als von den metrischen Eigenschaften her zu erwarten wäre. Dabei wiesen die Abweichungen immer in Richtung auf ein „gutes“ oder „ausgezeichnetes“ Wahrnehmungsergebnis.

Das ursprüngliche Prägnanzkonzept wurde von Rausch (1974, 1982) differenziert. Dieser nennt sieben Aspekte, die als eine Binnengliederung der Prägnanz zu verstehen sind: Gesetzmäßigkeit, Eigenständigkeit, Integrität, Einfachheit der Strukturierung, Ausdrucksfülle und Bedeutungsfülle. Jeder Aspekt ist bipolar konzipiert (z. B. Gesetzmäßigkeit versus Zufälligkeit). Es resultiert ein *Prägnanzprofil*, das die anschauliche Wirkung einer beliebigen Vorlage bei der Betrachterin oder dem Betrachter beschreibt. Kaum ein Konzept der Gestalttheorie hat so viele Missverständnisse hervorgerufen, wie das Prägnanzkonzept (vgl. Hüppe, 1984). Spätere Untersuchungen zum Thema Prägnanz bemühten sich, wie bei Symmetrie, um eine Formalisierung und Quantifizierung, denn der von den Phänomenen ausgehende Ansatz der Gestalttheoretiker ermöglichte kaum konkrete Vorhersagen. Heute wird sowohl *Prägnanz* wie auch *gute Gestalt* oft mit *Einfachheit* gleichgesetzt (vgl. deshalb auch die Forschungsergebnisse zu Einfachheit, z. B. van der Helm, 2015a). Auch beim Konzept der Prägnanz kamen – unter Anwendung des Redundanzkonzeptes – die ersten Vorschläge zur Formalisierung von der Informationstheorie (z. B. Attneave, 1954). Prägnanz wird als Invarianz gegenüber Transformationen beschrieben (vgl. z. B. Garner, 1974; Palmer, 1982; Zimmer, 1984; Kebeck, 1991).

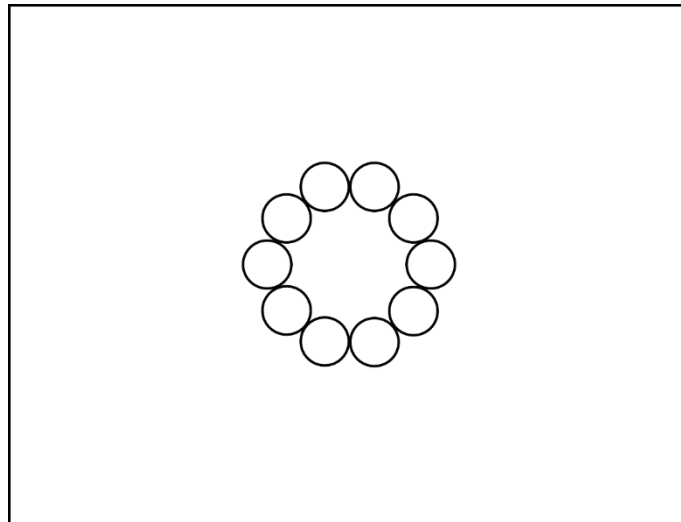


Abbildung 33: Einfachheit und Prägnanz. Die einzelnen Kreise werden zu einem größeren Kreis zusammengefasst.

(8) Nähe

Gruppieren Sie Elemente so, dass Elemente mit relativ geringer Entfernung zueinander zusammengefasst werden und dass möglichst dichte, voneinander abgegrenzte Gruppen entstehen.

Die neuroanatomische Basis für die Heuristik der Nähe bildet die *retinotopische Organisation* des visuellen Systems (vgl. Kapitel II). Sobald eine Ausgrenzung von Figuren erfolgt, kann die Heuristik angewendet werden (Abbildung 34). Insofern ist es eine sehr leistungsstarke Heuristik. Zwei Elemente, die sich nicht gegenseitig verdecken, weisen immer auch eine messbare Distanz zueinander auf. „Nähe“ ist nicht nur universell, sondern auch leicht quantifizierbar. Entscheidend ist die jeweils kürzeste Distanz zu einem Nachbargebilde. Die Wirkung dieser Distanz lässt sich mithilfe einer Potenzfunktion beschreiben (Stevens, 1957; Luce, 2002). Im Wettstreit mit anderen Heuristiken steht „Nähe“ in der Hierarchie allerdings weit unten. Anders formuliert: Sie ist fast immer wirksam, aber nie dominant.

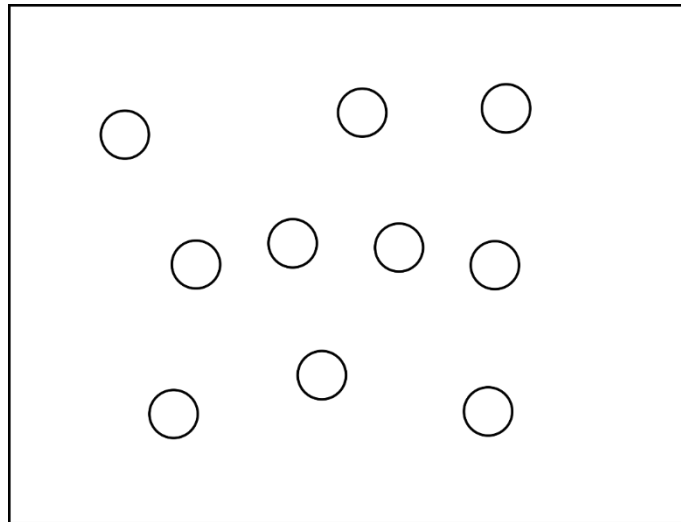


Abbildung 34: Nähe. Die vier mittleren Kreise werden als zusammengehörig gesehen.

(9) Gleichheit und Ähnlichkeit

Gruppierere Elemente so, dass gleiche oder ähnliche Elemente zusammengefasst werden und dass in sich einheitliche Gruppen entstehen.

Wie bereits erwähnt, hat diese Heuristik eine Sonderstellung, weil die übrigen Heuristiken erst wirksam werden, wenn durch Qualitätssprünge (als Zeichen von Inhomogenität) Teile des Sehfeldes ausgegliedert wurden (vgl. Abbildung 35). Das Kriterium der Gleichheit oder Ähnlichkeit kann sich auf unterschiedliche Stimuluseigenschaften wie Form, Helligkeit, Farbe, Orientierung oder Größe beziehen. Für die Wahrnehmungsforschung stellt sich die Frage einer mehrdimensionalen Skalierung: Welches mathematische Modell kann die Gewichtung der Merkmale vorhersagen?

Das Konzept der Ähnlichkeit hat für die Frage des Verhältnisses von physikalischer und phänomenaler Welt einen zentralen Stellenwert. Sein Erklärungswert allerdings ist heftig umstritten. Deshalb ist es notwendig, sich bei der Beschreibung von Ähnlichkeiten auf einen Messvorgang zu beziehen. In der empirischen Forschung lassen sich mindestens zwei

Herangehensweisen unterscheiden: 1. Ähnlichkeit kann indirekt gemessen werden. Die zugrunde liegende Annahme ist, dass zwei Objekte, die einander ähnlich sind, eher verwechselt werden. 2. Ähnlichkeit kann direkt gemessen werden. Demnach sollten Personen in der Lage sein, entweder auf einer Skala das Ausmaß an Ähnlichkeit oder Unähnlichkeit anzugeben oder verschiedene Vergleichsreize zu einem Standardreiz mithilfe einer „Größer-Relation“ zu beurteilen. Die beiden Vorgehensweisen führen nur selten zu identischen Ergebnissen. Die Forschung geht deshalb zunehmend davon aus, dass Ähnlichkeit kein einheitliches Konzept ist und dass es sinnvoll ist, verschiedene Formen von Ähnlichkeit zu unterscheiden (Palmer, 1999).

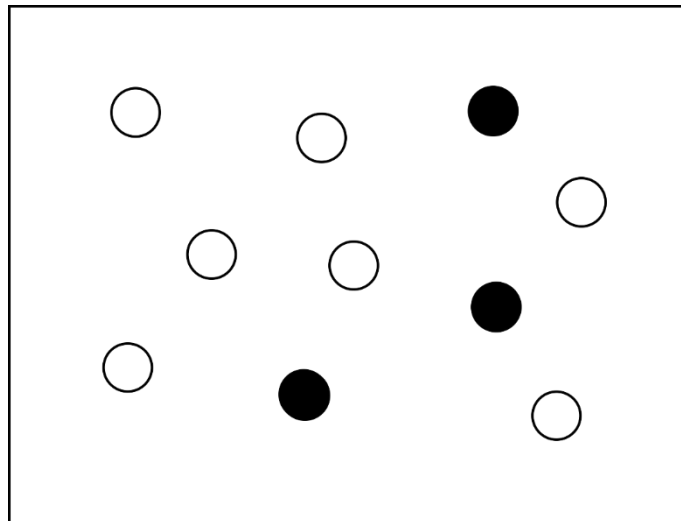


Abbildung 35: Gleichheit und Ähnlichkeit. Die drei gefüllten Kreise werden als zusammengehörig gesehen.

(10) Kollinearität

Gruppieren Sie Elemente, die gemeinsam glatt verlaufende Linien ergeben, als zusammengehörig.

Für diese Heuristik finden sich unterschiedliche Bezeichnungen: durchgehende Kurve, glatter Verlauf, gute Fortsetzung, Kollinearität. Veranschaulicht wird sie, indem an einer Stelle eines Linienverlaufs zwei

Fortsetzungen gezeigt werden (Abbildung 36). Die Frage lautet: Welche der Alternativen wird präferiert? „Gute Fortsetzung“ ist eine Heuristik, um an Kreuzungs- oder Berührungspunkten Eindeutigkeit zu erzeugen. Wie „Nähe“ für die Gruppierung von Punktmustern ist „Gute Fortsetzung“ bei Linienverläufen universell. Da das visuelle System prospektiv arbeitet, bestimmt Kollinearität die Richtung der nachfolgenden Blickbewegungen. Sie soll den wahrscheinlichsten Verlauf vorwegnehmen. Die Frage bleibt: Was determiniert eine gute Fortsetzung? Korrespondiert sie mit dem minimalen Richtungswechsel zwischen den Liniensegmenten oder mit der ungebrochenen Fortsetzung der Krümmung des ersten Liniensegments durch das zweite? In neueren Untersuchungen wird zur Klärung dieser Frage „Gute Fortsetzung“ oft im Zusammenhang mit zwei anderen Phänomenen untersucht: der Konturintegration und der visuellen Interpolation. Bei Experimenten zur Konturintegration werden Elemente, die eine gemeinsame, unterbrochene Kontur bilden, in einer ansonsten zufälligen Anordnung gezeigt (vgl. z. B. die Arbeiten von Field, Hayes und Hess, 1993; Fantoni und Gerbino, 2003; Hess, May und Dumoulin, 2015). Bei visueller Interpolation muss eine Kontur zwischen zwei Segmenten ergänzt werden und es wird nach dem wahrscheinlichsten Verlauf gefragt (vgl. z. B. Kellman und Shipley, 1991; Wouterlood und Bodelie, 1992; Kellman, Garrigan, Kalar und Shipley, 2010). Im Hinblick auf die Frage, ob der Richtungswechsel oder die Fortsetzung der Krümmung entscheidend ist, sind die bisherigen Ergebnisse nicht eindeutig.

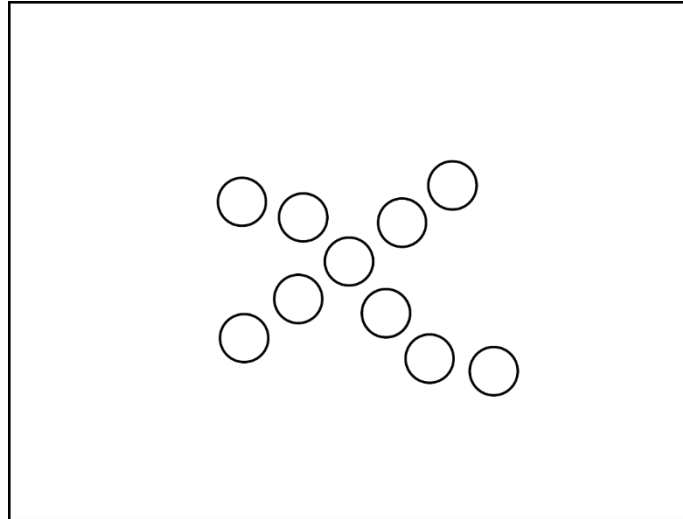


Abbildung 36: Kollinearität. Die Kreise werden in zwei Gruppen zusammengefasst, die jeweils eine glatt verlaufende Linie bilden.

(11) Geschlossenheit

Gruppierere Elemente, die geschlossene Teilverläufe bilden, als zusammengehörig.

Die Heuristik ist bedeutsam im Zusammenhang mit der Bildung von zwei- und dreidimensionalen Figuren (vgl. Abbildung 37). Nach Metzger (1967, S. 708) gilt: „Die Zusammenschlüsse von Linien im Sinne des Faktors der Geschlossenheit haben die Besonderheit, dass sie zur Ausbildung von Flächenfiguren führen, wobei die Linien den Charakter von Rändern oder Konturen einnehmen.“ Hierfür ist es nicht notwendig, dass sich die einzelnen Liniensegmente berühren: „Kurvenstücke oder Winkel, die einander benachbart sind, ohne sich zu berühren, schließen sich ebenfalls merklich bevorzugt mit ihren konkaven Seiten nach innen zu Paaren zusammen.“ Wichtig ist dieser Zusatz im Zusammenhang mit der Bildung von amodalen Figuren (vgl. Kapitel V). Gemeinsam mit „Guter Fortsetzung“ führt er dazu, dass es selbst bei großen Leerstellen, bei denen es keine Reizsprünge und damit keine Stimulation gibt, zur Wahrnehmung von Konturen kommt. Das visuelle System

komplementiert die Stimulation. Bei dieser Ergänzung hilft die Heuristik der Geschlossenheit. Sie berücksichtigt, dass in der Umgebungswahrnehmung Objekte oft teilweise verdeckt sind und die Netzhautprojektion in der Folge unvollständige Verläufe aufweist.

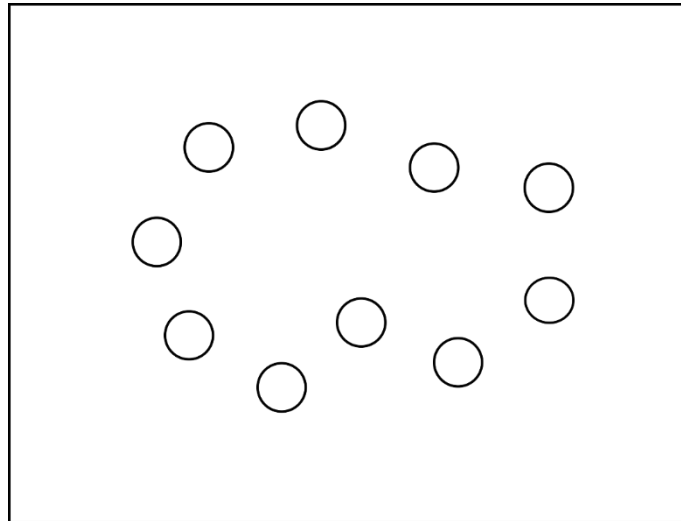


Abbildung 37: Geschlossenheit. Die Kreise werden als eine unregelmäßige, geschlossene Figur gesehen.

(12) Aufgehen ohne Rest

Gruppieren Sie Elemente so, dass die resultierenden Gruppen eine Gliederung in gleich große und gleiche oder nach demselben Prinzip konstruierte Untergruppen erlauben.

Verschiedene Fälle sind zu unterscheiden. So gliedern sich diskrete Elemente, die in einer Reihe angeordnet sind und sich aufgrund anderer Heuristiken, wie zum Beispiel Gleichartigkeit, zu Gruppen zusammenschließen, bevorzugt so, dass an den Enden dieser Reihen keine ungepaarten Einzelglieder übrigbleiben (vgl. Abbildung 38). Bei kontinuierlichen Gebilden führt der Restfaktor – gemeinsam mit Geschlossenheit – zu möglichst einfachen Figuren. Allgemein erfolgen die Gliederungen und Zusammenfassungen so, dass im Ganzfeld keine oder möglichst wenige „Anhängsel“ oder „Lücken“ entstehen. Wobei die Frage,

ob etwas anschaulich als Anhängsel oder Lücke aufgefasst wird, abhängig davon ist, was gerade zuvor als Ganzes verwirklicht worden ist. Etwas, was Anhängsel war, kann zur Lücke werden. Der Vorgang entspricht einem Figur-Grund-Wechsel.

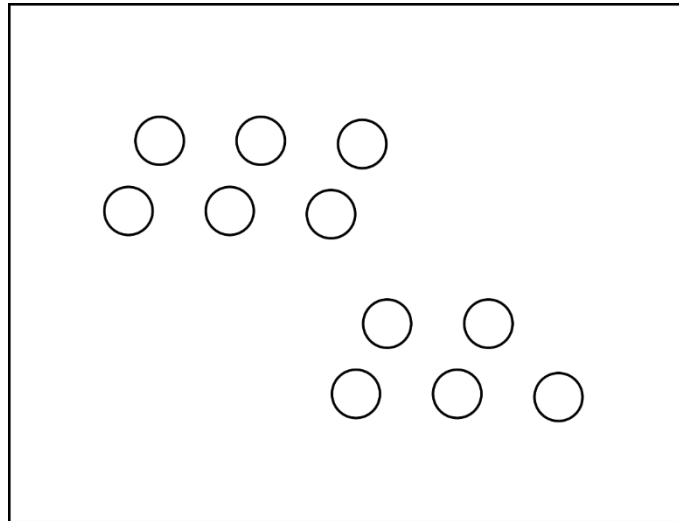


Abbildung 38: Aufgehen ohne Rest. Der Kreis in der unteren rechten Reihe bildet eine Einheit mit den benachbarten Kreisen und wird nicht als eine eigenständige Einheit gesehen, obwohl sein Abstand zu den benachbarten Elementen größer ist.

(13) Symmetrie

Gruppierere Elemente, deren Anordnung in Teilen oder als Ganzes symmetrisch ist, als zusammengehörig.

Im Zusammenhang mit der Mathematisierung der Gestaltgesetze und der Beschreibung der Gestaltgüte durch das Konzept der Transformationsinvarianz sind auch die Symmetrieeigenschaften der Stimulation in den Vordergrund des Forschungsinteresses gerückt. Bei Wertheimers Überlegungen spielt Symmetrie zwar eine wichtige Rolle, aber sie wird nicht als eigenständiges Gruppierungsgesetz aufgeführt. Nach Wertheimer kann Symmetrie nur eine Eigenschaft von Ganzen und nicht von Teilen sein: „Dabei spielt eine bestimmte Art prägnanter „Ganzeigenschaften“ eine ausgezeichnete Rolle: Eigenschaften wie

„Geschlossenheit“, „Symmetrie“, „inneres Gleichgewicht“ (dabei muß klar sein, daß z. B. Symmetrie durchaus nicht einfach eine „Gleichheit“ von Teilen ist, sondern logisch richtig nur vom Ganzen her, als Ganzeigenschaft gefaßt werden kann).“ (Wertheimer 1923, S. 325). Gruppentheoretischen Überlegungen folgend, wird heute zwischen „globalen“ Symmetrien, die die Gesamtfigur betreffen – und den Anforderungen von Wertheimer entsprechen – und „lokalen“ Symmetrien – die im mathematischen Sinne keine Gruppe bilden – unterschieden. Die seit Ernst Machs (1886) grundlegender Arbeit am häufigsten untersuchten Symmetriearten sind Spiegel- oder Axialsymmetrien. Sie entstehen, wenn eine Figur durch senkrechte Spiegelung an einer ihrer Symmetrieachsen auf sich selbst abgebildet wird (vgl. Abbildung 39). Neben den Spiegelsymmetrien wurden später vor allem Rotations- und Translationssymmetrien untersucht (vgl. z. B. Barlow und Reeves, 1979; Zimmer, 1984; Wagemans, 1993, 1995; für einen aktuellen Überblick van der Helm, 2015b; für Beispiele unterschiedlicher Arten von Symmetrie in der Natur vgl. Weil, 1952; für Beispiele in der Wissenschaft und in der Kunst vgl. Shubnikov und Koptsik, 1974).

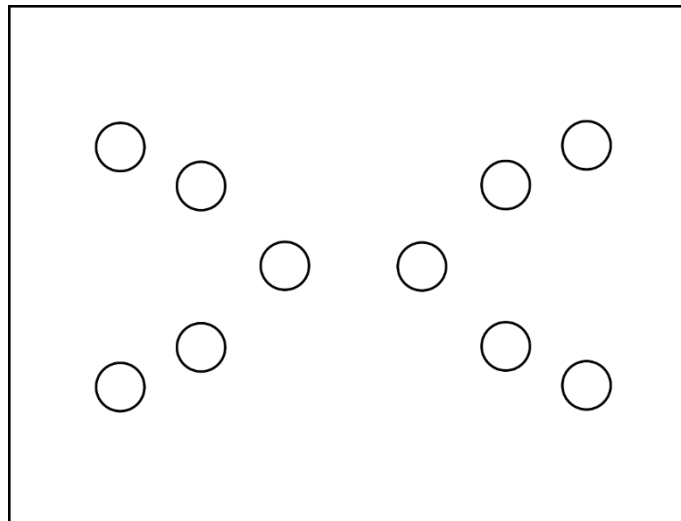


Abbildung 39: Symmetrie. Gesehen wird ein Muster aus Kreisen, das an einer senkrechten Mittelachse gespiegelt ist.

(14) Konnektivität

Gruppieren Sie Elemente, die miteinander durch ein weiteres Element verbunden sind, auch dann als zusammengehörig, wenn sie ansonsten selbständig sind.

Die Unterscheidung zwischen dem Ganzen und seinen Teilen bedarf in Bezug auf Fragen der Gliederung und Gruppierung einer Differenzierung. Eigenständige Ganze, die durch ein weiteres Element verbunden sind, können als Teile eines Ganzen gesehen werden. Sie werden in der Folge eher als zusammengehörig erlebt (vgl. Abbildung 40). Das verbindende Element muss keine gemeinsamen Merkmale mit den zu verbindenden Elementen aufweisen und die beiden eigenständigen Teile müssen keiner gemeinsamen Region angehören (Palmer und Rock, 1994).

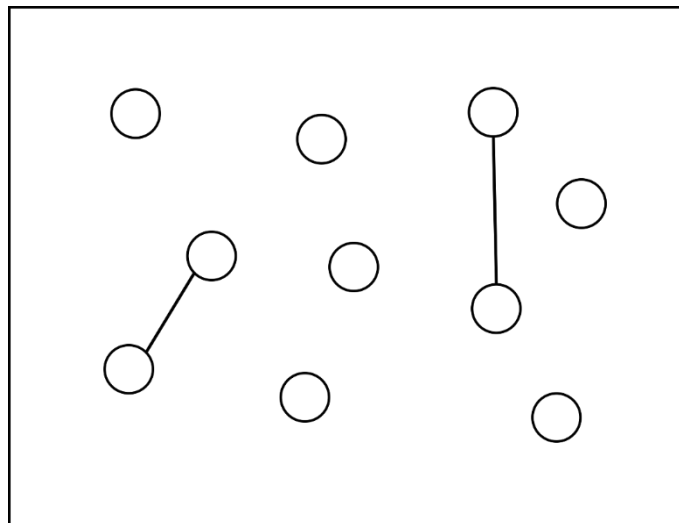


Abbildung 40: Konnektivität. Die Kreise, die durch eine Linie miteinander verbunden sind, werden als zusammengehörig gesehen.

Evidenz für eine Eigenständigkeit der Konnektivität kommt nicht nur aus experimentellen Untersuchungen, sondern findet sich auch bei Patienten mit neurologischen Störungen. Beim Balint-Syndrom etwa können Patienten nicht mehr als ein einziges Objekt zu einem Zeitpunkt wahrnehmen. Werden ihnen zwei separate Kreise auf einem Computermonitor dargeboten, dann berichten sie, nur einen Kreis zu

sehen. Werden die beiden Kreise jedoch durch ein drittes Element verbunden, dann werden beide Kreise erkannt (Humphreys und Riddoch, 1993).

(15) Gemeinsames Schicksal

Gruppiere Elemente, die eine gleichartige Veränderung erfahren, als zusammengehörig und bilde Ganze, die in sich beständig sind, auch wenn die Teile ihre Eigenschaften ändern.

Auf den Netzhäuten der beiden Augen wird, bedingt durch die relative Bewegung der Beobachterin oder des Beobachters zu den Oberflächen der Außenwelt und ihren Texturen, ein Fließmuster erzeugt (vgl. Abbildung 41). Sowohl bei Objektbewegungen wie bei Eigenbewegungen verändern sich die optischen Fließmuster (Hinweise, dass sich Punktmuster, die auf einen gemeinsamen Fluchtpunkt konvergieren oder eine gemeinsame Kreisbewegung ausführen, eher zusammenschließen, liefern u. a. die Ergebnisse der Experimente von Börjesson und Ahlström, 1993; Ahlström, 1995). In Abhängigkeit von der Form des Objektes und der Art der Bewegung bleiben jedoch einzelne Verhältnisse invariant, oder sie „erleiden ein gemeinsames Schicksal“. Hieraus zieht das visuelle System Rückschlüsse über die Art des Ereignisses, etwa über Veränderungen in der Form der Objekte oder ihre Entfernung, über die Richtung der Bewegung oder ihre Geschwindigkeit. Es verwundert nicht, dass die Heuristik des „gemeinsamen Schicksals“ zu einer der wichtigsten Variablen bei der Programmierung von Computerspielen oder bei der Entwicklung sehender Maschinen geworden ist.

Die Heuristik ist sehr wirkmächtig und nimmt in der Hierarchie der Heuristiken einen oberen Platz ein. Bisher finden sich aber kaum empirische Untersuchungen, die eine Quantifizierung dieser Wirkung erlauben. Bei Wertheimer war das Gesetz des gemeinsamen Schicksals

oder übereinstimmenden Verhaltens auf die Geschwindigkeit oder Richtung einer Bewegung begrenzt. Heute wird häufig auch die gemeinsame Veränderung anderer Merkmale der Stimulation wie Helligkeit, Farbe, Form oder Orientierung einbezogen (vgl. auch Levinthal und Frankoneri, 2011; speziell für Helligkeitsänderungen vgl. Sekuler und Bennett, 2001).

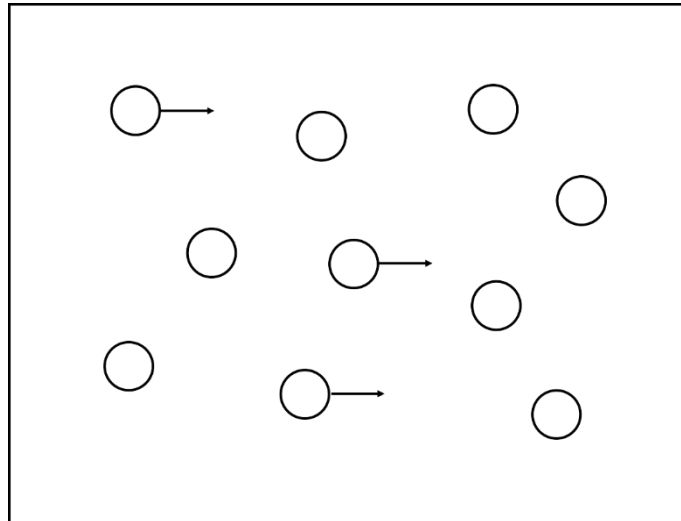


Abbildung 41: Gemeinsames Schicksal. Die Pfeile symbolisieren Bewegungsvektoren; Kreise, die sich in gleicher Weise bewegen, werden als zusammengehörig gesehen.

(16) Simultanität

Gruppierere Elemente, die in einer Stimulation simultan auftreten oder simultan ihre Eigenschaften verändern, eher als zusammengehörig als Elemente, die sukzessiv auftreten oder sukzessiv ihre Eigenschaften verändern.

„Gemeinsames Schicksal“ beschreibt, wie Zusammenfassungen bei einer Bewegung oder bei einer Veränderung der Leuchtdichte entstehen. Aber auch Elemente, die kein gemeinsames Schicksal erleiden, können aufgrund der zeitlichen Simultanität ihres Auftretens als zusammengehörig empfunden werden (vgl. Abbildung 42). Zumindest

sprechen hierfür die Ergebnisse einer Reihe von Experimenten mit Punktmustern (vgl. z. B. Alais, Blake und Lee, 1998 und Lee und Blake, 1999). Ausgehend von einer Matrix, bei der die Farbe der Punkte zufällig über die Zeit variierte, konnte gezeigt werden, dass, wenn bei einer Untergruppe der Farbwechsel synchron erfolgte, die Elemente eine Gruppe bildeten und als eigenständige Figur wahrnehmbar waren. Diese Untergliederung trat selbst dann auf, wenn sich die Punkte in verschiedene Richtungen bewegten und sich nach dem Gesetz des gemeinsamen Schicksals hätten zusammenschließen müssen.

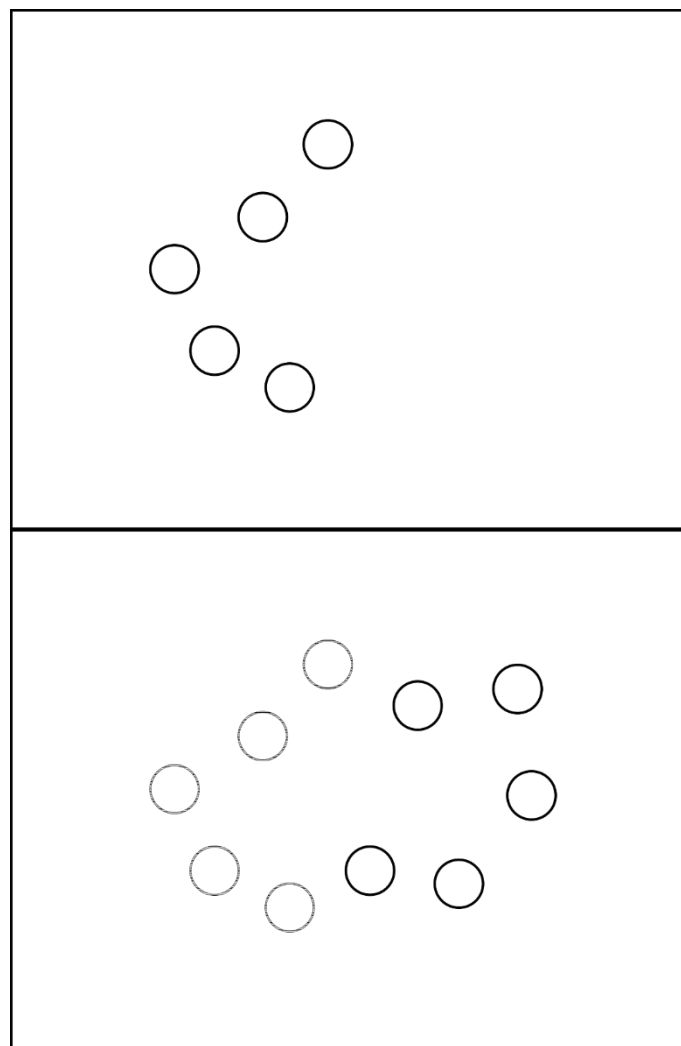


Abbildung 42: Simultanität. Gezeigt wird ein zeitlicher Verlauf. Die fünf Kreise, die im unteren Teil der Abbildung hinzukommen, werden wegen ihres gemeinsamen Auftretens als zusammengehörig gesehen.

Ob die rein zeitliche Synchronizität eines Ereignisses wirklich ein universeller Faktor ist oder ob eher die gemeinsame *zeitliche Struktur* von Ereignissen der übergeordnete Gesichtspunkt ist, bleibt derzeit noch umstritten (eine Argumentation für die Eigenständigkeit der Simultanität findet sich bei Palmer, 1999; kritisch äußern sich dagegen z. B. Farid und Adelson, 2001; Farid, 2002; Guttman, Gilroy und Blake, 2007).

(17) Hysterese

Gliedere das Ganzfeld so, dass vorherige Gliederungen beibehalten werden, wenn sich die Stimulation kontinuierlich oder in kleinen Schritten verändert.

Hysterese meint das Zurückbleiben einer Wirkung hinter der sie verursachenden, zeitlich veränderlichen physikalischen Größe. Ein bekanntes Beispiel ist das Magnetisierungsverhalten von ferromagnetischen Stoffen: Wird ein Eisenstück in ein äußeres Magnetfeld eingebracht, so baut sich die Magnetisierung verzögert auf. Wird das Eisenstück aus dem Feld entfernt oder wird das Magnetfeld abgeschaltet, so bleibt eine Restmagnetisierung zurück, die Remanenz. Diese wird erst durch ein entgegengesetzt gepoltes äußeres Feld aufgehoben.

In der Wahrnehmungsforschung beschreibt Hysterese die Senkung der Wahrnehmungsschwelle aufgrund einer kurz zuvor erfolgten Stimulation (vgl. Abbildung 43). Neurophysiologisch entspricht dies dem Prozess der *Bahnung*, bei dem es zu einer wenige Sekunden bis Minuten lang anhaltenden Verstärkung der Neurotransmission nach einer vorausgegangenen Aktivität kommt. Aus Experimenten, bei denen visuelles Material sehr kurzzeitig dargeboten wird, wissen wir, dass für eine Bahnung keine bewusste Wahrnehmung notwendig ist (sublimales Priming). Hysterese ist universeller als der von Wertheimer ursprünglich vorgeschlagene Faktor der „objektiven Einstellung“. Hysterese ist

Ausdruck einer konservativen Voreinstellung des visuellen Systems. Erwartet wird das, was gerade gesehen wurde. Eine unverzichtbare Heuristik, um die Beständigkeit der visuellen Welt zu garantieren. Alleine schon deshalb, weil etwa alle 4 bis 6 Sekunden ein Lidschlag erfolgt und die hiermit verbundene Dunkelphase etwa 400 bis 600 Millisekunden dauert.

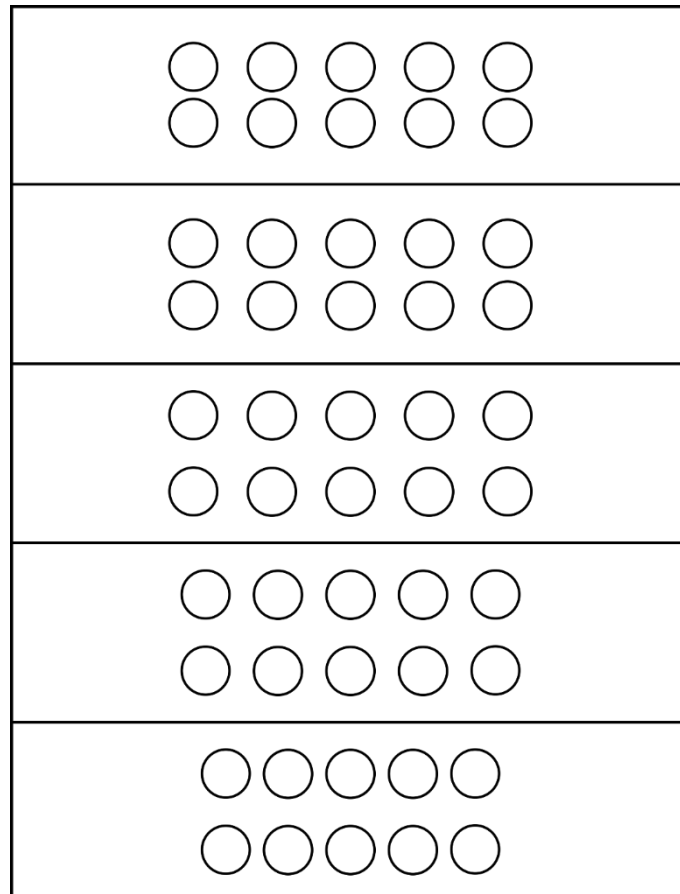


Abbildung 43: Hysterese. Die Abbildung kann von oben nach unten oder umgekehrt betrachtet werden. Im ersten Fall gliedert sich die mittlere Anordnung eher in Spalten, im zweiten Fall eher in Zeilen.

(18) Fixationswechsel

Nutze die Fixationswechsel während der Augenbewegungen zur Untergliederung des Ganzfeldes.

Typisch für die Wahrnehmung von natürlichen Ganzfeldern ist, dass nur ein kleiner Ausschnitt foveal abgebildet werden kann. Für die Beobachterin oder den Beobachter ist es unerlässlich, einzelne Regionen in aufeinanderfolgenden Fixationen zu fokussieren. Dies führt zu einer Untergliederung des Wahrnehmungsfeldes in Abhängigkeit von den Fixationswechseln (vgl. Abbildung 44).

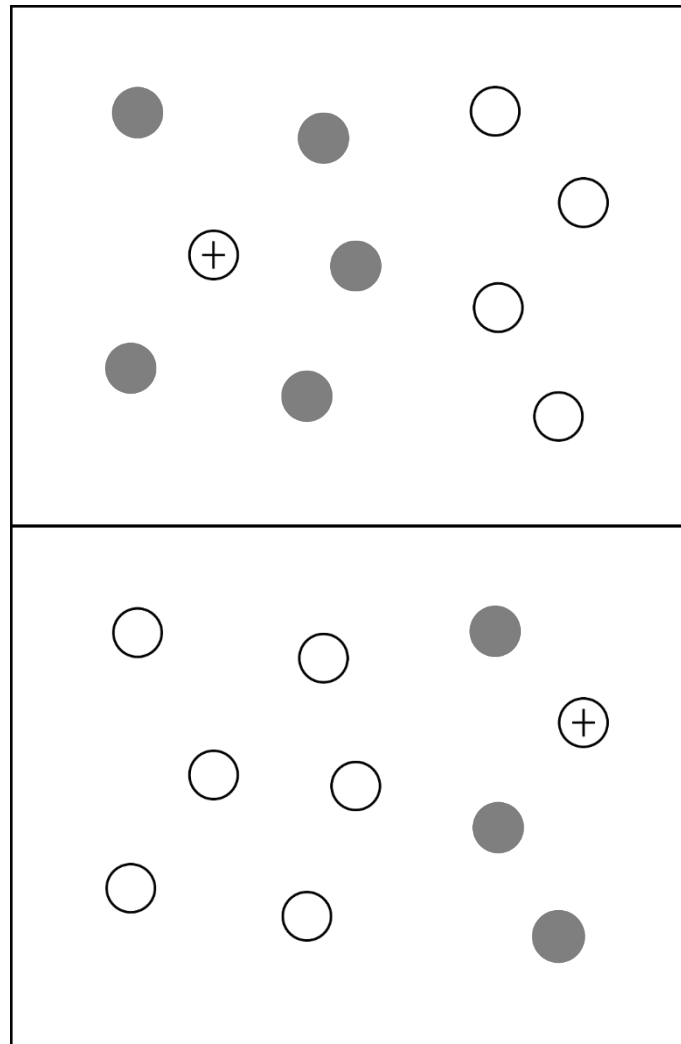


Abbildung 44: Fixationswechsel. Die Fixationskreuze zeigen an, auf welchen Kreis der Blick aktuell fokussiert. Mit dem Fixationswechsel werden die Kreise anders gruppiert.

Bewusste Wahrnehmung setzt eine Fixation voraus. Die Pfade der Augenbewegungen werden von zwei Faktoren bestimmt: dem Ziel der Informationssuche und den Merkmalen der Stimulation. Fast immer ist

das Bewegungsmuster Ergebnis ihres engen Wechselspiels (vgl. Kapitel V). An den Fixationswechslern zeigt sich das komplexe Zusammenspiel zwischen den objektbezogenen und den subjektbezogenen Gliederungsfaktoren.

(19) Eigenbewegung

Nutze die Eigenbewegung der Beobachterin oder des Beobachters für Gliederungen und Zusammenfassungen höherer Ordnung. Reorganisiere bestehende Gliederungen und Zusammenfassungen unter Einbeziehung der Eigenbewegung.

Die Eigenbewegungen umfassen Rumpf-, Kopf- und Augenbewegungen. Einerseits erzeugen sie bei der Beobachterin oder dem Beobachter systematische Veränderungen der Fließmuster auf den Retinae, andererseits steuern sie die Zusammenfassungen und Gliederungen innerhalb der optischen Fließmuster (vgl. Abbildung 45). Auf diese Weise reorganisieren sie bestehende Ordnungen. Die Geschwindigkeit und die Richtung der Eigenbewegung bildet eine Folie für die Einschätzung von Veränderungen in der Stimulation. Von besonderer Bedeutung sind Invarianzen, die die Eigenbewegung mit der Tiefenstruktur der Umwelt verbinden und für die eigene Handlungsplanung relevant sind. Beispiele für solche anspruchsvollen Aufgaben sind die Entdeckung eines Kollisionskurses und die Berechnung eines Ausweichkurses. Hierzu muss aus den Veränderungen der Fließmuster nicht nur eine Annäherungsrate und das Zeitintervall bis zur Kollision berechnet werden (time-to-collision), sondern auch eingeschätzt werden, inwieweit die Annäherung des Objektes in Abhängigkeit von der Eigenbewegung der Beobachterin oder des Beobachters symmetrisch erfolgt. Asymmetrische Veränderungen sprechen dafür, dass die Kollision vermieden werden kann.

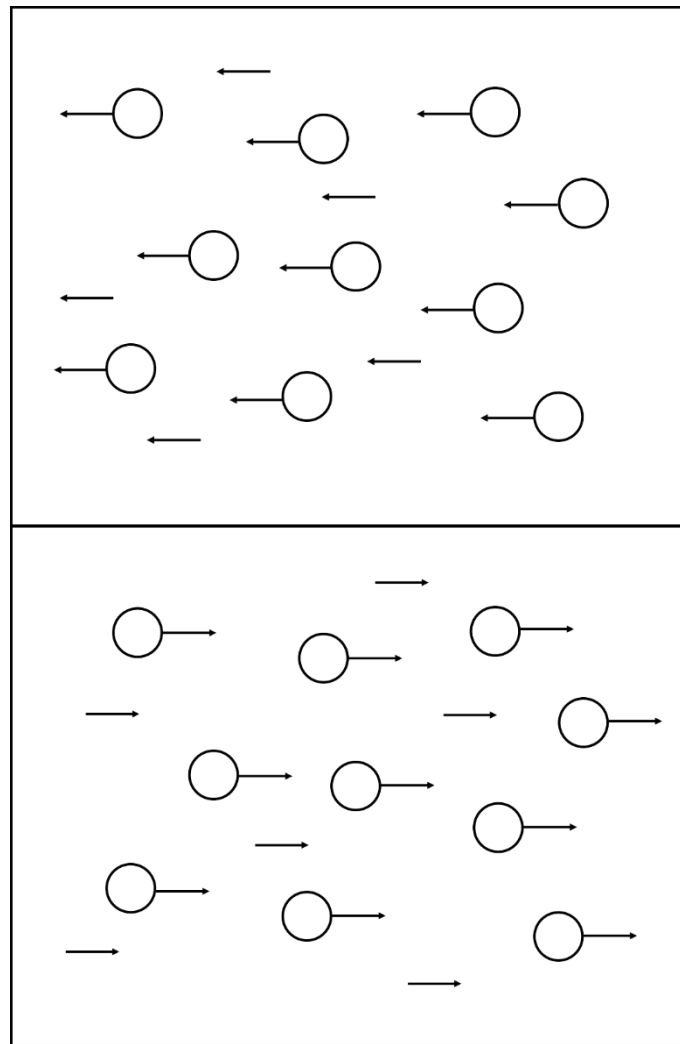


Abbildung 45: Eigenbewegung. Der obere Teil der Abbildung zeigt das Fließmuster einer Bewegung der Beobachterin oder des Beobachters nach rechts, der untere Teil der Abbildung nach links.

(20) Intention

Erlaube intentionalen Suchprozessen, bestehende autochthone Gliederungen und Gruppierungen zu revidieren.

Je komplexer und alltagsnäher die Stimulation ist, umso bedeutsamer werden hypothesengeleitete Gliederungsprozesse. Die Segmentierung einer Szene setzt eine Auffassungsabsicht bereits voraus. In Abhängigkeit von der Auffassungsabsicht zeigen sich systematische Unterschiede in den Augenbewegungen. Ihr Muster lässt sich nur vorhersagen, wenn auch die

Intentionen der Beobachterin oder des Beobachters einbezogen werden (vgl. Abbildung 46). Metzger (1967, S. 713), der den Begriff „Auffassungsabsicht“ als Zusammenfassung aller willentlichen Prozesse verwendet, ging noch davon aus, dass sie „ohne Einfluß auf die primäre Feldgliederung durch Qualitätssprünge“ bleibt. Heute wissen wir, wie eng das Zusammenspiel von datengesteuerter und hypothesengesteuerter Verarbeitung ist (vgl. Kapitel II und III). Zumindest indirekt, über die Fixationswechsel, können die Gliederungen und Zusammenschlüsse durch die Intentionen beeinflusst werden (für die Wirkung von „top down“ Prozessen auf Gliederungen und Gruppierungen vgl. auch Beck und Palmer, 2002).



Abbildung 46: Intention. Im Rahmen der intentionalen Suche kann das Ganzfeld neu gegliedert und können die Elemente neu gruppiert werden.

Ein Beispiel für den weitreichenden Einfluss der Intention auf Gliederungs- und Gruppierungsprozesse sind *Vexierbilder*. Suchbilder, die eine Figur enthalten, die nicht auf den ersten Blick erkennbar ist. Sie ist zwar vollständig vorhanden, aber sie bleibt zunächst unsichtbar. Erst der intentionale Blick macht die Figur sichtbar. Die automatischen Gruppierungsprozesse erlauben zwei Interpretationen derselben Stimulation, aber die beiden sind nicht gleich wahrscheinlich. Erhält das Bild jedoch einen Titel, der die zweite Deutung nahelegt, so wird die

zunächst favorisierte Gliederung aufgegeben, die versteckte Figur entdeckt. Schließlich kann die Betrachterin oder der Betrachter zwischen den beiden Versionen wechseln, gleichzeitig gesehen werden können sie nicht.

Kapitel V: Phänomene und Heuristiken

Das Wort „Phänomen“ bezeichnet im Altgriechischen ein mit den Sinnen wahrnehmbares Ereignis. Das den Sinnen Erscheinende steht im Gegensatz zum „Noumenon“, dem in Begriffen Gedachten. Es hat einen unmittelbaren Zusammenhang zur bewussten Wahrnehmung und damit zum Begriff des Perzeptes. Eine *phänomentreue Beschreibung* steht am Anfang jeder Wahrnehmungsforschung. Sie ist eine möglichst unverfälschte Beschreibung dessen, was die Beobachterin oder der Beobachter sieht. Eine „Verfälschung“ ist vor allem durch den Einfluss von Wissen zu erwarten. Die Gestalttheoretiker unterschieden deshalb zwischen einer *naiv-phänomenalen* und einer *kritisch-phänomenalen* Beschreibung. Diese Abgrenzung soll am Simultankontrast verdeutlicht werden.

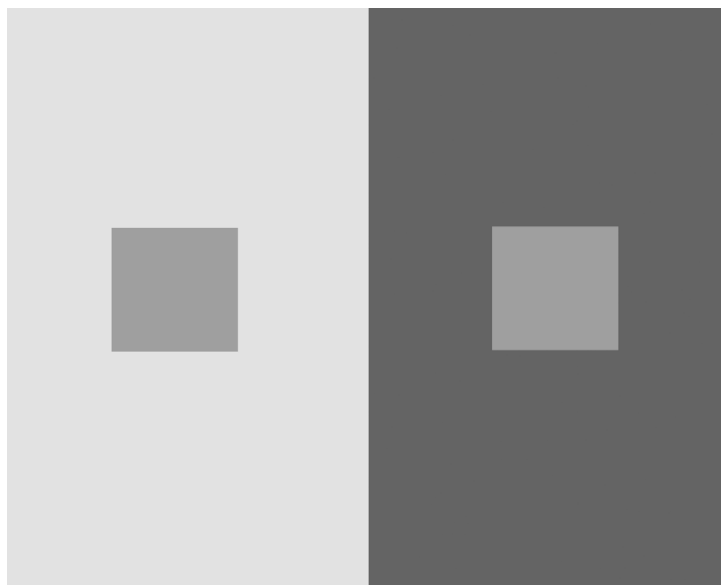


Abbildung 47: Simultaner Helligkeitskontrast. Das graue Quadrat hat in beiden Rechtecken dieselbe Leuchtdichte.

Mit *Simultankontrast* wird die Veränderung einer erlebten Helligkeit oder Farbe in Abhängigkeit von den umgebenden Helligkeiten oder Farben bezeichnet. Er ist von fundamentaler Bedeutung für die visuelle Wahrnehmung. Physikalisch betrachtet reflektieren die beiden Quadrate

in der Mitte der beiden Rechtecke dieselbe Menge an Lichtenergie (Abbildung 47). Phänomenal erscheinen sie unterschiedlich hell. Eine physikalische Beschreibung ergibt drei Flächen unterschiedlicher Luminanz, eine phänomenale Beschreibung vier Flächen unterschiedlicher Helligkeit. Die Formulierung einer Betrachterin oder eines Betrachters: „Ich sehe vier unterschiedlich helle Flächen...“ ist völlig korrekt. Es ist ein naiv-phänomenales Urteil.

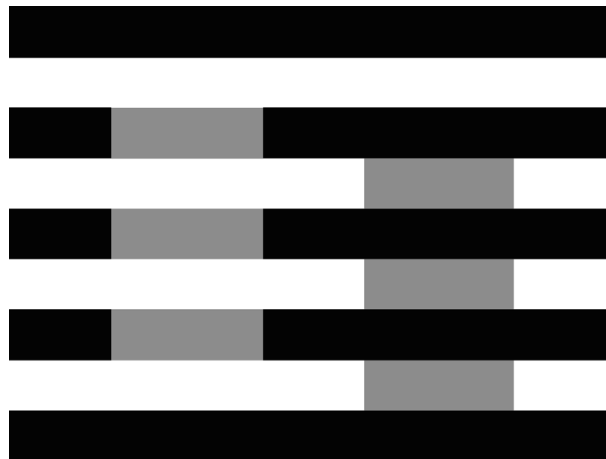


Abbildung 48: White'sche-Illusion. Die grauen Rechtecke im linken und rechten Teil der Abbildung haben dieselbe Leuchtdichte.

Der Eindruck eines Helligkeitsunterschieds ist zwingend und wird, wie sich durch Abdeckung und wieder Aufdeckung der Umgebungsfarbe nachweisen lässt, vom Wissen über die tatsächlichen physikalischen Verhältnisse (kritisch-phänomenales Urteil) kaum beeinflusst. Kritisch-phänomenales und naiv-phänomenales Urteil sind hier weitgehend unabhängig voneinander. Ein eindrucksvolles Beispiel für die Wirkung simultaner Kontraste ist die „White'sche-Illusion“ (Abbildung 48). Die grauen Rechtecke im linken Teil der Abbildung erscheinen heller als die Rechtecke im rechten Teil. Physikalisch sind sowohl die Lichtstärken als auch die Wellenlängen, die von den Streifen reflektiert werden, identisch. Selbst wer sich die Mühe macht und diese Vorlage eigenhändig erstellt, wird die Rechtecke als unterschiedlich hell erleben.

Bei den nun folgenden Beispielen gibt es fast immer einen Unterschied zwischen dem naiv-phänomenalen und dem kritisch-phänomenalen Urteil. Das kann als ein Hinweis darauf gewertet werden, dass die hier zur Anwendung kommenden Heuristiken zu Fehlern führen. Doch handelt es sich um „produktive“ Fehler. Sie haben einen funktionalen Wert, weil diese Heuristiken unter optimalen Wahrnehmungsbedingungen die Eindeutigkeit und Stabilität der visuellen Welt garantieren. Es wundert nicht, dass fast alle Beispiele aus dem Bereich der Bildwahrnehmung stammen (für eine detaillierte Analyse solcher Heuristiken vgl. auch Hoffmann, 2001).

1. Amodalität

Ständig tasten die Augen mithilfe von schnellen Bewegungen (Sakkaden) das Blickfeld ab. Figuren vom Hintergrund zu trennen ist das Ziel, eine wesentliche Voraussetzung für die Identifikation von Objekten. Sie werden in den Blick genommen und foveal (scharf, farbig und detailreich) auf der Retina abgebildet. Bei der Abgrenzung von Figur und Hintergrund und auch bei der Abgrenzung sich überlagernder Figuren spielen die Konturen bzw. Kanten eine besondere Rolle. Physikalisch sind diese Kanten in der Regel durch Intensitätssprünge markiert. Wie bereits erläutert sind die wahrgenommenen Kontraste, die durch Kanten entstehen, häufig deutlich ausgeprägter, als die ursprünglichen physikalischen Intensitätssprünge vermuten lassen.

In der Umgebungswahrnehmung finden sich viele Situationen, bei denen ein Objekt teilweise durch ein anderes Objekt verdeckt wird. Gerade hier ist es für uns wichtig, das teilverdeckte Objekt zu identifizieren, um die Situation richtig einschätzen zu können. Das gilt im Straßenverkehr ebenso wie bei einem Waldspaziergang. Das visuelle System ergänzt nach bestimmten Regeln Kanten, auch wenn sie im proximalen Reiz nicht oder

nicht vollständig enthalten sind. In der Folge behandelt es *amodale Kanten* wie modale Kanten, bei denen die zugehörige Stimulation vorliegt.

Amodale Kanten verdeutlichen, auf welcher geringen Datengrundlage das Gehirn in der Lage ist, ein vollständiges Perzept zu erstellen. Abbildung 49 zeigt ein berühmtes Beispiel, das „Kanizsa Dreieck“: ein weißes Dreieck, das drei schwarze Scheiben und ein weiteres weißes Dreieck mit schwarzem Rand überlagert. Physikalisch gesehen gibt es für große Teile des oberen weißen Dreiecks keine Luminanzunterschiede. Wie kann es zum Erkennen einer vollständigen und geschlossenen Figur kommen, wenn die Stimulation derart unvollständig ist?

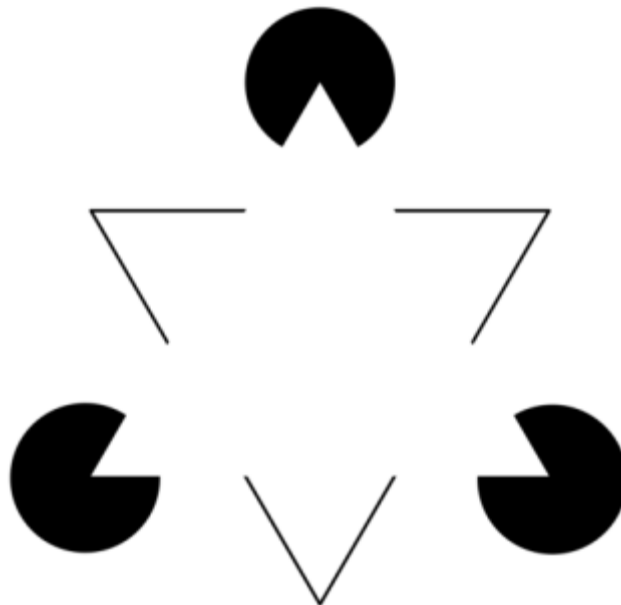


Abbildung 49: Amodales Dreieck (Kanizsa, 1955).

Nach Kanizsa (1979, S. 194) haben amodale Figuren auf der phänomenalen Ebene folgende Merkmale:

1. In einer Teilregion des visuellen Feldes erscheinen Veränderungen der Helligkeit und/oder der „Erscheinungsweise“, die diese Region von der umgebenden Region trennt (chromatische Transformation).

2. Die Teilregion wird als „vor“ oder „über“ dem Rest des Feldes gesehen (Verlagerung in die dritte Dimension).
3. Die Teilregion besitzt eine mehr oder weniger klare Grenze, die sie von der umgebenden Region trennt (Anwesenheit einer Grenze).
4. Wenn die Bedingungen optimal sind, dann sind alle hier genannten phänomenologischen Aspekte überaus zwingend und erreichen einen amodalen Charakter, der sich von der Wahrnehmung lediglich virtueller Linien unterscheidet.

Es ist irreführend, diese Figuren als „Scheinfiguren“ oder „subjektive Figuren“ zu klassifizieren, wie es häufig geschieht. Sie werden vom visuellen System wie reale Figuren behandelt. Nach dem derzeitigen Kenntnisstand werden die „Rohdaten“ für amodale Konturen vor allem parafoveal gewonnen, d. h. aus der Umgebung der unmittelbar fixierten Stelle. Bei der Identifikation von Objekten spielen Ecken und T-Stellen eine besondere Rolle. Wenn man die Kanten einer (dreidimensionalen) Ecke isoliert vom Rest betrachtet, fällt auf, dass sie aus den meisten Perspektiven eine Pfeilform aufweist und die Dreidimensionalität des Objektes unterstützt. T-Stellen hingegen zeigen in der Regel an, dass sich ein Objekt räumlich vor einem anderen befindet. Beide fungieren als Terminatoren und bilden eine Art Gerüst. Unter bestimmten Bedingungen (z. B. wenn die Balken der T-Stellen auf einer Linie liegen) werden Ergänzungen vorgenommen. Zwei wichtige Heuristiken für diese Ergänzungen lauten: 1. Identifiziere die Terminatoren und konstruiere auf dieser Basis vollständige Figuren. 2. Fehlen Teile der Konturen, ergänze sie durch Verläufe, die eine möglichst einfache Gesamtfigur ergeben.

Amodale Figuren zeigen eindringlich den produktiven Charakter der Wahrnehmung. Das Perzept ist nicht die einfache Eins-zu-eins-Abbildung des distalen Reizes. Die Ergänzungen sind jedoch nicht willkürlich oder subjektiv im Sinne von „individuell“, sondern sie folgen den Gesetzmäßigkeiten der visuellen Informationsverarbeitung.

Lassen sich automatische Heuristiken willentlich beeinflussen? Die Abbildung 50 von Walther Ehrenstein zeigt ein schwarzes Gitter mit weißen Lücken an einigen Kreuzungsstellen. Ist die wahrgenommene Form der Lücken willentlich steuerbar? Spontan wird bei den meisten Leerstellen ein Kreis gesehen. Mit Anstrengung ist es möglich, auch ein Quadrat zu sehen. Andere Figuren sind praktisch ausgeschlossen. Lediglich die größere weiße Fläche im unteren Teil der Abbildung wird spontan als Rechteck erkannt. Mit Anstrengung können hier vielleicht auch zwei Kreise gesehen werden. Das Beispiel zeigt die Möglichkeit einer willentlichen Einflussnahme. Die Heuristiken der „aufwärts“ gerichteten Verarbeitung jedoch geben den Interpretationsspielraum vor und schränken diese Optionen stark ein.

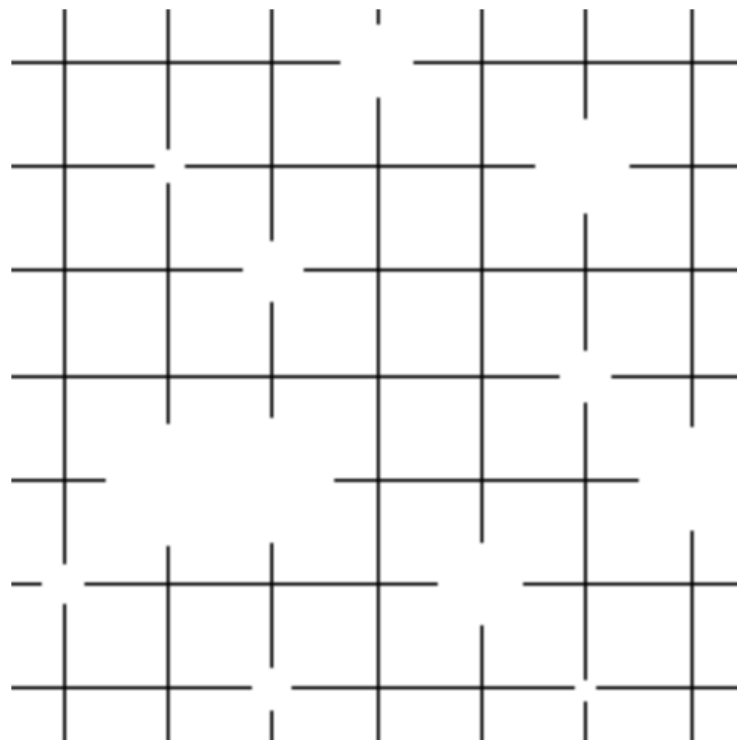


Abbildung 50: Ehrenstein-Figur. Erläuterung im Text.

Bei amodalen Figuren werden nicht nur Konturen gesehen, für die die zugehörigen Reizsprünge fehlen, sondern auch Helligkeitsunterschiede. So erscheint das obere Dreieck in Abbildung 49 heller als das untere bzw. als der Hintergrund. Physikalisch gibt es keine Luminanzunterschiede. Die

Funktion des Helligkeitseffekts ist, die Figur gegenüber dem Hintergrund deutlicher hervortreten zu lassen. Er tritt nur bei vollständigen Kanten auf. Ob die Vollständigkeit modal oder amodal hergestellt wird, spielt keine Rolle. Deutlich wird: Im „Notfall“ erzeugt das visuelle System eine Ergänzung, um die Figur gegenüber dem Hintergrund hervorzuheben. Dies dient der Ambiguitätsreduktion bei Objekterkennungen. Diese leistungsstarke Heuristik kann in besonderen Fällen auch zu Fehlinterpretationen führen, denn die Ergänzungen folgen den Probabilitäten der Außenwelt.

2. Transparenz

Phänomenale Transparenz meint, dass opake Oberflächen lichtdurchlässig erscheinen können. Die Betrachterin oder der Betrachter sieht ein transparentes graues Rechteck vor einem in einigen Bildteilen hell-grauen und in anderen Bildteilen dunkel-grauen Grund (Abbildung 51). Die Bedingungen für das Auftreten von phänomenaler Transparenz liegen nach Metelli (1974) auf drei Ebenen. Sie lassen sich als topologisch, figural und chromatisch beschreiben. Das visuelle Feld muss aus zumindest vier Regionen bestehen. Zwei dieser Regionen bilden ein gemeinsames visuelles Objekt. Da es transparent erscheinen soll, muss es dem Vordergrund zugeordnet werden. Die beiden übrigen Regionen müssen ebenfalls eine gemeinsame Grenze haben und eine zweifarbige Region bilden, die lichtundurchlässig erscheint und als „hinter“ dem lichtdurchlässigen Objekt gesehen wird. Abbildung 52 zeigt das Resultat der Verletzung dieser topologischen Voraussetzungen. Obwohl identische Flächen und dieselben Graustufen verwendet werden, entsteht in keinem Fall ein Transparenzeindruck.



Abbildung 51: Phänomenale Transparenz (Kanizsa, 1979).

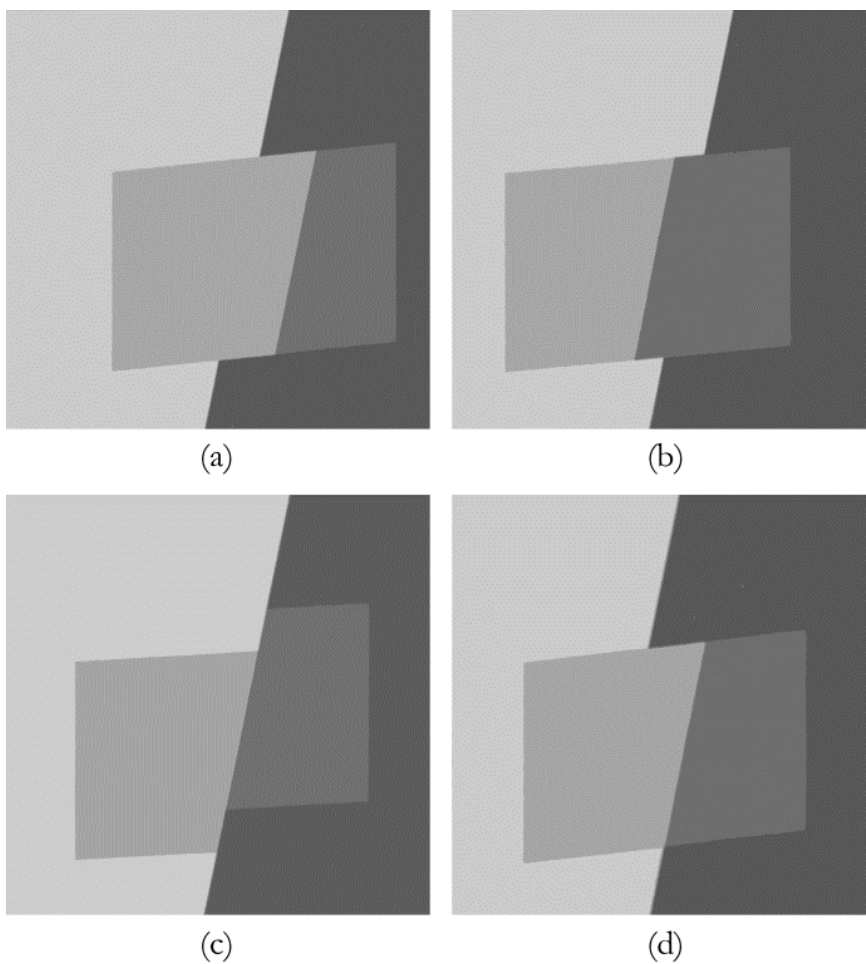


Abbildung 52: Die topologischen Voraussetzungen für phänomenale Transparenz sind verletzt. Es entsteht zwar Räumlichkeit, aber keine Durchsichtigkeit (Kanizsa, 1979).

Die topologischen Voraussetzungen ihrerseits sind zwar notwendig, aber nicht hinreichend. Die beiden Regionen, die als transparente Oberfläche gesehen werden sollen, müssen eine gemeinsame Figur bilden. Fehlt diese Einheitlichkeit, kommt es zu keinem Transparenzeindruck. Die Grenzlinie zwischen dem helleren und dem dunkleren Bereich der Figur hat eine besondere Bedeutung. Sie muss, wie in Abbildung 51 gezeigt, dem dahinterliegenden, undurchsichtigen Bereich zugeordnet werden (Metelli, 1974). Hier wird eine ganze Reihe von Heuristiken angewendet, die erst in ihrer Kombination zum Transparenzeindruck führen. Transparenz und räumliche Tiefe sind aneinander gebunden. Vier Heuristiken sind: 1. Interpretiere die kleinere Fläche als vorne liegend. 2. Verstärke den Helligkeitsunterschied zwischen Flächen unterschiedlicher Graustufen. 3. Interpretiere die Fläche, die durch eine Grenzlinie umschlossen wird, als Figur. 4. Interpretiere die Flächen mit den geringsten Helligkeitsunterschieden als zusammengehörig.

Physiologisch ist der Transparenzeffekt eine Folge der Helligkeitskorrekturen, die das visuelle System permanent vornimmt. Erste lokale Korrekturen beginnen bereits mithilfe der rezeptiven Felder und dem Prozess der lateralen Hemmung auf der Ebene der Retina. Globale Helligkeitskorrekturen, die größere Teile oder das gesamte visuelle Feld umfassen, werden in späteren Verarbeitungsschritten bis zur kortikalen Ebene vorgenommen. Beim Transparenzeffekt wirken alle Korrekturstufen zusammen. Letztlich dienen sie dem Prinzip der Kontrastbildung. Sie erleichtern die Differenzierung zwischen Figur und Grund und die Abgrenzung von Figuren. In der Umgebungswahrnehmung kommt eine weitere wichtige Funktion hinzu. Hier erlauben in vielen Situationen erst die Korrekturen die Trennung zwischen einer Veränderung der Beleuchtung und einer Veränderung des Objektes. Diese Unterscheidung ist im proximalen Reiz nicht enthalten. Hier sind beide Größen konfundiert, also nicht unabhängig voneinander bestimmbar.

Das Beispiel der Transparenz verdeutlicht charakteristische Eigenschaften des visuellen Systems. In jede Analyse sind das gesamte visuelle Feld und nicht nur einzelne Teilbereiche einzubeziehen. Unter dieser Voraussetzung kann das Perzept vorhergesagt werden. Die Wahrnehmung muss als *umgebungsabhängig* verstanden werden. Globale Merkmale können lokale Differenzierungen dominieren. Es gibt eine Ordnungstendenz zur einfachen und eindeutigen Gliederung des gesamten visuellen Felds. Das bedingt bei zweidimensionalen Abbildungen eine starke Tendenz zur Trennung von Bildebenen und zur Entwicklung einer Räumlichkeit: Sie schafft eine höhere Wahrscheinlichkeit für Eindeutigkeit. Die Betonung liegt hierbei auf Wahrscheinlichkeit, denn auch Räumlichkeit führt nicht zwingend zu einer eindeutigen und stabilen Wahrnehmung.

3. Multistabilität

Ein und derselbe distale Reiz kann zu verschiedenen Perzepten führen. Sie konkurrieren miteinander, ohne dass die Betrachterin oder der Betrachter das Ergebnis willentlich beeinflussen könnte. Ein prominentes Beispiel ist der „Necker-Würfel“ (Abbildung 53). Bei längerer Betrachtung zeigt sich neben der zunächst wahrgenommenen eine alternative räumliche Anordnung. Beide Lösungen sind annähernd gleich attraktiv. Gleichzeitig können beide Versionen aber nicht gesehen werden. Es kommt zu einem plötzlichen und willentlich kaum beeinflussbaren Sprung von einer Alternative zur anderen. Dieser folgt einem gewissen Rhythmus. Während die Stimulation gleichwertige Interpretationen zulässt, ist das Wahrnehmungsergebnis immer eindeutig, selbst wenn große Teile der Stimulation neu gruppiert und interpretiert werden müssen. Auch hier ist das Ziel die Ambiguitätsreduktion.

Auf der Ebene der Prozesse ist der Necker-Würfel ein schönes Beispiel für die Arbeitsweise eines nichtlinearen Systems: Zwei ausgezeichnete Zustände erreichen beide eine vergleichbare Stabilität. Das System pendelt zwischen ihnen hin und her. Eine Verallgemeinerung führt zur Multistabilität. Perzepte können immer nur dann entstehen, wenn auf der Ebene der Prozesse ein Zustand temporärer Stabilität erreicht ist.

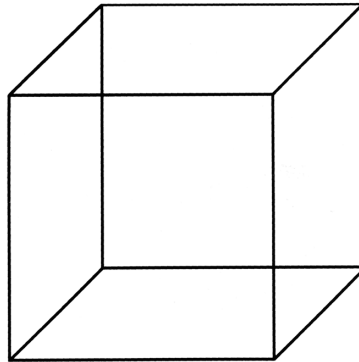


Abbildung 53: Necker-Würfel. Erläuterung im Text.

Betrachtet man den Drahtwürfel (nicht seine zweidimensionale Abbildung), ist er ein Objekt von hoher Symmetrie. Alle Kanten haben dieselbe Länge, alle Winkel sind gleich. Mathematische Transformationen des Würfels wie Rotationen und Spiegelungen führen zu identischen Lösungen. Es lassen sich verschiedene Projektionen des Drahtwürfels auf einer zweidimensionalen Fläche erzeugen (vgl. Abbildung 54). Lediglich die Versionen (d) bis (f) sind perspektivisch korrekt (mit einem, mit zwei oder mit drei Fluchtpunkten); (a) ist eine unmögliche Sicht auf den Würfel, (b) und (c) sind Parallelprojektionen ohne Fluchtpunkt bzw. einem Fluchtpunkt im Unendlichen. Damit sind sie – streng genommen – unmögliche Sichtweisen in einem endlichen Raum. Das heißt nicht, dass Ansichten auf bestimmte Objekte, wie (a) und (b), nicht existieren können. Sie stammen dann von einem nicht gleichseitigen Würfel.

Welche dieser Projektionen rufen einen räumlichen Eindruck hervor? Welche eine Bistabilität oder Multistabilität? Bei den perspektivischen Zeichnungen (a) bis (f) bleiben die Symmetriemerkmale in

unterschiedlichem Ausmaß erhalten. Vielfach sind die Linien nicht gleich lang und verlaufen nicht parallel. Die Winkel sind keine rechten Winkel. Interessanterweise zeigen gerade die beiden Versionen (c) und (d), die mit sechs bzw. acht Symmetrieachsen dem Drahtwürfel am nächsten kommen, den geringsten Tiefeneindruck.

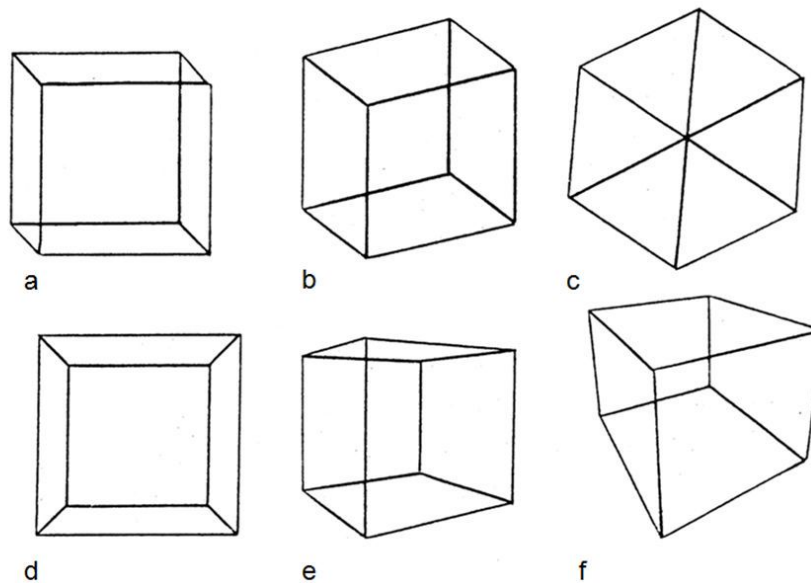


Abbildung 54: Projektionen des Necker-Würfels (Zimmer, 1995). Erläuterung im Text.

Was konstituiert den räumlichen Eindruck einer Konturzeichnung auf einer zweidimensionalen Fläche? Hochberg und Brooks (1960), die reversible figürliche Darstellungen dreidimensionaler Objekte untersuchten, fanden drei Faktoren: 1. Einfachheit versus Komplexität (gemessen durch die Anzahl der Winkel), 2. Kontinuität versus Segmentierung (Anzahl der Liniensegmente) und 3. Symmetrie versus Asymmetrie (Anzahl der unterschiedlichen Winkel relativ zur Gesamtzahl der Winkel). Sie postulieren eine Unabhängigkeit dieser Faktoren und ein additives Modell: Je größer die Komplexität, die Diskontinuität und die Asymmetrie, umso größer ist die zu erwartende räumliche Wirkung einer Projektion. Zimmer (1995) vermutet dagegen, dass es sich um einen kompetitiven Prozess handelt. Sein Modell nimmt an, dass der

Tiefeneindruck dann maximal wird, wenn die drei Faktoren vergleichbare Werte aufweisen. Demnach sollte in Abbildung 54 bei (a) und (b) der räumliche Eindruck stärker sein als bei (e) und (f), die Hochberg und Brooks (1960) favorisieren würden.

Offensichtlich sind Räumlichkeit und Multistabilität nicht unabhängig voneinander. Bei längerer Betrachtung lässt sich bei allen sechs Vorlagen eine Räumlichkeit herstellen. Während sich bei (a), (b), (e) und (f) dieser Eindruck unmittelbar ergibt und es etwas später zu einem Wechsel zwischen zwei räumlichen Lösungen kommt (Bistabilität), können bei (c) und (d) oft drei Lösungen gesehen werden (Tristabilität): eine nicht-räumliche, die zunächst in Erscheinung tritt, und zwei räumliche, die sich bei längerer Betrachtung abwechseln. Die räumliche Interpretation reduziert damit die Mehrdeutigkeit um eine Variante. Hier lässt sich wieder prüfen, inwieweit die automatischen Heuristiken willentlich beeinflussbar sind. Bei welcher Projektion kann die Tiefenwirkung am leichtesten beeinflusst werden?

Das visuelle System steht letztlich immer vor einer Aufgabe: Es muss entscheiden, welche Lichtreflexion eines dreidimensionalen Objektes die zweidimensionale Projektion auf der Retina erzeugt hat. Daher beziehen sich viele Heuristiken auf das Problem der inversen Projektion. Sie begrenzen die unendliche Zahl der Möglichkeiten. Nach dem Probabilitätsprinzip schätzen sie die Wahrscheinlichkeit für eine zufällige Ansicht eines Gegenstandes als sehr gering ein. Einem Objekt wird meist in einer charakteristischen Ansicht begegnet. Aus dieser Ansicht wird es schnell und ohne Mühe erkannt. Als Beispiel der vermeintlich einfachste Fall: eine gerade Linie auf einem weißen Blatt Papier. Hierfür kann es zahllose dreidimensionale Interpretationen geben: Es könnte ein Ring aus der Seitenansicht sein, ebenso auch ein flaches Objekt mit einem quadratischen Grundriss. Jede beliebig gekrümmte Kante, exakt von der Seite gesehen, würde diese Projektion erzeugen. Warum sehen wir eine Linie und keine Ansicht eines dreidimensionalen Objektes? Weil nicht alle

möglichen Ansichten von dreidimensionalen Objekten wahrscheinliche Ansichten sind. Gesucht werden charakteristische und stabile Ansichten. Diese werden auch als *reguläre Ansichten* bezeichnet (Hoffmann, 2001). Die Heuristiken lauten: 1. Wenn möglich, bilde reguläre Ansichten eines dreidimensionalen Objektes. 2. Findet sich keine reguläre Ansicht eines dreidimensionalen Objektes, dann interpretiere die Stimulation als zweidimensional.

Die einzige stabile Ansicht des Gebildes auf dem Papier ist eine gerade Linie. Auch bei einer Veränderung der Position kann diese Interpretation beibehalten werden. Donald Hoffman (2001) hat eine Reihe von Regeln formuliert, die sich auf dieses Problem beziehen. Hier eine Auswahl: 1. Interpretiere eine gerade Linie in einem Bild als eine gerade Linie dreidimensional. 2. Wenn die Enden zweier Linien in einem Bild zusammenfallen, interpretiere sie so, dass sie auch dreidimensional zusammenfallen. 3. Interpretiere Linien, die in einem zweidimensionalen Bild kollinear sind (sich entsprechen), auch dreidimensional kollinear. Alle Regeln sollen helfen, die Anzahl möglicher Objekte, die durch ihre Reflexion am Sinnesorgan dieselbe Stimulation hervorrufen können, zu reduzieren. Multistabilität ist aus diesen Gründen kein Spezialfall.

Wie gezeigt wurde, dass ein und derselbe distale Reiz verschiedene Wahrnehmungsergebnisse erzeugen kann, lässt sich ebenfalls nachweisen, dass unter verarmten Beobachtungsbedingungen unterschiedliche distale Reize dasselbe Wahrnehmungsergebnis produzieren können. Ein anschauliches Beispiel hierfür liefern die Experimente, die Adelbert Ames zwischen 1940 und 1950 durchführte. Die Abbildung 55 zeigt die „Amesschen Stühle“. In der oberen Reihe ist das Wahrnehmungsergebnis zu sehen: In allen drei Situationen ergibt sich für die Beobachterin oder den Beobachter phänomenal ein identischer Stuhl. In der unteren Reihe werden die zugehörigen distalen Reize präsentiert. Lediglich für das linke Bild der oberen Reihe bildet ein Stuhl die Grundlage. In den beiden anderen Fällen handelt es sich um ein kompliziertes, realitätsfernes

Gebilde, bei dem (in einem der Fälle) einzelne Teile nicht einmal miteinander verbunden sind (mittlere, untere Abbildung).

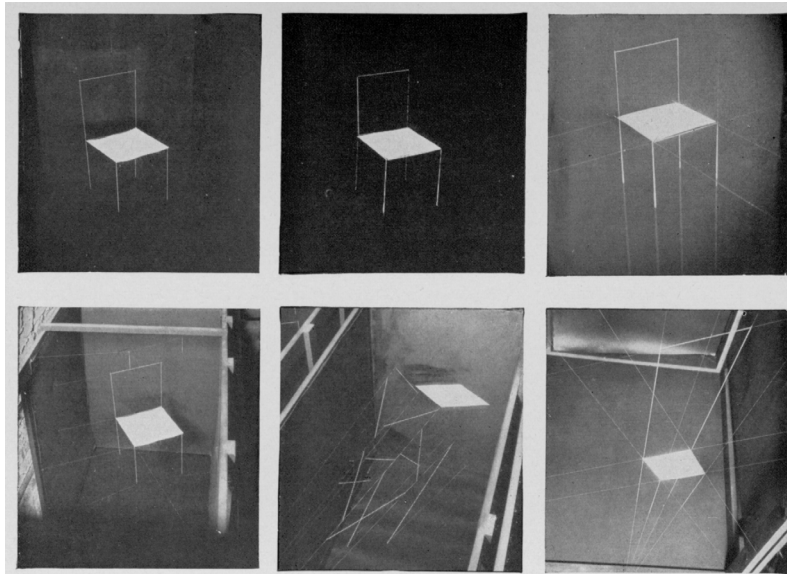


Abbildung 55: Amessche Stühle. Die obere Reihe zeigt die Sicht der Betrachterin oder des Betrachters, die untere das zugehörige Konstruktionsprinzip.

Ames stellte die distalen Reize mithilfe unterschiedlicher Materialien (z. B. Draht) und gezeichneter Flächen (die weiße Fläche in der mittleren unteren Abbildung ist ein Rhombus) her. Er schuf optische Anordnungen, die bei einer monokularen Betrachtung durch eine punktförmige Öffnung (Guckkasten) identisch aussehen. Die Betrachterin oder der Betrachter konnte erst bei einer Öffnung des Kastens erkennen, dass der distale Reiz keine Ähnlichkeit mit einem Stuhl hatte.

Die Beobachtungsbedingungen entsprechen prinzipiell den Restriktionen bei einer linearperspektivischen Darstellung dreidimensionaler Szenen auf einer zweidimensionalen Fläche (z. B. einem Foto): ein stationärer, genau definierter Standort und monokulare Beobachtungsbedingungen. Sie sind notwendig, um eine korrekte Abbildung zu garantieren. Dann sind die Projektionen auf der Retina identisch und mit ihr auch der proximale Reiz. Die Suche nach einer regulären Ansicht für diese Projektion führt für alle

drei distalen Reize zum Perzept des Stuhles. Der Mensch fällt hier auf zufällige Ansichten herein, die zwar potenziell in der Natur vorkommen können, aber sehr unwahrscheinlich sind. Hinzu kommt, dass wir durch die Zirkularität des Sehvorgangs unter nicht verarmten Bedingungen in kürzester Zeit eine Ansicht finden könnten, die die „Zufälligkeit“ des ursprünglichen Blickwinkels offensichtlich werden ließe.

4. Schatten

Schatten ist eine Region relativer Dunkelheit innerhalb einer beleuchteten Region. Er erscheint, wenn ein Objekt ganz oder teilweise das Licht verdeckt. Wegen der geradlinigen Ausbreitung des Lichtes stellt der Schatten eine Verlängerung der Verbindungslinien der Lichtquelle zum Körper dar. In Abhängigkeit von der Art der Lichtquelle verändert sich auch der Schatten. Punktförmige Beleuchtungskörper mit einer scharfen Strahlenbündelung werfen einen klar abgegrenzten, harten Schatten. Wird der Radius der Lichtquelle vergrößert, wird der Schatten weicher. Die Sonne und andere flächige Lichtquellen bilden einen Kernschatten (kein direktes Licht), der von einem Halb- oder Teilschatten umgeben wird.

Unter natürlichen Umgebungsbedingungen werden Schatten schnell sehr komplex, weil nicht nur das Strahlungslicht, sondern auch das Umgebungslicht zu berücksichtigen ist. Es überrascht nicht, dass die Wiedergabe natürlicher Schattenverhältnisse bei der Erzeugung virtueller Welten eine anspruchsvolle Aufgabe darstellt und für Theorien der Objekterkennung immer ein Prüfstein ist. Eine Vorstellung vermittelt das sogenannte „bump-mapping“-Verfahren (vgl. Abbildung 56). Ein Gegenstand wird mithilfe möglichst weniger Polygone dargestellt. Das Resultat ist ein geometrisches Gerüst aus aneinander grenzenden Flächen. In einem zweiten Schritt werden die Einzelflächen so manipuliert, dass der Eindruck von unebenen Flächen entsteht. Wie zu sehen ist, gewinnt die Gesamtoberfläche an Detailreichtum und weist differenzierte Übergänge in der Helligkeit auf, die vom visuellen System als Licht und Schatten bzw.

als Reflexion interpretiert werden. Insgesamt gewinnt die Figur an Plastizität.

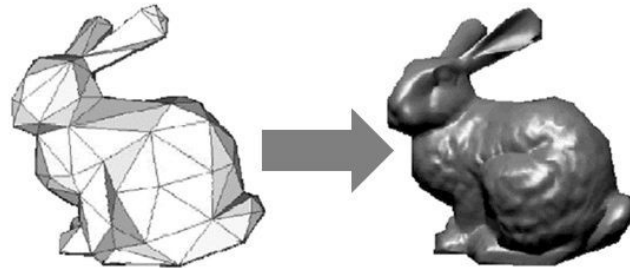


Abbildung 56: Konstruktion einer Figur mit hoher Plastizität und Lichtreflexionen nach dem „bump-mapping“-Verfahren. Erläuterung im Text.

Wie unterschiedlich die Schattenbildung bei gleichem Lichteinfall dargestellt werden kann, zeigt der Vergleich zweier Figurenstudien von Giovanni Battista Piazzetta (1683-1754), die als Stiche von Marco Pitteri (links) und Franco Bartolozzi (rechts) erhalten sind (vgl. Abbildung 57). Die Fassung von Bartolozzi zeigt den Akt mit klareren Konturen und deutlicheren Helligkeitssprüngen innerhalb der Figur. Schatten und Schattierung wirken schematisch. Der Kupferstich von Pitteri dagegen hat eine feinere Modellierung, die auch die Grautonwerte der Zeichnung wiederzugeben vermag. Die Beispiele zeigen, dass der Schatten selbst keine scharfe Kontur haben darf. Würde man den Schatten mit einer Grenzlinie versehen, entsteht ein eigenständiges Objekt. Schatten hat weder Grenzlinie noch Volumen. Gibt es in einem Bild Fehler bei der Schattierung, können sie lokal entdeckt werden. Die Betrachterin oder der Betrachter muss somit nicht immer das gesamte Bild in seine Analyse einbeziehen.

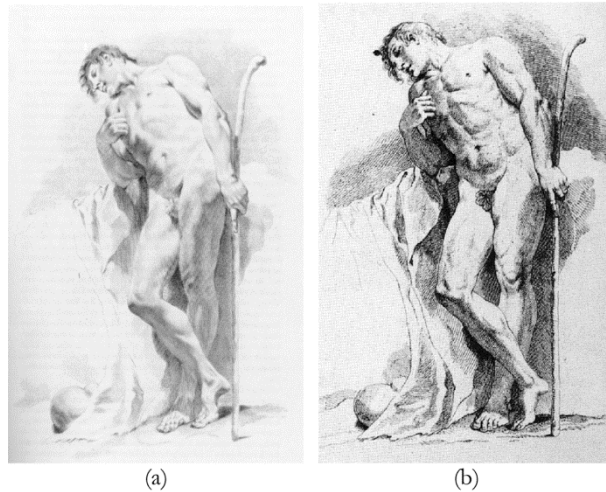


Abbildung 57: Figurenstudie von Marco Pitteri (a) und Francesco Bartolozzi (b) nach Giovanni Battista Piazzetta (1683-1754).

Auf die Fragen nach dem Informationsgehalt des Schattens ließe sich kurz antworten: Schatten schafft Räumlichkeit und reduziert Mehrdeutigkeit. Er unterstützt den Objektcharakter. Doch für die visuelle Wahrnehmung stellt die Identifizierung von Schatten ein Problem dar. Indem die aktuell wahrgenommene Helligkeit und Farbe einer Oberfläche sowohl von den Reflexionseigenschaften der Oberfläche als auch von der Beleuchtung abhängig ist, wird bei komplexeren Stimulationen der relative Einfluss der einzelnen Größen immer schwerer zu ermitteln. In der Tierwelt ist das Prinzip der „Gegenschattierung“ verbreitet. Der Körper wird gegen Angreifer getarnt, indem das von oben einfallende Licht bezüglich seiner Schattenwirkung an der Körperunterseite neutralisiert wird. Die Oberseite des Körpers ist dunkel und hebt sich nicht vom dunklen Untergrund ab, während die Unterseite heller ist, was den Kontrast zur (oberhalb liegenden) Lichtquelle verringert. Die Unterseite des Tieres erscheint, als wäre sie von der Sonne beschienen (z. B. bei manchen Fisch- und Raupenarten). Die relative Verbreitung dieses Prinzips lässt darauf schließen, dass Schattierung ein wichtiger Hinweisreiz für die Dreidimensionalität einer Form ist.

Um den Einfluss der Schattierung auf die Wahrnehmung der Räumlichkeit einer Form unabhängig von anderen Tiefenhinweisen untersuchen zu können, hat Ramachandran (1988, 1995) Vorlagen konstruiert, die aus einem homogenen Hintergrund und einer wechselnden Anzahl und Anordnung von Kreisen bestehen (vgl. Abbildung 58). Variiert wird die Richtung der Helligkeitsverteilung in den Kreisen. Eine phänomenale Beschreibung enthält in der Regel immer die dritte Dimension. Es werden keine Kreise gesehen, sondern Kugeln oder Mulden. In Bezug auf beide Interpretationen bleibt das Muster mehrdeutig. Interessant ist, dass entweder alle Kreise als Kugeln (konvex) oder als Mulden (konkav) erscheinen, nicht aber einzelne Kreise als Kugeln und andere als Mulden. Ramachandran (1988) vermutet als Grund für die einheitliche Interpretation, dass das Gehirn zu einem Zeitpunkt nur eine Lichtquelle für das gesamte visuelle Feld akzeptiert. Eine weitere wichtige Heuristik lautet folglich: Interpretiere alle Helligkeitsunterschiede im visuellen Feld als von einer Lichtquelle stammend. Das könnte in der natürlichen Beleuchtungssituation auf dem Planeten Erde begründet sein.

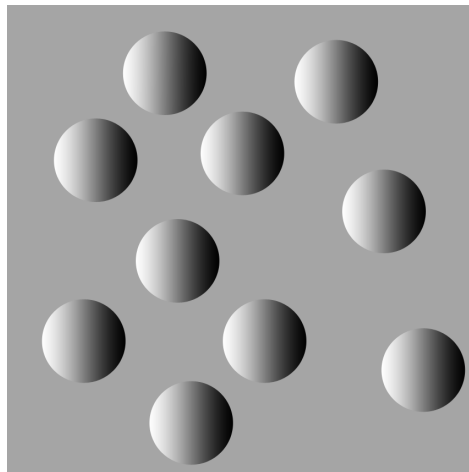
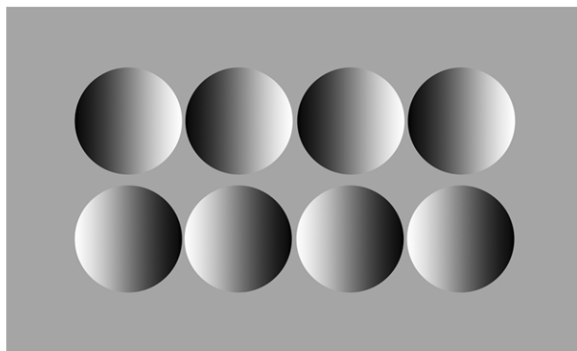


Abbildung 58: Mehrdeutige Figuren, die entweder als Kugeln oder als Mulden gesehen werden können (Ramachandran, 1995).

Die Abbildung 59 (a) zeigt zwei Reihen von Kreisen. Jeder Kreis der unteren Reihe wurde durch eine Spiegelung des Kreises in der oberen

Reihe an seiner vertikalen Achse erzeugt. Bei längerer Betrachtung wird eine Reihe oft als Kugeln, die andere als Mulden gesehen. Wird eine Reihe abgedeckt, kann die verbleibende Reihe sowohl als konvex als auch als konkav gesehen werden. Das kann als Hinweis verstanden werden, dass für das gesamte visuelle Feld nur eine Beleuchtungsquelle vorausgesetzt wird. Die Abbildung 59 (b) zeigt, wie die Ambiguität durch Einbindung in eine umgebende Form aufgelöst werden kann. Das schlauch- oder röhrenartige Gebilde wird von einer rechts außerhalb des Bildes befindlichen Lichtquelle beschienen. In Übereinstimmung mit diesem Lichtschema wird der untere Kreis als Mulde und der obere als Auswölbung gesehen. Ambiguität entsteht dann, wenn die Betrachterin oder der Betrachter die umgebende Form abdeckt und einen Kreis isoliert betrachtet. Dass er konvex erscheint, erklärt sich mit der generellen Tendenz des visuellen Systems, entsprechend der Häufigkeitsverteilung von gekrümmten Oberflächen in der Natur, Formen bevorzugt als konvex zu sehen (Prinzip der Probabilität).



(a)



(b)

Abbildung 59: (a) Eine Reihe wird zumeist als Mulden, die andere als Kugeln gesehen. Ein Hinweis darauf, dass für das gesamte Gesichtsfeld nur eine Beleuchtungsquelle angenommen wird. (b) Durch eine Einbettung in eine eindeutige Gesamtsituation kann die Mehrdeutigkeit reduziert werden (Ramachandran, 1995).

Ramachandran (1995) postuliert nicht nur eine einheitliche Lichtquelle, sondern konnte zeigen, dass die Beleuchtung auch als „von oben kommend“ gesehen wird. Die Bezeichnung „oben“ bezieht sich dabei auf die retinalen Abbildungsverhältnisse (proximaler Reiz). Die Abbildung 60 zeigt in Teil (a) Kugeln und in Teil (b) Mulden. Wird die Vorlage um 180° gedreht, invertieren auch die Formen.

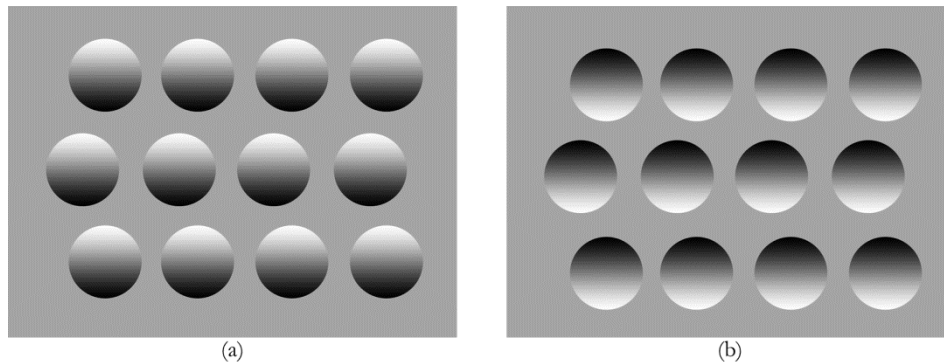


Abbildung 60: In (a) werden die Kreise als Kugeln gesehen, in (b) als Mulden. Dreht die Betrachterin oder der Betrachter das Bild um 180°, invertieren auch die Formen (Ramachandran, 1995).

Wie verhält sich das visuelle System, wenn es widersprüchliche Informationen über die Lokalisation der Lichtquelle erhält? Unter natürlichen Umgebungsbedingungen kommt diese Situation nicht vor, für die Analyse von Heuristiken ist sie jedoch interessant. Abbildung 61 zeigt Beispiele. Zu beurteilen ist jeweils, welche der beiden mittleren Kreise einen höheren Grad an Konvexität aufweist. In den Versuchen von Ramachandran gab es eindeutige Ergebnisse. Bei A ist es der Kreis in der Vorlage (b), bei B ist es der Kreis in der Vorlage (a) und bei C ist es wieder (b). Das Ergebnismuster stimmt mit der Annahme überein, dass das Ausmaß der wahrgenommenen Räumlichkeit im Sinne eines Zentrum-Umfeld-Effektes von den umgebenden Objekten beeinflusst wird.

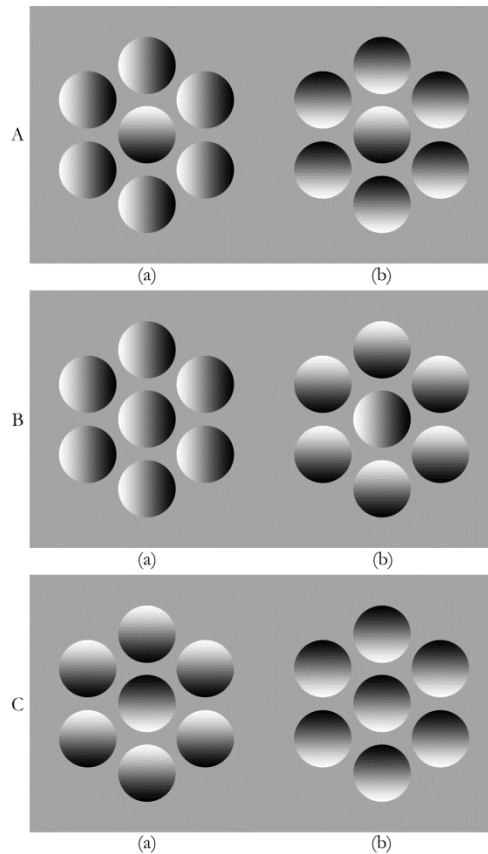


Abbildung 61: Demonstrationen des Zentrum-Umfeld-Effektes, bei dem die wahrgenommene Räumlichkeit von der Umgebung beeinflusst wird (Ramachandran, 1995). Erläuterung im Text.

Solche Effekte wurden für verschiedene Dimensionen wie Helligkeit, Farbe und Bewegung nachgewiesen. Sie besagen, dass die Charakteristika einer Bildregion durch die benachbarten Bildregionen immer dann beeinflusst werden können, wenn eine Variation auf derselben Dimension vorliegt. In diesem Fall wird die Tiefenwirkung für den mittleren Kreis gefördert, wenn die umliegenden Kreise eine umgekehrte Polarität haben. Ist sie nicht gegeben, dann besteht eine Tendenz, die Helligkeitsunterschiede der Oberflächen als Unterschiede der Reflektanz dieser Oberfläche zu interpretieren, was zu einer Reduktion des Tiefeindrucks führt. Der Zentrum-Umfeld-Effekt ist ein Beispiel für die Dominanz globaler Merkmale, die die Interpretation lokaler Merkmale bestimmen. Erst die globale Dominanz garantiert die Stabilität der visuellen Welt.

5. Farbausbreitung

Es ist schwer, die Bedeutung von Farbe für die Konstruktion einer visuellen Welt zu beurteilen. Es gibt viele Tierarten, deren visuelles System keine Farbunterscheidungen kennt. In der Regel reicht es aus, wenn Texturunterschiede über die Helligkeitskontraste bemerkt werden oder wenn sich Flächen mit gleicher Helligkeit, aber unterschiedlicher Struktur, relativ zueinander bewegen. Die Funktion der Farbe wird gewöhnlich definiert als die Empfindung, die es uns ermöglicht, zwischen zwei strukturlosen Flächen gleicher Helligkeit zu unterscheiden. Durch die Verwendung des Empfindungsbegriffs wird Newtons Einsicht berücksichtigt, dass das Licht selbst nicht farbig ist, sondern dass erst die neuronale Verarbeitung der unterschiedlichen Wellenlängen in der Retina und in den visuellen Arealen des Cortex zu der Empfindung „Farbe“ führt.

Rein spektrale Unterschiede zwischen Oberflächen finden sich unter den Bedingungen der Umgebungswahrnehmung ausgesprochen selten. Fast immer sind sie mit Intensitätsunterschieden verbunden. Die Forschung der letzten Jahrzehnte hat gezeigt, wie schwierig die Diskriminierung von isoluminanten Flächen ist. Eine Sichtweise gewinnt an Bedeutung, die die Farbe als diejenige Empfindung auffasst, die es erlaubt, Objekte zu unterscheiden, die aufgrund ihrer Textur nur schwer abzugrenzen sind (vgl. Gegenfurtner und Rieger, 2000). Farbe wäre eine Möglichkeit, Entscheidungen schneller und zuverlässiger zu treffen.

Für die Wahrnehmung von Farben gilt ein mit Helligkeiten vergleichbares Konstanzprinzip. Die spektrale Zusammensetzung des das Auge erreichenden Lichts wird von den Reflexionseigenschaften der Oberflächen und von den Merkmalen der Beleuchtung bestimmt. Würde sich die Farbwahrnehmung ausschließlich auf die sensorischen Informationen stützen, wäre sie permanenten Schwankungen unterworfen. Tatsächlich erleben wir – innerhalb gewisser Grenzen – Farbe als in ihrer Qualität identisch. Eine rote Oberfläche behält

anschaulich auch dann ihren roten Farbton, wenn sie unterschiedlichen spektralen Verhältnissen ausgesetzt ist und in der Folge der Wellenlängenbereich des Lichtes, der die Retina erreicht, objektiv variiert. Offensichtlich sind leistungsstarke Heuristiken am Werk. Dieser Aspekt der Farbwahrnehmung soll bei der Diskussion der Konstanzphänomene aufgegriffen werden.

Die maximale Information über die Form eines Gegenstandes erhält das visuelle System nicht über die Flächen, sondern von den Ecken und Kanten. Daher ist für Maler oder Zeichner nicht notwendig, an jeder Stelle im Bild Helligkeit und Farbe zu definieren. Es muss an den Übergängen geschehen. Ohne Reizänderung geht das visuelle System von einer Kontinuität der Fläche aus (Prinzip der Ergänzung). Auf die Entdeckung von Übergängen sind die antagonistisch aufgebauten rezeptiven Felder spezialisiert. Wird dieses Prinzip global, im gesamten Gesichtsfeld, angewendet, führt es dazu, dass die Beleuchtungsveränderungen von den Veränderungen der Reflektanz der Oberflächen getrennt werden können. Die Regel, dass nur die Übergänge markiert werden müssen, lässt sich exemplarisch an einem Aquarell von Paul Cézanne zeigen (Abbildung 62). Der weiße Untergrund wirkt auch dann farbig, wenn nur der kleinere Teil der zugehörigen Fläche aquarelliert wurde.



Abbildung 62: Paul Cézanne, Die Kalkbrennerei, 1890-1894.

Das Farbsystem hat eine überraschend geringe Auflösung. Neurone dieses Systems haben deutlich größere rezeptive Felder und ihre Zahl ist kleiner als die der Neurone für die Formwahrnehmung. Unsere Farbwahrnehmung ist gröber als die Formwahrnehmung. Verwendet ein Maler für den Hintergrund eine Farbe mit geringen Helligkeitskontrasten und setzt gleichzeitig Konturen mit hohen Kontrasten ein, sieht die Betrachterin oder der Betrachter eine Übereinstimmung der Kolorierung mit der Form, selbst wenn sie objektiv nicht gegeben ist (z. B. bei Aquarellen von Auguste Rodin). Das visuelle System verarbeitet nicht verschiedene Eigenschaften des retinalen Bildes isoliert voneinander, um sie später additiv zusammensetzen. Es gibt starke Interaktionen zwischen den Kanälen oder Modulen.

Das Farbsystem kodiert, wie am Beispiel des Cézanne-Aquarells gezeigt, lediglich die Ecken der Figuren und arbeitet dann mit einer *Ausfüllfunktion* (Paradiso und Nakayama, 1991; Caputo, 1998; De Weerd, Desimone und Ungerleider, 1998). Es reicht aus, die Kolorierung anzudeuten. Das Vorgehen entspricht der Funktionsweise der doppelten Gegenfarbentzellen, die auf die Grenzen von farbigen Flächen, nicht auf die Farbflächen antworten.

Damit muss die Frage, wann sich Farben phänomenal mischen, neu gestellt werden. Einzubeziehen ist das Verhältnis von Farbe und Form. Unter der Annahme, dass das visuelle System die Grenzen eines Gegenstandes zunächst mithilfe des hochauflösenden Formsystems definiert und dann mit dem niedrigauflösenden Farbsystem die Farbfläche ergänzt, ist die Farbmischung davon abhängig, ob die Farben als zur selben Oberfläche gehörig wahrgenommen werden. Wenn es einen großen Helligkeitskontrast zwischen zwei Farben gibt, ist es für das Formsystem schwierig, sie als zu einer Oberfläche gehörig zu interpretieren. Umgekehrt ist die Farbausbreitung dann am größten, wenn die Helligkeitsunterschiede zu benachbarten Farben gering sind (vgl. auch Livingstone, 2002).

6. Biologische Bewegung

Die Bewegungswahrnehmung war lange Zeit ein vernachlässigtes Forschungsthema. Dies liegt sicher nicht an ihrer geringen Bedeutung. Unter den Bedingungen der Umgebungswahrnehmung geht es fast immer um die Wahrnehmung von Eigen- oder Fremdbewegung. Zwei Gründe sind verantwortlich: Vor dem Einzug der Computer in der Forschung war es apparativ schwierig, grundlegende Formen der Bewegungswahrnehmung unter kontrollierten Bedingungen zu untersuchen. Zudem dominierte die Vorstellung, dass die Bewegungswahrnehmung ein komplizierter Spezialfall ist und es deshalb sinnvoll sei, zunächst den einfacheren und grundlegenderen Fall einer stationären Wahrnehmung zu verstehen. Die Arbeiten von Gibson (1950, 1966, 1979) haben diese Sichtweise infrage gestellt. Heute gilt eine dynamische Stimulation als die Grundsituation der visuellen Wahrnehmung.

Der schwedische Forscher Gunnar Johansson entwickelte in den 1970er Jahren ein Paradigma, um der Frage nachzugehen, ob das visuelle System spezifische Bewegungsinformationen nutzt. Kann Bewegung auch ohne Form erkannt werden? An den Hauptgelenken von Darstellern wurden sechs bis zwölf punktförmige Lichtquellen befestigt (vgl. Abbildung 63). In einem abgedunkelten Raum wurden Bewegungsabläufe gefilmt: gehen, hüpfen, schreiten. Charakteristische Bewegungsmuster konnten losgelöst von ihrer Gestalt aufgezeichnet werden. Sieht ein Beobachterin oder ein Beobachter ein Standbild oder eine unbewegte Person, so sieht er eine zufällige Anordnung von hellen Punkten auf einer dunklen Fläche. Bei Bewegung jedoch wird in Bruchteilen einer Sekunde nicht nur erkannt, dass es sich um einen Menschen handelt, sondern auch, welche Bewegung er gerade ausführt (vgl. Abbildung 64).

Nachfolgende Untersuchungen haben gezeigt, dass sehr viel weitergehende Differenzierungen möglich sind. So lässt sich beurteilen, ob eine Person eine körperliche Beeinträchtigung hat oder ob ein Objekt, das gerade hochgehoben wird, leicht oder schwer ist. Die Identifikationsrate für Freunde und Bekannte liegt dabei deutlich über der Ratewahrscheinlichkeit.

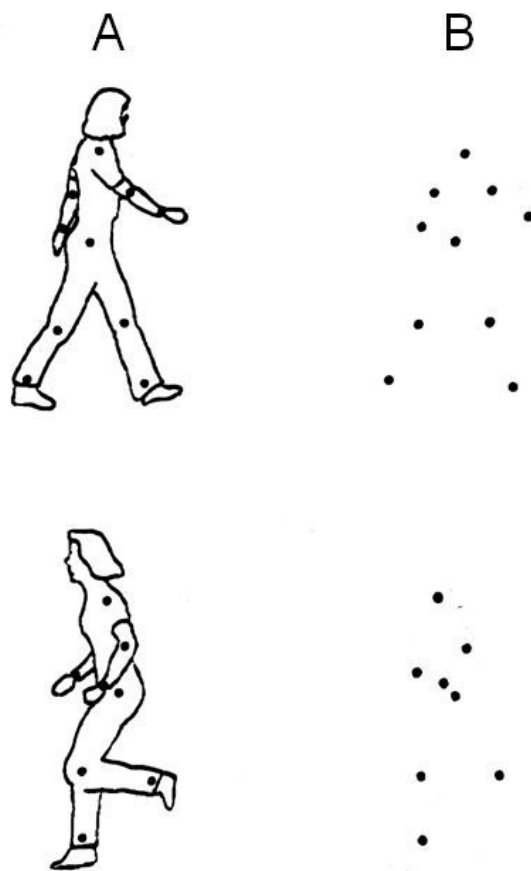


Abbildung 63: Anordnung der Markierungspunkte bei Johansson. Konturen einer gehenden und einer laufenden Person und den daraus resultierenden Lichtpunktkonfigurationen (Johansson, 1973).

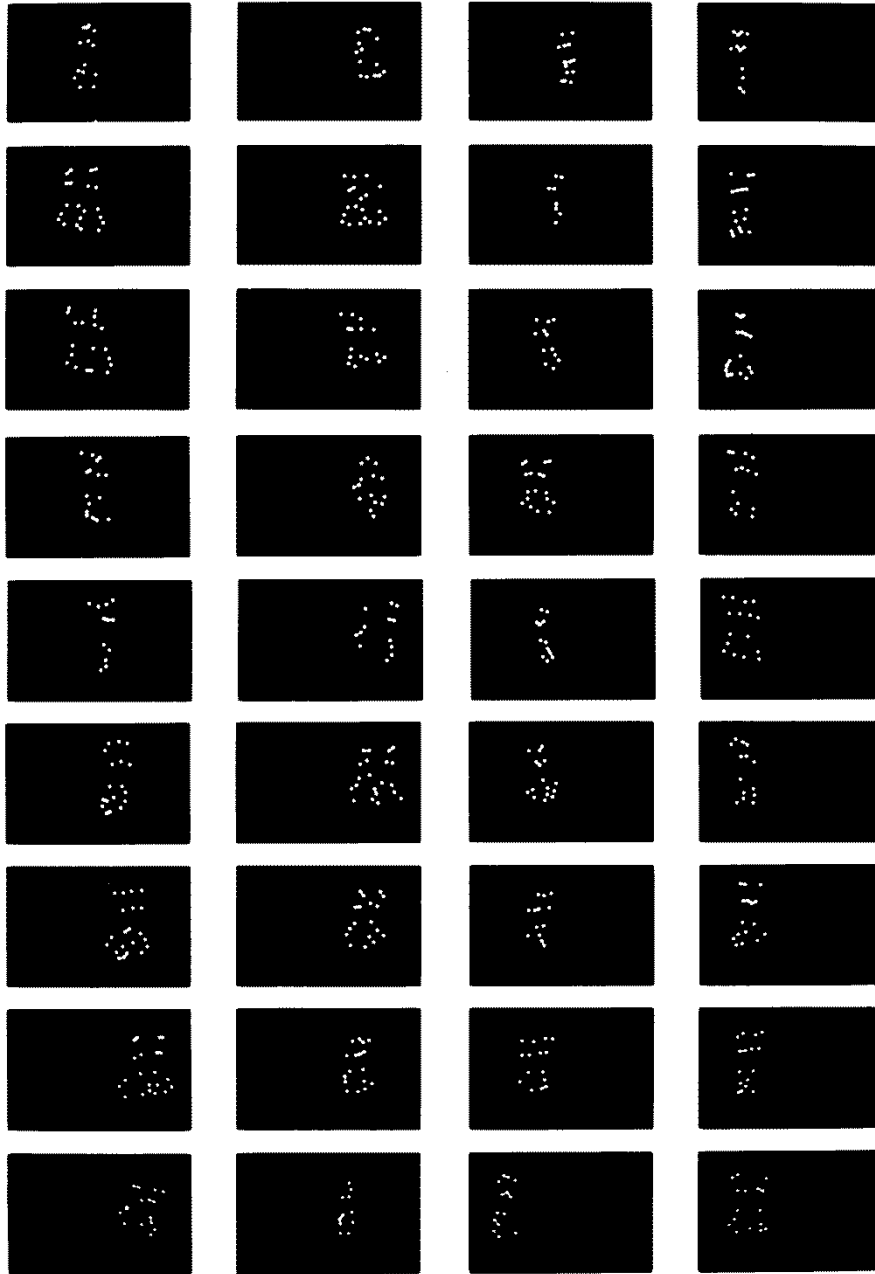


Abbildung 64: Punktmuster eines tanzenden Paares, das von Johansson gefilmt wurde. Je zwei Lämpchen befanden sich an Schultern, Ellbogen, Handgelenken, Hüften, Knien und Fußgelenken. In spaltenweiser Anordnung ist jedes sechste Bild eines 35-mm-Films dargestellt. Versuchspersonen können beim Betrachten des Films in Sekundenbruchteilen angeben, dass sie die Bewegung von zwei Personen sehen (Johansson, 1986).

Johansson deutete die überraschenden Leistungen des visuellen Systems als Spezialisierung im Laufe der Evolution. Mit seinem *vektoranalytischen*

Modell können Bewegungsmuster in ihre Komponenten zerlegt werden. Punkte, die sich gleichzeitig in eine Richtung bewegen, bilden ein Bezugssystem, das den Rahmen für andere Bewegungen liefert. Wenn die Stimulation aus einer Anzahl verschiedener Bewegungen besteht, bildet sich eine Hierarchie der Bezugssysteme. Der statische Hintergrund ist der äußerste Rahmen. Ein Beispiel: Bei einem stehenden Auto ist die Windschutzscheibe das Bezugssystem für die Bewegung der Scheibenwischer. Letztere bewegen sich relativ zu ihr. Fährt das Auto, wird die Bewegung der Scheibenwischer weiterhin relativ zur Windschutzscheibe gesehen, obwohl sie sich rechnerisch aus der Bewegung des Autos nach vorne und der seitlichen Bewegung der Scheibenwischer zusammensetzt. Die Ausrichtung der Windschutzscheibe dagegen wird relativ zur Straße wahrgenommen. Sie stellt das äußerste Bezugssystem dar. Nach dem Modell von Johansson ist bei biologischen Bewegungen der gemeinsame Bewegungsvektor aller Lichtpunkte die Pendelbewegung des Rumpfes. Sie bildet das invariante Erkennungsmerkmal, das uns hilft, menschliche Bewegungsmuster zu klassifizieren.

Heute werden Proband*innen künstlich erzeugte Punktmuster auf einem Monitor dargeboten. Sie können durch Linien, die die Objektform andeuten, ergänzt werden. Gelegentlich werden unmögliche oder zufällige Bewegungen gezeigt. Obwohl die Versuchsanordnungen verändert wurden, sind die Effekte robust. Die Ergebnisse deuten darauf hin, dass das visuelle System nicht auf die Spezifika des menschlichen Körpers anspricht, sondern auf typische Bewegungsmuster einer Klasse von Organismen (für Wirbeltiere: starre Teile und Gelenke zwischen diesen Teilen). Auf diese Weise können Bewegungen verschiedener Tierarten erkannt werden. Die Formerkennung hat, anders als von Johansson vermutet, eine unterstützende Funktion (für einen Überblick vgl. Blake und Shiffrar, 2007).

Neben einer sorgfältigen phänomenalen Beschreibung der Wahrnehmungsleistung findet mithilfe von fMRT- und PET-Studien eine intensive Suche nach den entsprechenden hirneurophysiologischen Prozessen statt. Die Studien zeigen, dass die Bewegungsinformationen vorwiegend im dorsalen parietalen Pfad verarbeitet werden. Bei der Wahrnehmung biologischer Bewegungen sind Zentren aktiv, die an der Steuerung der eigenen Motorik beteiligt sind. Das könnte als ein Vergleich mit unseren eigenen „Motorprogrammen“ interpretiert werden. Möglicherweise wird getestet, ob wir die gesehene Bewegung selbst ausführen könnten (Prinzip der Zirkularität). Neuere computationale Modelle bemühen sich, Bewegungs- und Forminformation zu integrieren. Nach der Theorie von Giese und Poggio (2003) werden beide Informationen in zwei parallelen aufsteigenden Bahnen analysiert, die sich dem ventralen bzw. dorsalen Strang zuordnen lassen. Kernstück des Modells ist die Annahme, dass menschliche Bewegungen durch gelernte Sequenzen von „Schnappschüssen“ des menschlichen Körpers (Form) und durch charakteristische optische Fließmuster (Bewegung) repräsentiert werden. Prozesse der Hemmung stellen sicher, dass die „Schnappschüsse“ nur in der korrekten zeitlichen Abfolge aktiviert werden können. Diese Annahme macht das Modell empfindlich gegenüber Störungen der zeitlichen Struktur von Bewegungen. Das Ergebnis wurde in vielen empirischen Untersuchungen bestätigt: Ist die Zeitgestalt gestört, sinkt die Erkennungsrate für eine Bewegung drastisch.

Auf der Basis der Prinzipien lässt sich der Wahrnehmungsprozess bei biologischen Bewegungen beschreiben: In der statischen Anordnung (Einzelbild) können keine eindeutigen Gruppierungen gebildet werden. Ein Punktmuster erscheint auch bei längerer Betrachtung zufällig. Erst die Bewegung der Punkte erlaubt die Detektion von stabilen Gruppierungen. Da die Stimulation vieldeutig ist (der distale Reiz besteht lediglich aus hellen Punkten auf einer schwarzen Fläche), werden als Mittel der Ambiguitätsreduktion die Probabilitäten für die Bewegung starrer Teilkörper mit Gelenken herangezogen; diese haben nur bestimmte

Freiheitsgrade in den Bewegungen. Im Laufe des Sehprozesses muss die Frage nach einer regulären Ansicht der Bewegung, die von einem Körper erzeugt werden kann, beantwortet werden. Die Frage ist mit der nach der Multistabilität vergleichbar. Allerdings zielt sie nicht ab auf Veränderungen der Form (Raumgestalt), sondern auf Veränderung in der Zeit (Zeitgestalt). Wenn eine stabile und eindeutige Interpretation des Punktmusters vorgenommen werden kann, die diese Bedingungen erfüllt, ist die Identifikation abgeschlossen. Bei diesem Prozess kommen folgende Heuristiken zur Anwendung: 1. Interpretiere Elemente, die eine gleichförmige Veränderung der Bewegung erleiden, als zusammengehörig. 2. Interpretiere die Richtungs- und Geschwindigkeitsveränderungen so, dass stabile Zeitgestalten entstehen. 3. Interpretiere die zweidimensionalen Ansichten als dreidimensionale Objekte mit natürlichen Bewegungsabläufen. Sicherlich werden nicht alle möglichen Körper gleichzeitig getestet. Im Idealfall gibt es eine Hierarchie weiterer Heuristiken, die die Priorität festlegen. Hier kommt die Formwahrnehmung ins Spiel. Sie erleichtert die Hierarchisierung und schließt Alternativen aus.

Die visuelle Wahrnehmung ist konservativ. Das System behält eine gewonnene Hypothese bei, bis mit ihr dauerhaft unvereinbare Stimulationen auftreten. Sie hält an der ersten Interpretation fest, bis eine alternative, ebenfalls eindeutige Interpretation zur Verfügung steht und die ursprüngliche unmöglich oder zumindest unwahrscheinlich geworden ist. Ambiguität wird unbedingt vermieden. Das lässt sich am Problem der *Identität* beobachten. Werden in einem Johansson-Experiment zwei Bewegungsarten nacheinander von zwei verschiedenen Darstellern ausgeführt (z. B. schreitet der erste Darsteller und der zweite hüpfte), wird eine einzige Person gesehen, die einen Wechsel der Bewegung ausführt. Erst wenn gleichzeitig weitere Parameter geändert werden, die hierzu im Widerspruch stehen (z. B. solche, die darauf schließen lassen, dass es sich im ersten Fall um einen Erwachsenen und im zweiten Fall um ein Kind handelt), wird die Identitätsannahme aufgegeben. Phänomenal gibt es

einen Widerstand gegen die Aufgabe einer eindeutigen Interpretation (vgl. auch den Gestaltfaktor der „objektiven Einstellung“).

7. Kausalität

Die Wahrnehmung von Ereignissen geht über die Wahrnehmung der physikalischen Struktur von Bewegungen hinaus. Menschen sehen nicht nur Bewegungen, sie sehen Beziehungen zwischen den Bewegungen. Eine für das Verhalten besonders wichtige Beziehung ist die Kausalität: Die Bewegung eines Objektes A bewirkt die Bewegung eines Objektes B. Die grundlegende Frage ist: Sehen wir Kausalität oder erschließen wir sie aufgrund früherer Erfahrung? Beide Positionen haben wichtige Befürworter gefunden. Nach David Hume (1777/1966, S. 79) wird Kausalität zwischen Ereignissen immer erschlossen, sie kann nicht gesehen werden:

„It appears that, in single instances of the operation of bodies we never can, by our utmost scrutiny, discover anything but one event following another... All events seem entirely loose and separate. One event follows another, but we can never observe any tie between them. They seem *conjoined*, but never *connected*.“

Albert Michotte entwickelte in den 1940er Jahren eine Versuchsanordnung, mit deren Hilfe er zu zeigen beabsichtigte, dass Kausalität gesehen werden kann (vgl. Abbildung 65). Seine Experimente gliedern das Ereignis in drei Phasen: 1. Annäherung, 2. Kontakt und 3. Ergebnis. Zu sehen sind zwei Objekte X und Y (z. B. Rechtecke), die sich entweder bewegen oder stationär sind. In der ersten Phase (Annäherung) bewegt sich X auf Y zu und Y ist stationär. In der zweiten Phase (Kontakt) berühren sich X und Y und in der dritten Phase (Ergebnis) bewegt sich entweder Y oder es bewegen sich beide Objekte. In Abhängigkeit von der relativen Geschwindigkeit der Objekte in den Phasen eins und drei sind auf der phänomenalen Ebene drei qualitativ verschiedene Ereignisse zu

beobachten: Stoßen (A), Schießen (B) und Schieben (C). Damit ein Stoßen gesehen werden kann, müssen laut Michotte folgende Bedingungen erfüllt sein: Die Bewegung von Y muss innerhalb eines Zeitraums von ca. 200 Millisekunden nach dem Kontakt mit X beginnen. Die Bewegungsrichtung von Y muss annähernd dieselbe sein wie die von X. Die Geschwindigkeit von Y muss langsamer sein als die von X oder darf maximal dieselbe Geschwindigkeit aufweisen. Ist sie deutlich schneller, wird nicht ein Stoßen (A), sondern ein Schießen (B) gesehen. Y erhält von X nur den Impuls für eine eigenständige Bewegung. Für ein Schieben (C) ist es notwendig, dass sich X und Y nach dem Kontakt in dieselbe Bewegungsrichtung und mit derselben Geschwindigkeit weiterbewegen.

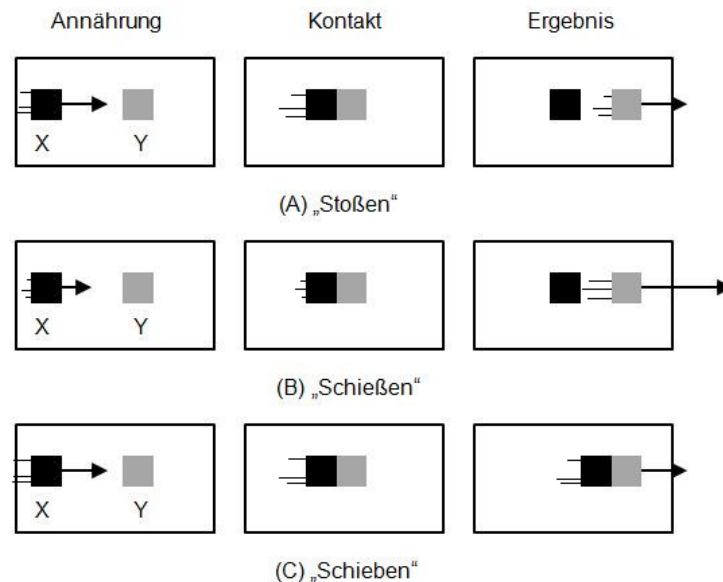


Abbildung 65: Wahrgenommene Kausalität nach Albert Michotte: Wenn ein sich bewegendes Objekt (X) auf ein stationäres Objekt (Y) trifft und das zweite Objekt sich in dieselbe Richtung bewegt, dann wird spontan eine Verursachung dieser zweiten Bewegung wahrgenommen. In Abhängigkeit von den genauen zeitlichen Verhältnissen und der Geschwindigkeit der Objekte lassen sich drei Ereignisse unterscheiden: Stoßen (A), Schießen (B) und Schieben (C).

Heute werden derartige Experimente computergestützt am Monitor durchgeführt. Sowohl die Form der Objekte, als auch ihre Bewegungsrichtung und Geschwindigkeit lassen sich leicht manipulieren.

Runeson (1977) konnte zeigen, dass Beobachter*innen, auch wenn sie keine expliziten Informationen erhalten, in der Lage sind, die relative Masse der Objekte X und Y zu beurteilen, die sich in der Kontakt- oder Kollisionsphase berühren. Nach dem Modell von Runeson wird das Verhältnis der Massen aus dem Verhältnis der Geschwindigkeitsdifferenzen (Anfangs- und Endgeschwindigkeit) der beiden Objekte berechnet. Demnach gäbe es im visuellen System einen Algorithmus, der Geschwindigkeits- und Masseinformation bei Kollisionsereignissen zueinander in Beziehung setzt. Nachfolgende Untersuchungen haben präzisiert, dass die Proband*innen zwar zuverlässig beurteilen können, welches der Objekte schwerer ist, dass sie aber nicht in der Lage sind, die genaue Differenz anzugeben. Es ist auch in diesem Fall davon auszugehen, dass Heuristiken zum Einsatz kommen. Palmer (1999, S. 515) hat sie als Abprall-Heuristik (ricochet heuristic) und als Bombardierungs-Heuristik (clobbering heuristic) bezeichnet. Wenn sich das Objekt X nach dem Zusammenprall mit einer größeren Geschwindigkeit rückwärts bewegt als das Objekt Y vorwärts, ist das stationäre Objekt (Y) schwerer. Wenn Y sich nach der Kollision mit einer größeren Geschwindigkeit bewegt als X in der Annäherungsphase, hat das Objekt X die größere Masse. Das dürfte für viele Wahrnehmungssituationen charakteristisch sein: Es kann zwar ein zuverlässiges ordinales Urteil abgegeben werden (X bewegt sich schneller als Y), aber kein metrisches Urteil (X bewegt sich doppelt so schnell wie Y). Interessant ist der Konflikt, der entsteht, wenn das „abgeprallte“ Objekt sich langsamer bewegt als das getroffene. Die meisten Beobachter*innen haben zwar eine klare Vorstellung, welches der beiden Objekte schwerer ist, stimmen aber in ihrem Urteil untereinander nicht überein.

Obwohl das gewählte Beispiel sehr speziell ist, wird etwas von grundsätzlicher Bedeutung für die Wahrnehmung von Ereignissen deutlich. In der Umgebungswahrnehmung kommt es häufig darauf an, Entscheidungen schnell zu treffen und mit diesen Entscheidungen auf der „sicheren Seite“ zu sein. Die Präzision des Differenzbetrags zwischen den

Alternativen spielt eine untergeordnete Rolle. Einfache Heuristiken und Grenzwerte reichen. Ein Beleg hierfür sind die Untersuchungen zur Wahrnehmung von Kollisionen mit der eigenen Person (time-to-collision). Die experimentelle Situation: Ein Objekt bewegt sich auf die Beobachterin oder den Beobachter zu. Die Fragen: Wird es mit der Beobachterin oder dem Beobachter kollidieren? Muss eine Ausweichbewegung ausgeführt werden? Die Ergebnisse zeigen, dass in fast allen Situationen Kollisionskurse zuverlässig vorhergesagt werden können. In Abhängigkeit von der Beschleunigung des Objektes bestehen jedoch Schwierigkeiten, den genauen Zeitpunkt zu prognostizieren. Dies gilt insbesondere bei nichtlinearen Beschleunigungen. Wenn die Größe der Abbildung des Objektes auf der Retina (Sehwinkel) einen Grenzwert überschreitet, wird eine automatische Ausweichbewegung eingeleitet. In vielen Ballsportarten kommt es deshalb trotz hoher Geschwindigkeiten zu relativ wenigen Gesichtstreffern.

Etwas, das physikalisch ein Kontinuum darstellt (z. B. die Annäherung eines Objektes), zerfällt phänomenal in qualitativ unterschiedliche Ereignisse (z. B. eine gefährliche, eine potentiell gefährliche und eine ungefährliche Situation). Eine solche Klassifikation kann mit Heuristiken leicht erreicht werden. Eine metrische Skalierung der Gefahr (etwa auf einer Skala von 1 bis 100) würde Algorithmen und eine exakte Berechnung voraussetzen. Sie ist für die meisten Entscheidungen und Handlungen nicht notwendig. In vielen Wahrnehmungssituationen dürfte überdies die Datengrundlage unvollständig sein. Eine plausible Option besteht darin, dass das visuelle System die für die Beurteilung der Situation wichtigen Parameter aufgrund der aktuellen Stimulation und der bekannten Wahrscheinlichkeiten für vergleichbare Ereignisse schätzt. Bei dieser Schätzung spielen Heuristiken eine zentrale Rolle. Solche Entscheidungsprozesse lassen sich gut mit dem im ersten Kapitel angesprochenen Bayes'schen Modell beschreiben. Die aktuellen Informationen (Daten) werden mit dem Vorwissen verknüpft und je nach Situation gewichtet. Wenn sich ein Objekt so schnell auf eine Person zu

bewegt, dass eine Reflexhandlung ausgelöst wird, spielt das Vorwissen eine geringe Rolle; die Reaktion wird allein von den Daten gesteuert. Wenn genug Zeit bleibt, ein Objekt auf seine Gefährlichkeit hin zu bewerten, gewinnt das Vorwissen an Einfluss.

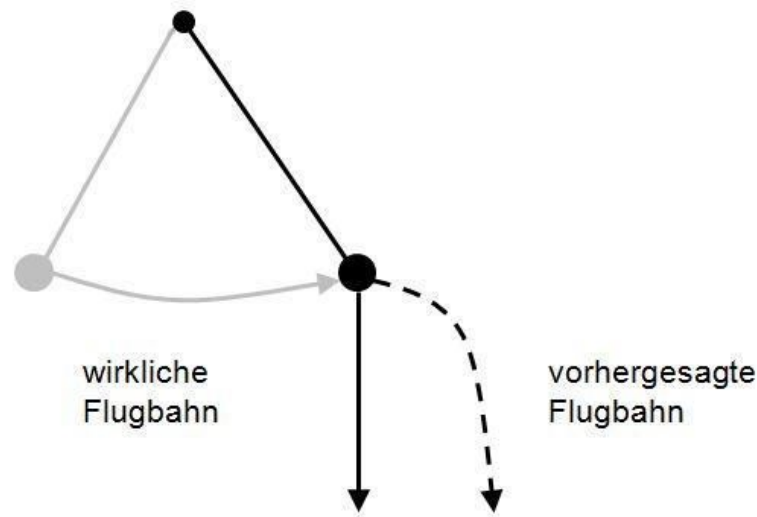


Abbildung 66: Pendelproblem. Erläuterung im Text.

Die Experimente zur Bewegungs- und Ereigniswahrnehmung sind abzugrenzen von Aufgaben zur „intuitiven Physik“. Unter dieser Bezeichnung werden Vorstellungen zusammengefasst, die Lai*innen von physikalischen Vorgängen haben. Eine typische Aufgabe ist das „Pendelproblem“ (vgl. Abbildung 66). Wenn die Bewegung des Pendels an seiner höchsten Stelle unterbrochen wird, wie ist die weitere Flugbahn eines am Pendel befestigten Gewichtes? Viele Menschen glauben, dass das Gewicht in einer gekrümmten Bahn zur Erde fällt. Werden sie aufgefordert eine Zeichnung anzufertigen, skizzieren sie eine Fortführung des bisherigen Verlaufs (vgl. Abbildung 66). In Wirklichkeit fällt das Gewicht senkrecht zu Boden. Die Forschung zur intuitiven Physik hat gezeigt, dass Menschen selbst von einfachen physikalischen Abläufen oft falsche Vorstellungen haben. Das hier verlangte Urteil ist kritisch-phänomenal. Es bezieht sich auf das Wissen, nicht auf das Sehen. Die Antworten sind abhängig von der physikalischen Vorbildung. Klärt man die

Proband*innen auf, verschwinden die Effekte. Anders bei den Experimenten zur Ereignis- und Kausalitätswahrnehmung. Selbst der Experimentator, der die Bewegung der Objekte X und Y programmiert hat, sieht zwingend ein Stoßen, Schieben oder Ziehen. Das Wissen, dass es sich hier ausschließlich um einen zeitlichen, nicht aber um einen kausalen Zusammenhang handelt, hat keinen Einfluss auf die Wahrnehmung des Ereignisses.

8. Visuelle Illusionen

Visuelle Illusionen sind definiert als systematische Abweichungen der phänomenalen von den physikalischen Verhältnissen (für anschauliche Beispiele und ihre Erklärungen vgl. Shapiro und Todorovic, 2017). Eine Wahrnehmung ohne Messgeräte wird mit einer Wahrnehmung unter ihrer Zuhilfenahme verglichen. Wendet man diese Definition auf die Umgebungswahrnehmung an, dann kommt es auch hier zu Illusionen. Viele hängen allerdings entweder direkt mit den Leistungsgrenzen des visuellen Systems zusammen, z. B. bei der Beurteilung von Beschleunigungen, oder basieren auf einer zu hohen Komplexität der Stimulation im Verhältnis zur Darbietungszeit und überschreiten damit die Verarbeitungskapazität. Die meisten bekannten Illusionen jedoch sind Fälle der Bildwahrnehmung.

„Die Illusionen, Mehrdeutigkeiten und andere visuelle Anomalitäten, die von Künstlern und Wahrnehmungspsychologen untersucht wurden, sind keine Manifestationen von willkürlichen Eigenarten, Störimpulsen oder Designfehlern des visuellen Systems des Menschen. Vielmehr liegen die Ursachen dieser Abweichungen in den Tätigkeiten von leistungsstarken und automatischen Inferenzprinzipien, die an die allgemeinen Gegebenheiten der natürlichen Welt gut angepasst sind“ (Shepard 1990, S. 212).

Eine Illusion wie die „gedrehten Tische“ von Shepard (1990), bei der beide Tische exakt dieselben Kantenlängen haben, ist nur in der Bildwahrnehmung möglich (vgl. Abbildung 67). Die perspektivischen Hinweisreize deuten an, dass die längere Kante des Tisches auf der linken Seite in die Tiefe geht, während die längere Kante des rechten Tisches eher orthogonal zur Sichtlinie verläuft. Wenn der linke Tisch sich allerdings in die Tiefe erstreckt, dann muss das korrespondierende Netzhautbild Verkürzungen aufweisen. Die Tatsache, dass die beiden retinalen Bilder der Vierecke, die als Oberseiten von Tischen interpretiert werden, eine identische Länge haben, impliziert, dass die wirkliche Länge des linken Tisches größer sein muss als die wirkliche Länge des rechten Tisches. Umgekehrtes gilt für ihre Breite. Weil die Schlussfolgerungen über Orientierung, Tiefe und Länge automatisch und ohne Beteiligung des Bewusstseins erfolgen, bleibt Wissen über die Mechanismen der Illusion ohne Einfluss auf die Größe des Täuschungsbetrags (vgl. Shepard, 1990). Je weiter die Verarbeitung des retinalen Signals voranschreitet, desto mehr nimmt die Möglichkeit zur bewussten Beeinflussung des Gesehenen zu. Dieses Beispiel macht den großen Anteil der automatischen Verarbeitung im visuellen System deutlich – und ebenso die kleinen „Freiheiten“ des Bewusstseins.

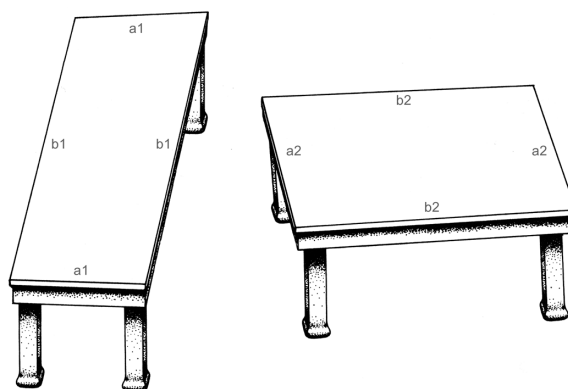


Abbildung 67: Gedrehte Tische (Shepard, 1990). Die Länge der Kante (a1) des linken Tisches entspricht der Länge der Kante (a2) des rechten Tisches und die Länge der Kante (b1) des linken Tisches entspricht der Länge der Kante (b2) des rechten Tisches.

Die Verwechslung der Realität mit dem Bild ist die bemerkenswerteste Illusion bezüglich der Bildwahrnehmung. Es stellt sich die Frage nach den Voraussetzungen für eine solche Verwechslung. Diese Randbedingungen spiegeln ihrerseits die Spezifika der Bildwahrnehmung und den Unterschied zur Umgebungswahrnehmung wider. Sie betreffen sowohl das Bild und seine Beleuchtung als auch die Beobachtungsbedingungen. Hier zehn Regeln für ein Trompe-l'œil (für einzelne Aspekte vgl. auch Pirenne, 1970; Kubovy, 1979; Shepard, 1990):

1. Die dargestellten Gegenstände müssen in der Vorstellung der Betrachterin oder des Betrachters unveränderliche und unbewegliche Gegenstände sein.
2. Die verwendeten Farben sollten den Gedächtnisfarben der Gegenstände entsprechen.
3. Die Art der Darstellung muss detailliert, der Auftrag der Materialien so flach sein, dass die Strukturmerkmale der Textur nicht zu erkennen sind. Es gilt, einen Wechsel zum „Bild als Oberfläche“ zu erschweren.
4. Der Bildraum sollte optisch ein Teilraum des Realraumes sein.
5. Die Lichtquelle des Bildlichts muss außerhalb des Bildes liegen und dem Realraum zugeordnet werden. Der im Bild dargestellte Lichteinfall sollte mit den Lichtverhältnissen im Realraum übereinstimmen. Beide „Beleuchtungen“ müssen einer gemeinsamen Lichtquelle zugeordnet werden können.
6. Der Standpunkt der Betrachterin oder des Betrachters sollte in Entfernung und Winkel möglichst nah an dem für das Bild konzipierten Beobachtungspunkt liegen.
7. Die Betrachtungsbedingungen sind idealerweise stationär und monokular.
8. Der Betrachtungsabstand sollte so groß sein, dass die visuelle Auflösung nicht ausreicht, um die Texturmerkmale zu

identifizieren. Dies betrifft alle Hinweise auf einer zweidimensionalen Oberfläche (z. B. leere Stellen einer Leinwand).

9. Die Beleuchtung darf weder die Textur des Bildes noch den Übergang zwischen Bild und Realraum sichtbar werden lassen. In der Regel können nur mittlere oder geringe Lichtintensitäten verwendet werden. Spiegelnde Reflexionen und Schattenbildungen sind zu vermeiden.
10. Die Präsentation sollte ohne Bildrahmen und möglichst gut integriert an der Stelle im Raum erfolgen, in der sich auch die abgebildeten Gegenstände befinden könnten. Wird ein Rahmen verwendet, muss dieser die Grenzen des Bildes als Oberfläche verdecken und die Funktion eines Fensters ausüben.

Sind diese Voraussetzungen erfüllt, zieht das visuelle System entsprechend der evolutionären Dominanz der Umgebungswahrnehmung den Schluss, dass es sich um eine dreidimensionale Szene handelt. Es ist erstaunlich, dass solche Illusionen nicht häufiger stattfinden. Jede zweidimensionale Abbildung einer Szene ist konsistent mit einer unendlichen Anzahl dreidimensionaler Anordnungen. Die Mehrdeutigkeit wird reduziert, indem das visuelle System die für das Überleben notwendigen Regelmäßigkeiten der physikalischen Welt internalisiert und als Wahrscheinlichkeiten auf die Bildwahrnehmung überträgt. Zu diesen Regelmäßigkeiten gehören u. a. die Texturgradienten der Oberflächen und die Symmetrieeigenschaften der Gegenstände.

Wird der Prozess der Bildwahrnehmung als eine Reduktion von Mehrdeutigkeit rekonstruiert, ergeben sich folgende Schritte (vgl. auch Pirenne, 1970; Shepard, 1990): Auf der Basis der Texturinformation sowie der Informationen über die Grenzen des Bildes ermittelt das visuelle System die Entfernung und den Neigungswinkel der Bildoberfläche. Das Vorgehen entspricht der Bestimmung anderer Oberflächen im visuellen Feld. Das visuelle System benutzt diese Informationen, um die

Verzerrungen der retinalen Abbildung, die sich durch die Differenz zwischen idealem und realem Standort und durch die Bewegung der Betrachterin oder des Betrachters vor dem Bild ergeben, zu kompensieren. Schließlich konstruiert das visuelle System eine korrekte Repräsentation, gerade so, als würde die Betrachterin oder der Betrachter die Szene ohne Neigungswinkel und vom idealen Standort aus sehen.

Kapitel VI: Beständigkeit der visuellen Welt

Das Gesehene entspricht fast nie dem Netzhautbild. Der proximale Reiz ist zwar entscheidend für die Stimulation, aber das Perzept weicht in systematischer Weise davon ab. Unter optimalen Wahrnehmungsbedingungen korrespondiert es mit dem distalen Reiz. Unsere visuelle Repräsentation entspricht in wesentlichen Aspekten der Außenwelt, aber wie wird diese Korrespondenz erreicht? Je nach theoretischer Ausrichtung werden entweder eher die Reize oder eher die Prozesse als verantwortlich angesehen.

Das Konzept der *Invarianzen* nimmt an, dass verborgene Informationen im Reizangebot enthalten sind. Es handelt sich um die Merkmale, die bei einer räumlichen und zeitlichen Veränderung des proximalen Stimulus unverändert bleiben. Beispiele sind der Texturgradient oder das optische Fließmuster. In ihnen lassen sich Parameter finden, die auch bei einer Objektbewegung oder sich bewegenden Beobachter*innen konstant bleiben. Die Stimulation enthält Variablen höherer Ordnung, die die Eindeutigkeit und Stabilität des Perzeptes garantieren. Sie müssen vom visuellen System entdeckt werden. Die Detektionsleistung wird über spezialisierte Neurone erreicht. Die Ursache für die Stabilität liegt auf der Reizseite.

Das Konzept der *Korrekturen* betont die konstruktiven Prozesse bei der Informationsverarbeitung. Ausgehend vom retinalen Bild werden auf den Verarbeitungsstufen unterschiedliche Formen der Korrektur vorgenommen. Sie betreffen zunächst lokale, später auch globale Merkmale. Die Ursache für die Stabilität der visuellen Welt liegt in den Prozessen. Die Helligkeitswahrnehmung ist hierfür ein Beispiel. Um eine natürliche Beleuchtungssituation mit ihrem Wechsel von Licht und Schatten adäquat wiederzugeben, reicht die aktuelle Stimulation nicht aus. Zusätzliche Inferenzprozesse sind notwendig; Heuristiken, die die Richtung der Korrekturen vorgeben. Über Hemmung und Aktivierung

wird sichergestellt, dass schließlich eine Lösung präferiert wird, die die größtmögliche Übereinstimmung mit den Regelmäßigkeiten der Außenwelt aufweist.

Unter einer probabilistischen Auffassung der Wahrnehmung sind beide Konzepte nicht so gegensätzlich wie sie zunächst erscheinen. Denn die Prozesse haben sich in Anpassung an die Lebensbedingungen und damit an die Reizverhältnisse entwickelt. Bei vielen Phänomenen dürften beide Erklärungsansätze eine Rolle spielen. Lediglich ihre relative Bedeutung variiert. Die Zuverlässigkeit des visuellen Systems unter den Bedingungen der Umgebungswahrnehmung suggeriert ein deterministisches Verhältnis von Wahrnehmung und Umgebung. Ein solcher Determinismus ist allerdings eine Illusion. Die meisten Wahrnehmungsleistungen sind das Resultat einer Kombination vieler Informationsquellen.

Das wird bei den *Konstanzleistungen* besonders deutlich. Jede einzelne Informationsquelle liefert eine ungefähre Vorhersage in Bezug auf die anstehende Aufgabe. Die wahrgenommene Größe und die Form eines Objektes basieren auf einer Kombination unterschiedlicher Quellen wie der Verteilung von Licht und Schatten, der Textur, Verdeckung, Bewegung und binokularer Disparität. Hinzu kommt, dass die sensorische Analyse aufgrund von Zufälligkeiten bei der Lichtabsorption und aufgrund chemischer Ereignisse in den Photorezeptoren immer verrauscht ist. Die adäquate Beschreibung natürlicher Stimuli ist daher eine stochastische (für Methoden einer statistischen Analyse von natürlichen Szenen vgl. Geiser, 2008). Jede einzelne Informationsquelle steht in einem probabilistischen Zusammenhang zur physikalischen Größe und Form des Objektes. Zusammen bilden sie eine zuverlässige Grundlage für die visuelle Welt.

Der Begriff der *Konstanz* bezieht sich auf die phänomenale Ebene. Er verweist darauf, dass es im Erleben des Wahrnehmungssubjektes eine Stabilität gibt, obwohl sich die zugehörige Stimulation verändert hat.

Präziser formuliert: Quantitative Veränderungen in der Stimulation führen nicht zu einem qualitativ neuen Wahrnehmungserlebnis. Unter den Bedingungen der Umgebungswahrnehmung ist diese Situation die Regel. Invarianzen auf der Reizseite und Korrekturmechanismen bei den Prozessen garantieren die Stabilität der visuellen Welt trotz permanenter Veränderung der proximalen Stimuli. Für das Verständnis des Sehens kommt den Konstanzphänomenen eine besondere Bedeutung zu. Die visuelle Welt wäre ohne Konstanz chaotisch.

Konstanzphänomene treten grundsätzlich in allen Bereichen der Wahrnehmung auf. In der Forschung wurden vor allem Helligkeits-, Farb-, Größen- und Formkonstanzen intensiv untersucht. Aus zugehörigen Experimenten ist bekannt, dass extrem künstliche und sehr rigide Beobachtungsbedingungen geschaffen werden müssen, damit Konstanz zusammenbrechen. Erst unter stark verarmter Wahrnehmung (reduzierte Beleuchtungsverhältnisse, fehlende Vergleichsreize oder sehr kurze Darbietungszeiten) bestimmen die Eigenschaften des proximalen Reizes das Perzept.

1. Helligkeits- und Farbkonstanz

Die wahrgenommene Helligkeit eines Gegenstandes scheint direkt von der Menge des Lichts abzuhängen, das von ihm ausgeht. Oberflächen, die nicht selbst Lichtquellen sind, reflektieren bzw. absorbieren das auftreffende Licht in unterschiedlichen Anteilen. Weiße Flächen reflektieren etwa 90 Prozent und erscheinen daher sehr hell, während tiefschwarze Flächen weniger als fünf Prozent reflektieren. Sie wirken entsprechend dunkel. Natürlich hängt die Menge des reflektierten Lichts nicht allein vom Reflexionsgrad der Oberfläche, sondern auch von der Intensität des auftreffenden Lichts ab. Sie kann sehr stark – zwischen dem Sonnenlicht und dem Licht einer Taschenlampe – variieren. Die Intensitätsunterschiede sind erheblich größer als die

Reflexionsunterschiede zwischen hellen und dunklen Flächen. Von einem schwarzen Blatt Papier geht im Sonnenlicht objektiv wesentlich mehr Licht aus als von einem weißen Blatt Papier im schwachen Licht einer Glühlampe. Trotz der Luminanzunterschiede wird das schwarze Papier auch im Sonnenlicht als Schwarz wahrgenommen, es wird durch die hohe Lichtstärke nicht plötzlich Grau oder Weiß. Ebenso behält das weiße Papier unter unzureichenden Beleuchtungsverhältnissen phänomenal seine weiße Farbe, es erscheint lediglich „schwach beleuchtet“. Die wahrgenommene Helligkeit ändert sich nicht linear mit Veränderungen der objektiven Lichtmenge, sondern bleibt für die betreffende Oberfläche über einen weiten Bereich konstant. Erst in den Grenzbereichen einer extrem schwachen oder einer extrem starken Beleuchtung bricht diese Helligkeitskonstanz zusammen. Ein wichtiger Aspekt für die Wirksamkeit dieses Mechanismus ist die Unterscheidung zwischen Objektveränderungen und Beleuchtungsveränderungen im visuellen Feld. Die Wahrnehmung geht durch ihre Adaptation an die terrestrischen Verhältnisse zunächst davon aus, dass die Reflexionseigenschaften der Objekte konstant bleiben, während sich die Beleuchtungsverhältnisse ändern (vgl. Abbildung 68). Die Helligkeitskonstanz hat eine Ordnungsfunktion. Sie macht die Wahrnehmung weitgehend unabhängig von lokalen Lichtintensitäten, indem sie nicht die absoluten, sondern die relativen Helligkeiten im Gesichtsfeld berücksichtigt.



Abbildung 68: Helligkeitskonstanz in natürlicher Umgebung. Trotz unterschiedlicher Leuchtdichten aufgrund von Licht und Schatten wird der Lattenzaun als ein einheitliches Objekt wahrgenommen (Foto von Joel Meyerowitz).

In zahlreichen Experimenten wurden diverse Faktoren der Helligkeitskonstanz ermittelt. Die Versuchsanordnungen sehen meist folgendermaßen aus: Die Proband*innen sehen einen Standardreiz (z. B. eine rechteckige Fläche in einer mittleren Graustufe) und eine Reihe von Vergleichsreizen (mit unterschiedlichen Graustufen). Die Beleuchtung für Standardreiz und Vergleichsreize kann getrennt variiert werden. Aufgabe der Proband*innen ist es, denjenigen Vergleichsreiz auszuwählen, der ebenso hell wie der Standardreiz erscheint. Auf diese Weise lässt sich feststellen, inwieweit die Beleuchtungsstärke oder der Reflexionsgrad des distalen Stimulus die Einschätzung bestimmt. Fast immer zeigt sich, dass ein Vergleichsreiz mit ähnlichem oder identischem Reflexionsgrad ausgewählt wird, ungeachtet der verschiedenen Beleuchtungsstärken. Allerdings stellt sich das Ergebnis nur dann ein, wenn die Beobachter*innen gleichzeitig eine Anzahl von Vergleichsreizen zur Verfügung haben, das Gesichtsfeld verschiedenartige

Helligkeitsrelationen beinhaltet. Ist das nicht der Fall, stützen die Proband*innen ihr Urteil ausschließlich auf die Lichtmenge, die die Retina erreicht. Helligkeitskonstanz ist immer kontextgebunden. Ohne einen Vergleich, der Informationen über die Lichtverhältnisse in der Umgebung bietet, gibt es weder algorithmische noch heuristische Möglichkeiten, die Beleuchtungsstärke von der Reflexionseigenschaft des Objektes zu trennen.

Die Helligkeitskonstanz hat Grenzen. Wenn sich die Lichtintensitäten im Gesichtsfeld sehr stark unterscheiden (im Verhältnis größer als zehn zu eins), lässt die Wahrnehmungskonstanz nach. In der natürlichen Umgebung oder auch bei der Bildbetrachtung treten diese Ereignisse kaum auf; nur in Laborexperimenten lassen sie sich herstellen.

Die Bedeutung der Helligkeitskonstanz für die Bildwahrnehmung kann kaum hoch genug eingeschätzt werden. Ohne die Helligkeitskonstanz gibt es, überspitzt formuliert, kein Bildlicht. Sie stellt sicher, dass die Betrachterin oder der Betrachter zwischen der Beleuchtung im Raum und dem im Bild dargestellten Licht unterscheiden kann. Das Bildlicht ist weitgehend unabhängig von der Umgebungsbeleuchtung. Das Bild einer sonnendurchfluteten Landschaft erscheint auch unter den Bedingungen einer stark reduzierten Halogenbeleuchtung im Museum als „südländische Sommerlandschaft“.

Am Beispiel von Edward Hoppers (1882-1967) „Sonne in einem leeren Zimmer“ von 1963 (vgl. Abbildung 69) wird die Bedeutung der Helligkeits- und der mit ihr eng verwandten Farbkonstanz im Einzelnen gezeigt. Das Bild zeigt ein leeres Zimmer. Durch ein Fenster auf der rechten Seite fällt Licht. Farbton und Einfallswinkel lassen auf einen Spätnachmittag schließen. Gezeigt werden eine vorspringende Wand und ein Fußboden. Teile der Wand werden unterschiedlich beleuchtet, und es resultieren starke Helligkeitskontraste. Wären die Abbildungsverhältnisse auf der Retina entscheidend, müsste die Wand in unterschiedlichen

Farben, ähnlich einem Teppich mit geometrischem Muster, wahrgenommen werden. Tatsächlich wird eine einzige Wandfarbe mit Unterschieden in Beleuchtung und räumlicher Tiefe gesehen. Der Eindruck ist zwingend. Er kann nur ausgeschaltet werden, indem Teile des Bildes abgedeckt werden.

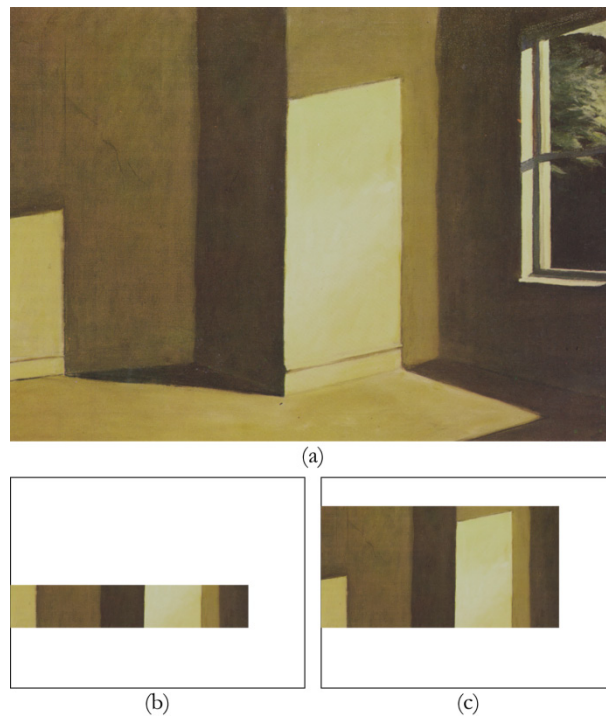


Abbildung 69: Edward Hopper, Sonne in einem leeren Zimmer, 1963. Im Ausschnitt (b) können keine Bildebenen unterschieden werden. Es entsteht lediglich ein Nebeneinander von Farbflächen. Ausschnitt (c) ist mehrdeutig: Wird eine räumliche Anordnung gesehen, wird zumeist auch Licht und Schatten gesehen.

Abbildung 69 zeigt zusätzlich zum Original zwei verschieden große Bildausschnitte. Im Ausschnitt (b) fehlt sowohl die räumliche Tiefe als auch die Differenzierungsmöglichkeit zwischen Licht und Schatten. Diese Situation entspricht den lokalen Abbildungsverhältnissen auf der Retina. Die Detailaufnahme verweist auf die Textur und damit einseitig auf das „Bild als Oberfläche“. In Abbildung 69 (c) ist der Ausschnitt größer, die Situation wird mehrdeutig. Eine räumliche Interpretation ist möglich, aber nicht zwingend. Kommt es zu einer Unterscheidung verschiedener

Bildebenen, ist die Voraussetzung für eine Interpretation der Farb- und Helligkeitsunterschiede als Bildlicht gegeben. Eine illusionistische Wirkung setzt die zuverlässige Trennung von Bildebenen voraus. In dem Augenblick, wo die Wand in (c) nicht räumlich gesehen wird, bricht die Farb- bzw. Helligkeitskonstanz zusammen. Das geschieht, obwohl sich die lokalen Verhältnisse auf der Retina nicht verändert haben.

Der Helligkeit vergleichbar gibt es auch bei der Farbwahrnehmung Konstanzleistungen. Die spektrale Zusammensetzung der das Auge erreichenden Lichtstrahlen wird einerseits von den Pigmenteigenschaften der Oberflächen, andererseits von den Merkmalen der Beleuchtung bestimmt. Würde sich die Farbwahrnehmung ausschließlich auf die sensorischen Informationen stützen, wäre sie ständigen Schwankungen unterworfen. Tatsächlich wird die Farbe innerhalb bestimmter Grenzen als in ihrer Qualität identisch erlebt. Eine rote Oberfläche behält anschaulich auch dann ihren roten Farbton, wenn sie unterschiedlichen spektralen Verhältnissen (z. B. Sonnenlicht und Neonlicht) ausgesetzt ist und damit der Wellenlängenbereich, der auf die Retina trifft, objektiv variiert. Der Farbeindruck ist, wie der Helligkeitseindruck, weitgehend unabhängig von den aktuellen Beleuchtungsverhältnissen. In der Regel wird eine annähernde Konstanz erreicht: Bei einer Veränderung der Beleuchtung ändert sich der wahrgenommene Farbton nicht in dem Ausmaß, wie es aufgrund der Verschiebung des Lichtspektrums zu erwarten wäre.

Die spezifischen Probleme der Bildwahrnehmung im Vergleich zur Umgebungswahrnehmung können an der Gegenüberstellung von Fotografien eines Spaziergangs durch eine Wiesenlandschaft zu drei Tageszeiten (Morgendämmerung, Mittagszeit und Abenddämmerung) illustriert werden (vgl. Abbildung 70). Das Gras wird trotz wechselnder Beleuchtungsverhältnisse immer grün sein. Die Lichtfarbe wird jedoch als veränderlich wahrgenommen. Auf den Fotografien ist der Farbeindruck der Wiese von den Tageszeiten abhängig. Die Verschiebungen der

spektralen Zusammensetzung, vor allem entlang der blau-gelben Gegenfarbenachse, sind auffällig. In der natürlichen Situation tritt eine „Korrektur“ über die Farbkonstanz auf. Bei der Betrachtung der Fotografien wirkt dieselbe Wiese verfärbt. Einzeln präsentiert können vor allem die Aufnahmen vom frühen Morgen und späten Abend unnatürlich wirken.

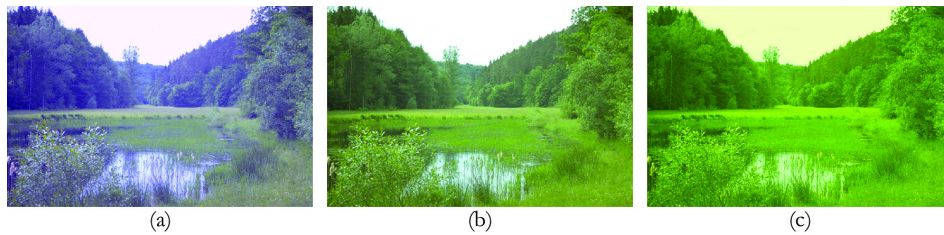


Abbildung 70: Fotos einer Wiesenlandschaft zu drei Tageszeiten: (a) Morgendämmerung, (b) Mittagszeit und (c) Abenddämmerung (Gegenfurtner, 2003).

Für die Farbkonstanz zum Teil verantwortlich ist die Farbadaptation. Bei längerer Einwirkung einer chromatischen Farbe ändert sich die Empfindlichkeit der Rezeptorzellen. Das Auge wird für diese Wellenlängen weniger empfindlich. Das adaptierte Auge nimmt z. B. Rot und auch Orange als weniger gesättigt wahr. Ein Beispiel für die Bedeutung der Adaptationsmechanismen ist der Wechsel von einem Raum mit Tageslicht zu einem Raum mit künstlicher Beleuchtung. Angenommen, dass bei der künstlichen Beleuchtung das langwellige Licht vorherrschend ist. Dann führt die Adaptation an das Kunstlicht dazu, dass das Auge weniger sensitiv für langwelliges Licht wird. Da jedoch mehr langwelliges Licht von den Gegenständen reflektiert wird, kompensieren sich die Ergebnisse der beiden Vorgänge. Die Farben werden als beständig wahrgenommen.

Farbkonstanz ist bereits kurz, nachdem der Raum betreten wurde, zu beobachten. Es müssen zusätzliche Faktoren wirksam sein. Besonders wichtig scheint die Umgebung der Farbfläche. Die Abdeckung des Umfeldes zeigt, dass die Farbkonstanz deutlich nachlässt, wenn

ausschließlich eine monochrome Fläche im Gesichtsfeld angeboten wird. Die Oberfläche nimmt dann die Farbe der Beleuchtung an. Eine grüne Fläche wird bei langwelligem Licht als gelblicher wahrgenommen als bei Tageslicht. Die Farbkonstanz ist maximal, wenn eine Farbfläche von einer Vielzahl chromatischer Farben umgeben ist. Einiges spricht dafür, dass die Stabilität der Farbwahrnehmung auf ähnlichen Mechanismen beruht wie die Helligkeitskonstanz. Adh mar Gelb hatte bereits um 1930 vermutet, dass wir Farbgradienten f r alle farbigen Fl chen im Gesichtsfeld bestimmen, auf deren Basis sich physikalisch unterschiedliche Farbinformationen als „invariant“ beurteilen lassen. In diesem Fall m sste die Farbkonstanz „versagen“, wenn keine weiteren Farben im Gesichtsfeld zu sehen sind. Tats chlich l sst sich ein fast vollst ndiger Verlust unter entsprechenden experimentellen Bedingungen nachweisen. Die Stabilit t der Farbwahrnehmung setzt Farbrelationen im visuellen Feld voraus.

F r Experimente zur Farbkonstanz werden h ufig sogenannte Mondrian-Vorlagen verwendet. Semir Zeki (1983a, 1983b) hat in Untersuchungen an Affen gezeigt, dass es in V4 (Temporallappen) Neurone gibt, die f r die Leistung der Farbkonstanz bedeutsam sind. Er variierte die Beleuchtung eines Quadrates in einer Mondrian-Vorlage, das vorwiegend Licht mittlerer Wellenl nge reflektiert (Gr n), indem er es abwechselnd mit st rker mittelwelligem oder st rker langwelligem Licht beleuchtete. Die Impulsrate des Neurons ver nderte sich nicht dem physikalischen Reiz entsprechend, sondern parallel zum eigenen Farbeindruck des Versuchsleiters bei der Betrachtung des Gesamtmusters. Das entspricht den Vorhersagen der Farbkonstanz.

Zur Erkl rung der Farbkonstanz ist gelegentlich auf die Bedeutung von Ged chtnisfarben verwiesen worden. Als Ged chtnisfarbe wird das vorab erworbene Wissen  ber eine charakteristische Zuordnung von Gegenstand und Farbe bezeichnet. Gras wird unabh ngig von der Beleuchtung immer als Gr n gesehen, weil es im Ged chtnis mit Gr n fest assoziiert wird.  hnlich vermutet Rock (1977), dass die Farbkonstanz im

Sinne von Helmholtz auf unbewussten Schlüssen basiert, sodass die wahre Farbe aus dem Gedächtnis ergänzt werden kann. Gegen diese Auffassung spricht, dass Farbkonstanz bei künstlichem, unvertrautem Labormaterial (z. B. Farbscheiben), bei dem ein Vorwissen nicht verfügbar sein kann, ebenfalls auftritt. Offensichtlich benutzt das visuelle System Hinweisreize, um eine „stabile Situation“ herzustellen (Kraft und Brainard, 1999). Wichtig scheint zu sein, dass die Mittelwertfarbe in einer Szene stark von der Beleuchtung abhängig ist und dass die Beleuchtungsveränderung unter natürlichen Bedingungen graduell und kontinuierlich erfolgt, während ein Wechsel der Reflektanz eher abrupt oder dichotom ist (ein Gegenstand ist vorhanden oder nicht). Die Unterschiede beim charakteristischen Verlauf von Beleuchtungs- und Reflektanzänderungen können vom visuellen System dann effektiv genutzt werden, wenn mehrere große und homogene Flächen für einen Vergleich zur Verfügung stehen, wenn in einer Szene z. B. neben den Informationen über die Wiese die Farbe des Himmels, eines Teichs oder des Waldrandes vorhanden sind.

Die Grenzen der Farbkonstanz zeigen sich dort, wo die Beleuchtung starke Anteile monochromatischen Lichts aufweist. In Diskotheken etwa, wo von starken Deckenstrahlern monochromatisches Licht erzeugt wird, wechselt je nach Standort der Tänzer die Farbe ihrer Kleidung. Dies verdeutlicht, wie gering unter solchen Verhältnissen die Einflüsse des Gedächtnisses auf die Farbwahrnehmung sind. Das Wissen, welchen Farbton die Kleidung eigentlich hat, spielt keine Rolle. Die spektrale Zusammensetzung des reflektierten Lichts ist so stark beeinträchtigt, dass die Korrekturmechanismen zu keiner Kompensation führen. Bei komplexen visuellen Szenen, mit denen wir es in der Umgebungswahrnehmung häufig zu tun haben, ist der aktuelle Farbeindruck nicht nur durch die drei Parameter Farbton, Sättigung und Helligkeit bestimmt. Die Kontextfaktoren – die Form des Objektes, die Textur der Oberfläche oder die Orientierung relativ zur Lichtquelle – haben eine entscheidende Bedeutung (für einen Überblick vgl. Shevell und Kingdom, 2008).

2. Größenkonstanz

Die meisten Gegenstände der visuellen Welt verändern ihre tatsächliche Größe kaum oder nur über einen sehr langen Zeitraum hinweg. Die physikalische Größenstabilität zeigt sich auch phänomenal: Die wahrgenommene Größe eines Gegenstandes bleibt unter verschiedenen Beobachtungsbedingungen erhalten. Die Konstanz tritt auf, obwohl die retinale Abbildung in Abhängigkeit von der Distanz variiert. Entfernte Objekte werden bei gleicher physikalischer Größe auf der Netzhaut kleiner abgebildet als nähere, da der zugehörige Sehwinkel abnimmt. Diesen Zusammenhang kann man sich bei Objekten, deren tatsächliche Größe bekannt ist, zunutze machen, um ihre Entfernung zu schätzen. Stützte sich die Größenwahrnehmung ausschließlich auf das Netzhautbild, würde die wahrgenommene Länge und Breite eines Gegenstandes je nach Entfernung variieren. Ein sich entfernender Fußgänger würde mit der Entfernung kleiner wirken. Tatsächlich tritt dieser Effekt nicht auf. Die Größe des Fußgängers wird bei einer Veränderung der Entfernung als konstant erlebt. Offensichtlich werden bei Längen- und Breitenwahrnehmungen andere Informationen als die Ausdehnung des retinalen Bildes herangezogen (vgl. Abbildung 71).

Eine Erklärung für das Auftreten von Größenkonstanz hat Gibson (1950) vorgeschlagen. Seiner Meinung nach basiert sie auf den Größenrelationen zwischen den Elementen des visuellen Feldes. Ursache wären invariante Reizmerkmale. Wenn zwei Gegenstände die gleichen Anteile einer regelmäßigen Textur verdecken, erscheinen sie als gleich groß, denn nicht nur die Abbildung eines entfernteren Gegenstandes, sondern auch die der zugehörigen Textur wird im gleichen Verhältnis auf der Netzhaut kleiner bzw. dichter. Gibson zufolge stammt die Information über die Größe der wahrgenommenen Gegenstände unmittelbar aus der optischen Anordnung und muss nicht induktiv geschlossen werden. Die Relationen, die sich z. B. im Texturgradienten einer Fläche zeigen, sind Invarianten,

die auch bei einer Objekt- oder Eigenbewegung die direkte Wahrnehmung der tatsächlichen Größe erlauben.



Abbildung 71: Ein Spiel mit der Größenkonstanz. Das Figurenpaar links im Vordergrund nimmt den gleichen Sehwinkel ein wie das Figurenpaar im Hintergrund.

Anders als die Helligkeits- und Farbkonstanz tritt Größenkonstanz auch dann auf, wenn keine Informationen über Größenverhältnisse auf der Basis eines Vergleichs zwischen verschiedenen Elementen verfügbar sind. Es ist zu vermuten, dass die Relationen im visuellen Feld alleine zur Erklärung der Konstanzleistung nicht ausreichen. Alternative Erklärungsansätze nehmen an, dass bei der Größenschätzung zwei Arten von Information herangezogen und miteinander in Beziehung gesetzt werden: die Größe des Netzhautbildes und die wahrgenommene Entfernung des Gegenstandes vom eigenen Standort. Beide Variablen verhalten sich reziprok zueinander: Je größer die Entfernung, desto kleiner wird das Netzhautbild und umgekehrt. Werden beide Faktoren bei der Größenschätzung berücksichtigt, ergibt sich trotz Variation der

Entfernung ein konstanter Schätzwert, weil die veränderliche Netzhautinformation durch den Entfernungsparameter korrigiert werden kann. Voraussetzung ist, dass die Entfernungsschätzung unabhängig von der Größe des retinalen Bildes erfolgt. Als räumliche Hinweisreize werden vor allem die binokulare Disparität und, bei kleinen Entfernungen, die Okulomotorik diskutiert.

Die Bedeutung der Information über räumliche Entfernungen haben Holway und Boring (1941) in einem klassischen Experiment untersucht. Sie gingen von der Überlegung aus, dass die Größenwahrnehmung prinzipiell von zwei Determinanten bestimmt werden kann: 1. vom Sehwinkel, also von der Größe des Netzhautbildes und 2. von der tatsächlichen Größe eines Objekts. Holway und Boring interessierte, unter welchen Bedingungen welches Ausmaß an Konstanz auftritt. Ihre Proband*innen saßen am Schnittpunkt zweier schwach erleuchteter Gänge. In einem Gang wurde mithilfe eines Projektors und einer Leinwand der „Standardreiz“ (z. B. Dreieck oder Kreis) dargeboten. Der Abstand des distalen Reizes zur Beobachterin oder zum Beobachter wurde ebenso variiert wie seine Größe, sodass bei Entfernungen zwischen drei und fünfunddreißig Metern der gleiche Sehwinkel resultierte. Im zweiten Gang befand sich in unveränderlichem Abstand eine Apparatur, mit der ein Vergleichsreiz erzeugt wurde. Hier bestand nur die Möglichkeit, die Größe des Bildes einzustellen. Aufgabe der Proband*innen war es, die Ausdehnung des Vergleichsreizes auf die wahrgenommene Größe des Standardreizes zu regulieren. Holway und Boring führten vier Bedingungen ein: (1) normales binokulares Sehen, (2) monokulares Sehen, (3) monokulares Sehen durch eine zusätzliche Blende und (4) monokulares Sehen durch eine Blende und einen schwarzen Tunnel, der alle Lichtreflexionen verhinderte. Zusätzliche Tiefenhinweise sollten Schritt für Schritt unterbunden werden. Wenn bei der Wahrnehmung des Standardreizes die Größenkonstanz wirksam ist, ist zu erwarten, dass der Vergleichsreiz in Abhängigkeit von der wahrgenommenen Entfernung des Standardreizes unterschiedlich eingestellt wird. Tatsächlich trat dieses

Resultat aber nur unter der Bedingung (1), normales binokulares Sehen, auf. Bei den übrigen, stärker eingeschränkten Bedingungen richteten sich die Proband*innen beim Größenurteil stärker nach dem Netzhautbild bzw. dem Sehwinkel. Besonders deutlich wurde es in der Bedingung (4), in der alle Tiefenhinweise eliminiert sind. Holway und Boring schlossen aus diesem Ergebnis, dass unter normalen Beobachtungsbedingungen die Entfernung mit der Größe verrechnet wird. Je weniger Hinweise räumlicher Tiefe verfügbar sind, desto stärker stützt sich das Urteil auf das Netzhautbild.

Größenkonstanz ist offensichtlich kein universelles Phänomen. In manchen Situationen verliert der Mechanismus seine Wirkung: Von einem Wolkenkratzer aus betrachtet sehen Autos und Menschen wie Spielzeug aus. Aus dem Fenster eines Flugzeugs verlieren auch Schiffe und ganze Städte ihre wahre Größe. Solche Effekte, die gelegentlich mit dem Begriff Unterkonstanz bezeichnet werden, könnten darauf beruhen, dass der für die Größenkonstanz wichtige Faktor „Entfernung“ nicht mehr valide eingeschätzt werden kann. Korrekte Entfernungsschätzungen sind zum einen in ihrer Reichweite begrenzt, zum anderen abhängig von entsprechenden Hinweisen im Gesichtsfeld. Wenn eine zuverlässige Entfernungsschätzung unmöglich ist, wird die Größe des retinalen Bildes zum entscheidenden Bestimmungsfaktor, die Korrekturmechanismen sind unwirksam.

3. Formkonstanz

Die Form eines Objektes nehmen wir als gleichbleibend wahr, obwohl die optische Projektion auf der Netzhaut in Abhängigkeit vom Beobachtungswinkel variiert. Die Variationen können sowohl durch eine Objekt- als auch durch eine Eigenbewegung verursacht werden. Eine Fläche wird auch dann als rechtwinklig wahrgenommen, wenn sie mehr oder weniger geneigt ist und das Netzhautbild einem Trapez gleicht (vgl. Abbildung 72). Auch die Formwahrnehmung stützt sich offenbar nicht

ausschließlich auf das Netzhautbild. Um die Form eines Gegenstandes unabhängig vom retinalen Bild wahrnehmen zu können, müssen Informationen über seine räumliche Orientierung bzw. seinen Dreh- oder Neigungswinkel relativ zur Position der Beobachterin oder des Beobachters als weitere Parameter einbezogen werden. Ähnlich, wie bei der Größenkonstanz die Ausdehnung des Netzhautbildes und die Entfernung des Gegenstandes miteinander in Beziehung gesetzt werden, lässt sich die „wirkliche“ Form aus der Relation zwischen der Form des Netzhautbildes und der räumlichen Orientierung des Gegenstandes bestimmen.

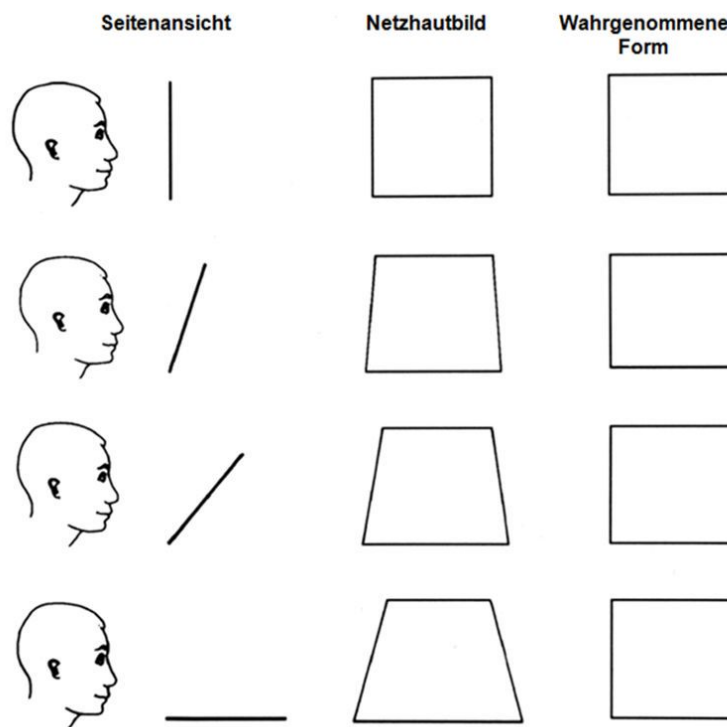


Abbildung 72: Formkonstanz. Unterschiedliche Orientierungen eines Quadrats im Raum führen zu unterschiedlichen Netzhautbildern. Die wahrgenommene Form ist jedoch jeweils ein Quadrat (Lindsay und Norman, 1977).

Voraussetzung ist, dass sich die entsprechenden Informationen dem visuellen Feld entnehmen lassen. Um die räumliche Lage eines Wahrnehmungsobjektes beurteilen zu können, darf es nicht isoliert

erscheinen. Es muss im Zusammenhang mit anderen visuellen Strukturen gesehen werden. Wichtig sind strukturierte Hintergründe, andere Vergleichsobjekte sowie die Textur der Oberfläche des Wahrnehmungsobjektes. Wie bei allen anderen Konstanzphänomenen gilt auch bei der Formkonstanz, dass sie ohne den Kontext als Informationsquelle nicht denkbar wäre und dass die Komplexität der Stimulation eine Voraussetzung für die Stabilität des Perzeptes ist.

In Laboruntersuchungen (Beck und Gibson, 1955; Leibowitz und Bourne, 1956) konnte gezeigt werden, dass die Formkonstanz weitgehend zusammenbricht, wenn entsprechende Hinweisreize im visuellen Feld ausgeblendet werden. Beim Fehlen solcher Informationen nähert sich die gesehene Form dann der Projektion auf der Netzhaut. Formkonstanz kann diesen Überlegungen zufolge nur dann auftreten, wenn das retinale Bild und die räumliche Orientierung des Objekts in einem festen Verhältnis zueinander stehen. Ändert sich die Form des Netzhautbildes, muss gleichzeitig eine Veränderung der Raumlage, z. B. infolge eigener Bewegung, wahrgenommen werden. Gibson (1950) geht von einem solchen Invarianzprinzip aus. Aus seiner Sicht lässt sich das Auftreten von Formkonstanz durch die im visuellen Feld vorhandenen Informationen hinreichend begründen (z. B. durch Überlagerungen von Konturen sowie durch Oberflächen- und Hintergrundtextur).

Spätere Untersuchungen sprechen dafür, dass das Auftreten von Formkonstanz auch von Top-down-Prozessen abhängig ist: Wurden die Proband*innen ausdrücklich instruiert, ihre Wahrnehmung auf die Form des Netzhautbildes zu stützen, reduzierte sich die wahrgenommene Konstanz erheblich. Diese Resultate sprechen für eine Auffassung, die die Formkonstanz als Ergebnis höherer kognitiver Prozesse sieht und sie nicht allein durch die Merkmale der visuellen Umgebung zu erklären versucht.

Aus dieser Forschungsrichtung stammt die weitergehende Überlegung, dass Formkonstanz im Wesentlichen ein Ergebnis vorhergehender

Erfahrungen sein könnte. Die charakteristischen Formen für eine Vielzahl von Objekten werden im Gedächtnis gespeichert und auf neue Wahrnehmungsinhalte angewendet. Wird ein Objekt durch den Vergleich mit Gedächtnisinhalten identifiziert, ist auch seine Form unabhängig von den Merkmalen des Netzhautbildes bestimmt. Gegen die Theorie der gelernten Konstanz spricht, dass das Phänomen auch bei unbekanntem, künstlichen Gegenständen im Labor und bei Neugeborenen auftritt. Eine Reihe der die Formkonstanz beeinflussenden Faktoren ist bekannt, ihr genaues Zusammenspiel jedoch nicht geklärt.

Abbildung 73 zeigt den amerikanischen Präsidenten Richard Nixon vor seinem eigenen Wahlkampfplakat. Unabhängig vom Standort erscheint der Redner auf dem Foto korrekt wiedergegeben. Irritierend ist die verzerrende Darstellung Nixons auf dem Plakat.



Abbildung 73: Nixon vor seinem Wahlkampfplakat. Die Fotografie einer Fotografie (Time Magazine vom 19. März 1968).

Warum führt das visuelle System das eine Mal eine Korrektur durch, das andere Mal nicht? Shepard (1990) gibt folgende Erläuterung: „Der Rahmen und die Textur einer Seite, auf der eine Fotografie erster Ordnung

gedruckt ist, dominiert und bestimmt deshalb die Korrektur der Neigung. Eine Fotografie zweiter Ordnung wird verzerrt belassen und erscheint flach, egal mit welchem Blickwinkel man sie betrachtet” (Shepard, 1990, S. 195).

Wie hat das visuelle System die erstaunliche Fähigkeit erworben, den Abbildungswinkel von Figuren zu korrigieren? Diese Leistung der Bildwahrnehmung entspricht dem Wiedererkennen von Objekten bei der Umgebungswahrnehmung, bei der sich die Perspektive ständig ändert. Zur Erklärung kann die Invarianzermittlung herangezogen werden, deren Basis die Wahrscheinlichkeiten der physikalischen Welt sind. Sie führt zu dieser Korrektur. Das Gehirn ist offensichtlich nur zu einer Berichtigung und nicht zu einer doppelten Anpassung in der Lage. Wird das erste Bild noch einmal aus einer anderen Perspektive auf einem Bild gezeigt, versagt die Formkonstanz, da die Lage und die Grenze dieser zweiten Bildebene nicht zu identifizieren sind. Im Normalfall führt sie dazu, dass eine Betrachterin oder ein Betrachter von fast jedem Standort aus glaubt, das Bild so zu sehen, wie der Maler oder Fotograf es gesehen hat. Eine alternative Erklärung wäre, dass gar keine Korrektur vorgenommen werden muss, weil auch die Repräsentation von Objekten nicht einer dreidimensionalen Darstellung entspricht, sondern mit den Merkmalen gelingt, die bereits in zwei Dimensionen vorhanden sind (vgl. Kapitel VII: Analyse und Konstruktion).

Für die Betrachterin oder den Betrachter werden die Konstanzleistungen des visuellen Systems, die in der Umgebungswahrnehmung lebensnotwendig sind, bei der Bildwahrnehmung zu attraktiven Fallen, denen er sich kaum entziehen kann. Die Möglichkeit ergibt sich bei der Unterstellung, dass ein Bild ein Abbild sein muss. Die Konstanzleistungen führen automatisch von instabilen und mehrdeutigen Zuständen zu invarianten Objekten. Sie sind ein effektives Mittel der Umgebungswahrnehmung, welches auf die Bildwahrnehmung angewandt leicht in eine Sackgasse führt. Der Grund liegt in einer falschen Prämisse,

die sich auf das Verhältnis von Bild und Wirklichkeit bezieht. Entgegen der Auffassung, dass wir heute, weil wir in einer „Bilderwelt“ leben, die Welt auch als Bild sehen (vgl. etwa Lopes, 1996), wird hier die Ansicht vertreten, dass wir, durch eine lange Evolutionsgeschichte geprägt, in der artifizielle Bilder bis vor Kurzem keine oder eine nur untergeordnete Rolle gespielt haben, auch im 21. Jahrhundert ein Bild als Abbild der Welt sehen und in Bildern nach Korrespondenzen zur Wirklichkeit suchen.

Kapitel VII: Analyse und Konstruktion

Die Aufgabe von Theorien und Modellen ist es, die Arbeitsweise des visuellen Systems zu beschreiben und seine erstaunlichen Leistungen zu erklären. Modelle der visuellen Informationsverarbeitung unterscheiden zwischen *Repräsentationen* und *Prozessen*. Repräsentationen entsprechen den Zuständen des visuellen Systems; es sind verschiedene Arten, eine Information darzustellen. Prozesse beschreiben die Übergänge zwischen den Repräsentationen, die Art und Weise, wie aus einer Repräsentation eine andere entsteht. Steven Palmer (1999, S. 80) formulierte: „Without representations, processes would have nothing to work on; without processes, no work would get done.“ Das visuelle System lässt sich als ein *repräsentationales System* auffassen, in dem zwei Welten aufeinander bezogen werden: die Außenwelt (Represented World) und die Innenwelt (Representing World). Zwischen beiden Welten muss es strukturelle Ähnlichkeiten geben, die sich formal als *Homomorphismus* beschreiben lassen (vgl. Abbildung 74).

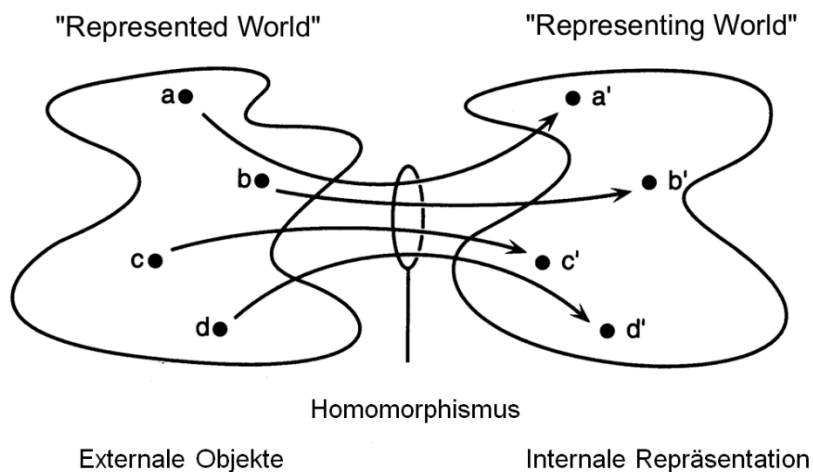


Abbildung 74: Homomorphismus (Palmer, 1999). Erläuterung im Text.

Die interne Abbildung der externalen Objekte muss in einer Art und Weise geschehen, dass die Relationen zwischen den externalen Objekten

auch in der internalen Abbildung erhalten bleiben. Ein Beispiel: Drei Objekte haben dieselbe Farbe, aber unterschiedliche Größen. Die Repräsentation muss diese Relationen enthalten. Das bedeutet nicht, dass Bilder im Kopf gespeichert werden müssen. Eine mentale Repräsentation muss und kann keine Eins-zu-eins-Abbildung der Wirklichkeit sein. Es geht darum, die relevanten Merkmale einer Szene so zu speichern, dass auch unter sich verändernden Bedingungen mit der Außenwelt sinnvoll interagiert werden kann.

Bei den Prozessen wird die Richtung unterschieden. Die von der Retina zum Cortex verlaufenden Prozesse werden, wie in Kapitel II erläutert, als „bottom-up“ oder „datengesteuert“ bezeichnet, die abwärts gerichteten Prozesse als „top-down“ oder „hypothesengesteuert“. Die Bottom-up-Verarbeitung fasst alle Prozesse zusammen, die eine Repräsentation auf einer niedrigeren Ebene in eine Repräsentation auf einer höheren Ebene überführt. Die Top-down-Verarbeitung beschreibt die umgekehrte Richtung. Die Modelle der visuellen Informationsverarbeitung unterscheiden sich a) in der Beschreibungsebene, b) im Auflösungslevel, c) in der Anzahl der Stufen der Informationsverarbeitung, d) im Repräsentationsformat und e) in der relativen Gewichtung von daten- und hypothesengesteuerten Prozessen. Bedauerlicherweise verwendet jede Theorie bzw. jedes Modell eine eigene, häufig sehr künstliche Begrifflichkeit: Was, zum Beispiel, könnte eine „Erstskizze“ sein oder ein „FIDO“? Das hat den Nachteil, dass die Leserin oder der Leser sich in diese Begrifflichkeit einarbeiten muss und dass bei einem Vergleich der Theorien nicht leicht zu entscheiden ist, ob hinter unterschiedlichen Bezeichnungen auch unterschiedliche Konzepte stehen. Es hat aber den Vorteil, dass der Konstruktcharakter der Begrifflichkeit betont wird. Die theoretischen Begriffe werden in der Folge nicht leicht ontologisiert.

1. Autonome Fahrzeuge

Bevor die anspruchsvollen Modelle der visuellen Informationsverarbeitung ins Zentrum des Interesses rücken, soll eine Einstimmung an einem Beispiel erfolgen. Im Mittelpunkt steht das Zusammenwirken der Analyse- und Konstruktionsprozesse. Was bis vor Kurzem noch in den Bereich der Science-Fiction gehörte, befindet sich heute bereits im Stadium der praktischen Erprobung: Fahrzeuge, die ohne Eingriffe des Menschen die Umgebung analysieren, Objekte und Hindernisse ausfindig machen, selbstständig und unter Berücksichtigung der Verkehrsregeln Entscheidungen treffen und sich autonom durch den Straßenverkehr bewegen. Als ein erster Höhepunkt dieser Entwicklung kann im Rückblick die amerikanische Urban Challenge 2007 gesehen werden, in der sich autonome Fahrzeuge erstmals in der Automobilgeschichte in einer urbanen Umgebung gemessen haben. Der von der Technologieabteilung des amerikanischen Verteidigungsministeriums (DARPA) ausgeschriebene Wettbewerb war mit Prämien von insgesamt 3,5 Millionen US-Dollar für diejenigen drei Teams dotiert, deren Fahrzeuge ohne menschliche Interventionen eine Strecke von knapp 90 Kilometern möglichst schnell und mit möglichst wenigen Verstößen gegen die kalifornischen Verkehrsregeln zurücklegten. Elf Teams qualifizierten sich für das Finale, sechs Fahrzeuge erreichten das Ziel. Das Siegerfahrzeug – ausgestattet mit einer zweistelligen Anzahl Kameras, Lasern, Radargeräten und sehr leistungsstarken Computern – beendete den Parcours in vier Stunden und zehn Minuten mit einer Durchschnittsgeschwindigkeit von 23 km/h. Die Urban Challenge zeigte, wie differenziert intelligente Fahrzeuge ihre Umwelt analysieren und wie effektiv sie darin navigieren. Die Entwicklung ist rasant: Im Frühjahr 2012 haben im Bundesstaat Nevada acht Prototypen eines autonomen Fahrzeuges aus den Laboratorien des Technologiekonzerns Google eine Zulassung erhalten. Im Herbst des gleichen Jahres erfolgte eine Zulassung von Testfahrzeugen für den Bundestaat Kalifornien. Die Fahrzeuge dürfen

sich auf öffentlichen Straßen bewegen. Zur Sicherheit muss derzeit jedoch eine Fahrerin oder ein Fahrer an Bord sein, um bei einem Ausfall der Systeme durch Lenk- und Bremsvorgänge eingreifen zu können. Viele Autohersteller arbeiten gegenwärtig an der Entwicklung autonomer Automobile bis hin zur Serienreife.

Doch warum sollten Entwicklungen im Bereich der Ingenieurwissenschaften eine Rolle in einem Buch über die visuelle Wahrnehmung des Menschen spielen? Die Forschung hat in beiden Fällen ein vergleichbares Anliegen. In einem Fall soll ein technisches System entwickelt werden, das ähnliche Leistungen wie ein bekanntes biologisches System vollbringt. Im anderen Fall soll die Arbeitsweise eines biologischen Systems mit Mitteln der Technik und der Informatik analysiert werden, um seine grundlegenden Probleme besser zu verstehen. Die Zugangsweisen könnten voneinander profitieren. Die Technik nutzt neurowissenschaftliche Erkenntnisse, um „sehende Maschinen“ zu konstruieren, die Wahrnehmungsforschung ihrerseits profitiert von den entwickelten Modellannahmen und Algorithmen. Wenn sich Modelle im technischen Bereich als sinnvoll und funktionstüchtig erweisen, dann lässt das womöglich Rückschlüsse auf das visuelle System des Menschen zu.

Ein Teil der Aufgaben weist eine hohe Ähnlichkeit auf: Die eigene Position muss mit jeder ihrer Veränderungen bestimmt werden. Objekte müssen klassifiziert und identifiziert werden. Sie müssen lokalisiert und in Bezug auf ihre Bewegung muss sowohl die Richtung als auch die Geschwindigkeit ermittelt werden. All dies muss in Echtzeit geschehen. Darüber hinaus müssen zukünftige Ereignisse antizipiert werden, Verhaltensoptionen bewertet und ausgewählt werden. Kurz: Es muss ein dreidimensionales Modell der Welt erstellt werden, das auch unter schwierigen Witterungs- und Beleuchtungsverhältnissen zuverlässige Entscheidungen erlaubt. Eine besondere Anforderung stellt die hohe Geschwindigkeit der Veränderungen und die Komplexität der Fahrumgebung dar. Wie bei anderen Modellen hängt die Repräsentation von der Zielsetzung ab. Im

Straßenverkehr hat Sicherheit die oberste Priorität, aber auch der Verkehrsfluss muss gewahrt bleiben.

Autonome Fahrzeuge gewinnen nur einen Teil ihrer Informationen über Kameras und die computergestützte Auswertung ihrer Bilder. Andere wichtige Hilfsmittel sind z. B. Radar, Laser und Infrarot. Wie beim Zusammenspiel zwischen verschiedenen Sinnesmodalitäten kommt es letztlich darauf an, auf der Basis dieser unterschiedlichen Informationen eine möglichst widerspruchsfreie und realistische Repräsentation der Umgebung zu erstellen. Hierzu müssen Informationen selektiert, ergänzt und gewichtet werden. Wie die biologischen Systeme nutzen auch die technischen Systeme hierzu die Probabilitäten der Außenwelt und versuchen über die Zirkularität die Ambiguitäten zu reduzieren.

Technische Systeme zur visuellen Objekterkennung greifen oft auf ein zweistufiges Verfahren zurück. Am Anfang steht die *Hypothesengenerierung*. Mögliche Objekte und deren Konturen sollen in einem Bild lokalisiert und separiert werden. Im Fall autonomer Automobile sind diese Objekte vor allem andere Fahrzeuge, Fußgänger*innen, Radfahrer*innen, Hindernisse sowie Spur- und Fahrbahnbegrenzungen. Auf die Hypothesengenerierung folgt die *Hypothesenverifikation*. In ihr wird überprüft, ob die lokalisierten Objekte tatsächlich zur vermuteten Objektgruppe gehören. Die Prozesse der Hypothesenverifikation haben ein detaillierteres Auflösungsvermögen, wären aber zu langsam, wenn die Prozesse der Hypothesengenerierung die Informationen nicht schon vorstrukturieren würden. Für jede Verarbeitungsstufe existieren unterschiedliche Lösungsansätze, die spezifische Stärken und Schwächen aufweisen (für einen Überblick vgl. Sun, Bebis und Miller, 2006). Das Ziel dieser Methoden ist eine Minimierung der falsch-negativen Identifizierungen (Objekt wird nicht als solches erkannt, obwohl es im Blickfeld vorhanden ist) und der falsch-positiven Identifizierungen (es wird ein Objekt identifiziert, das sich nicht im Sichtbereich befindet). Die Vorgehensweise soll am Beispiel der

Fahrzeugeterkennung, dem Auffinden anderer Fahrzeuge im Sichtbereich der Kamera, veranschaulicht werden. Hierfür wurden unterschiedliche Verfahren der Hypothesengenerierung und -verifikation vorgeschlagen und erprobt. Generell gilt, dass sich die Verfahrensarten gegenseitig nicht ausschließen, sondern vielmehr ergänzen und die Schwachstellen der anderen wechselseitig ausgleichen können.

Methoden der Hypothesengenerierung

Hypothesengenerierende Verfahren suchen nach allgemeinen Informationen. Sie müssen nicht unbedingt nur auf Fahrzeuge zutreffen. Falsch identifizierte Objekte müssen bei der Hypothesenprüfung verworfen werden. Das Hauptziel dieser Methoden liegt in einem schnellen und (im Hinblick auf die benötigte Rechenleistung) ökonomischen Auffinden möglicher Fahrzeuge, Hindernisse etc. Die Verfahren lassen sich in drei Gruppen gliedern: wissensbasierte, stereobasierte und bewegungsbasierte Verfahren.

a) wissensbasierte Verfahren

Wissensbasierte Verfahren arbeiten auf Basis von Informationen, die a priori in das System eingespeist und in konkreten Situationen zur Identifizierung von Fahrzeugen abgerufen werden. Es handelt sich vor allem um Wahrscheinlichkeiten in Bezug auf Symmetrien, Kanten und Ecken, Textur, Farbe, Schatten und Beleuchtung. Jedes dieser Verfahren hat bezüglich einer sicheren Objektidentifikation bestimmte Fehlerwahrscheinlichkeiten, die von den aktuellen Verkehrs- und Sichtbedingungen abhängen. Einzeln sind die verwendeten Heuristiken nicht zuverlässig. Erst die Kombination unterschiedlicher Regeln garantiert den Erfolg. Die Tatsache, dass nahezu alle Autos, wenn sie in einem 90° Winkel von vorne oder hinten betrachtet werden, spiegelsymmetrisch in Bezug auf ihre vertikale Achse erscheinen, kann eine wichtige Heuristik für ein wissensbasiertes System darstellen. Das aus

der Digitalkamera gewonnene Bild wird auf diese Art von Symmetrie hin ausgewertet. Die Methode ist am zuverlässigsten bei einfachen, reizarmen Umgebungen, wie bei einer Fahrt über eine Landstraße. In urbanen Situationen, mit vielen anderen symmetrischen Objekten im Blickfeld, steigt das Risiko falsch-positiver Identifizierungen. Problematisch ist in der Praxis zudem, dass die Methode ihrerseits die zuverlässige Erkennung homogener Flächen voraussetzt, doch diese treten aufgrund von Störeinflüssen oft nicht klar genug im Kamerabild hervor.

Bei den anderen Verfahren sind die Probleme ähnlich. Immer gibt es Bedingungen, unter denen die verwendeten Heuristiken fehleranfällig sind. Ein zweites Beispiel: Da die meisten Fahrzeuge in der frontalen wie rückwärtigen Ansicht als charakteristische Form mit vier deutlich sich abzeichnenden Ecken sowie jeweils zwei horizontalen und vertikalen Kanten erscheinen, können diese Merkmale zur Identifikation herangezogen werden. Hierfür müssen die Kanten von Objekten mithilfe verschiedener Bildbearbeitungsalgorithmen verstärkt und extrahiert werden. Dann kann das Bild nach entsprechenden rechteckigen Formen abgesucht werden. Dieses Verfahren hat sich als sehr erfolgreich erwiesen, ist jedoch ebenfalls mit spezifischen Problemen verbunden. Vor der Extraktion der Kanten muss ein Schwellenwert festgelegt werden, ab dem eine lokale Helligkeitsdifferenz für bedeutsam erachtet und damit als Kante (und nicht etwa als ein Schatten) identifiziert wird. Diese Schwelle muss so gewählt werden, dass alle wichtigen Objekte erkannt werden, andererseits aber das Risiko falsch-positiver Identifizierungen nicht erhöht wird. Die „ideale Schwelle“ ist stark von Beleuchtungs- und Umgebungseinflüssen abhängig und muss ständig adjustiert werden.

Der Asphalt der Fahrbahn weist ein spezifisches Farbspektrum auf. Wird aus dem Kamerabild eine Farbverteilung berechnet, in der sich ein besonders hoher Anteil eines bestimmten Grautons findet, kann davon ausgegangen werden, dass es sich bei der Fläche um die Fahrbahn handelt. Theoretisch ist die Heuristik „Interpretiere Flächen einer bestimmten

Ausdehnung und eines bestimmten Grautons als Fahrbahn“ eine effiziente Möglichkeit der Oberflächenidentifikation. Praktisch wurde der Farbton bisher aber kaum zur Hypothesengenerierung verwendet, da die Feststellung mit einer Reihe von Schwierigkeiten verbunden ist. Die Farbaufzeichnung durch Kameras hängt zum einen von der Farbtemperatur des Umgebungslichtes ab. Dessen spektrale Zusammensetzung unterscheidet sich bei morgendlicher und abendlicher Beleuchtung deutlich. Zum anderen ist die Farbe von der Kameraeinstellung am Fahrzeug beeinflusst (Weißabgleich). Die oben formulierte Heuristik bedarf in der Praxis einer Kombination mit anderen Regeln. Es fehlt sonst die Helligkeits- bzw. Farbkonstanz (vgl. Kapitel V).

Auch Schatten geben deutliche Hinweise auf die im Bild vorhandenen Objekte. Untersucht man die Helligkeitsverteilungen auf während der Fahrt aufgezeichneten Kamerabildern, zeigt sich, dass der Bereich unter den Fahrzeugen meist die dunkelste Fläche im Bild ist. Eine Heuristik „Interpretiere rechteckige Flächen oberhalb der dunkelsten Fläche im Bild als Fahrzeug“ wäre leistungsstark. Doch ihre Schwächen sind ebenso offensichtlich. Die Bildung der Schatten hängt von der Umgebungsbeleuchtung ab: Zeichnen sich die Schatten der Fahrzeuge aufgrund diffuser oder schwacher Umgebungsbeleuchtung (z. B. bei Regen oder Schnee) nicht deutlich genug ab, versagt auch die Objekterkennung.

Die meisten der bisher beschriebenen Verfahren zur Hypothesengenerierung sind ineffizient in Situationen mit unzureichender Umgebungsbeleuchtung. Eine extreme Testsituation sind Nachtfahrten. Hier könnte die Fahrzeugbeleuchtung hilfreich sein, um andere Verkehrsteilnehmer zu identifizieren. Eine ähnliche Form, eine ähnliche Helligkeit sowie ein bestimmter Abstand zwischen zwei Lichtquellen könnte ein Hinweis darauf sein, dass es sich bei den betreffenden Lichtquellen um die Beleuchtung eines Autos handelt. Die Schwierigkeit besteht darin, zwischen Fahrzeuglichtern und

Hintergrundlichtern bzw. Straßenbeleuchtungen zu differenzieren. Gut, dass autonome Fahrzeuge nicht nur über optische Sensoren verfügen.

b) stereobasierte Verfahren

Bei stereobasierten Verfahren der Hypothesengenerierung werden zwei versetzt montierte Digitalkameras zur Erfassung des Sichtbereiches verwendet. Dementsprechend dienen zwei Kamerabilder als Grundlage der Fahrzeugerkennung. Die Parallelität zum binokularen Sehen des Menschen ist offensichtlich. Es stellt sich dasselbe *Korrespondenzproblem*: Welche Pixel beider Bilder entsprechen dem gleichen Punkt im Sichtfeld? Das Problem ist bei sich bewegenden Objekten keineswegs trivial: Durch Erschütterungen während der Fahrt, durch die Eigenbewegung und durch weitere Störfaktoren verändern sich die exakte Position der Kameras und das aufzunehmende Objekt ständig. Eine – allerdings rechnerisch relativ aufwändige – Möglichkeit, das Korrespondenzproblem trotz dieser Widrigkeiten zu lösen, besteht darin, die jeweiligen Bildpixel anhand der sie umgebenden Pixel in verschiedene Kategorien zu unterteilen (z. B. horizontale oder vertikale Kanten zu bilden). Hierfür müssen die bekannten Heuristiken eingesetzt werden. Danach lassen sich leichter Entsprechungen zum zweiten Kamerabild finden.

Wurde für alle Bildbereiche eine Korrespondenz zwischen beiden Aufnahmen hergestellt, wird eine *Disparitätskarte* berechnet, die ihrerseits in eine dreidimensionale Karte der Szene konvertiert werden kann. Peaks in einer *Disparitätsverteilung* können helfen, Objekte zu identifizieren, da sie eine größere Fläche mit einer einheitlichen Disparität darstellen. Es ist wichtig, dass die Disparitätskarte eine ausreichend hohe Auflösung besitzt, um alle in der Szene vorhandenen Objekte zu erfassen. Die Erstellung immer neuer Disparitätskarten sowie deren Auswertung in Echtzeit bei einem sich bewegenden Fahrzeug erfordern eine hohe Rechenleistung.

c) bewegungsbasierte Verfahren

Bewegungsbasierte Verfahren gehen von der Berechnung des optischen Fließmusters aus, mit der die Bewegung der Objekte relativ zur Bewegung des Sensors bestimmt werden kann. Das Bild wird in Teilbereiche oder Regionen zerlegt, für die die durchschnittliche Bewegungsgeschwindigkeit und die Bewegungsrichtung bestimmt werden. Unterschiedliche Objektbewegungen bedingen unterschiedliche Fließmuster (vgl. auch Abbildung 12): Während der weit entfernte Hintergrund und Fahrzeuge, die sich in gleicher Richtung mit gleicher Geschwindigkeit bewegen, praktisch keinen Flow verursachen, rufen entgegenkommende Fahrzeuge divergierende Fließmuster hervor. Überholende oder sich entfernende Fahrzeuge sind durch ein konvergierendes Muster zu lokalisieren. Unter Einbeziehung der eigenen Fahrzeuggeschwindigkeit kann die Geschwindigkeit anderer Objekte geschätzt werden. In Teilbereichen des Bildes, für die eine wesentlich geringere Geschwindigkeit als für das eigene Fahrzeug ermittelt wird, kann ein Hindernis vermutet werden. Entscheidend für eine gute Objekterkennung ist die Berechnung eines optischen Fließmusters in ausreichend hoher Auflösung. Hier wird ein Optimum angestrebt, denn jeden Pixel separat zu berechnen, das würde nicht nur sehr große Leistungsressourcen erfordern, sondern wäre auch sehr anfällig für Störeinflüsse. Daher werden Informationen „höherer Ordnung“ über Kanten, Ecken oder Texturen einbezogen. Rechenleistung kann eingespart werden und die Bildauswertung wird weniger zufallsabhängig. Trotzdem kommt es bei sehr schnellen Bewegungen des Sensors oder bei größeren Flächen ohne bedeutsame Texturdifferenzen zu Problemen.

Methoden zur Hypothesenverifikation

Nachdem ein bestimmter Bildbereich als bedeutsames Objekt identifiziert wurde, wird mit Methoden der Hypothesenverifikation geprüft, ob es sich um das vermutete Objekt handelt. Die verwendeten Heuristiken können

wesentlich differenziertere, auf spezifische Objekte ausgerichtete Informationen prüfen, sind jedoch zur Analyse des gesamten Bildbereiches nicht geeignet. Man unterscheidet zwischen *vorlagenbasierten* und *erscheinungsbasierten* Verfahren. Erstere nutzen kritische Merkmale von Fahrzeugen, um gefundene Objekte dieser Klasse zuzuordnen. In ihrer Grundstruktur ähneln sie wissensbasierten Verfahren der Hypothesengenerierung, doch verwenden sie spezifischere Informationen. So kann z. B. das Vorhandensein von Nummernschildern sowie einer Front- bzw. Heckscheibe überprüft werden. Da diese Heuristiken je nach Entfernung des betreffenden Objektes unterschiedlich effektiv sind, müssen sie miteinander kombiniert werden. Der erscheinungsbasierte Ansatz geht einen anderen Weg: Der Klasse der Fahrzeuge wird eine zweite Klasse, die Nicht-Fahrzeuge, gegenübergestellt. Das System muss entscheiden, welcher Klasse das betreffende Objekt zuzuordnen ist. Dieses Vorgehen führt zu besseren Ergebnissen, erfordert allerdings auch einen höheren Rechenaufwand. Die Erscheinungsformen sowohl innerhalb der Klasse der Fahrzeuge, als auch innerhalb der Klasse der Nicht-Fahrzeuge weisen eine große Variabilität auf. Um eine gute Differenzierung zwischen beiden Klassen zu erzielen, müssen Eigenschaften verschiedener Fahrzeuge als auch verschiedener Nicht-Fahrzeug-Objekte in Form von „Trainingsbildern“ in das System eingegeben werden. Danach können Korrelationen der aktuellen Bilder mit den jeweiligen Klassen berechnet werden. Im abschließenden Schritt wird aufgrund dieser Korrelation das Objekt einer Klasse zugeordnet.

Fazit: Im Hinblick auf das hohe Risiko, das von Fehlentscheidungen im Straßenverkehr ausgeht, kommt es darauf an, die verschiedenen Systeme mit ihren je eigenen Stärken und Schwächen zu kombinieren. Wie bei der visuellen Wahrnehmung des Menschen muss eine möglichst große Anzahl verschiedener Heuristiken integriert werden, um die Fehlerrate zu minimieren.

Fahrzeugsteuerung

Bisher wurde beschrieben, anhand welcher Methoden visuelle Informationen in autonomen Fahrzeugen erfasst und ausgewertet werden können. Wie werden diese Informationen eingesetzt, um das Fahrzeug sicher im Straßenverkehr zu bewegen? Eine Möglichkeit, aus den Sensordaten konkrete Handlungsentscheidungen abzuleiten, besteht in der Verwendung von *Markov-Logik-Netzen*. Die Informationen aller aktiven und passiven Sensoren sowie weitere Informationen (z. B. GPS-Daten) werden in Form von Netzwerken miteinander in Verbindung gebracht. Diese Netze basieren ebenfalls auf einem probabilistischen Grundgedanken: Nachdem jedem erfassten Objekt mit einer gewissen Wahrscheinlichkeit eine bestimmte Klasse (z. B. „Die Wahrscheinlichkeit, dass Objekt X zur Klasse der Fahrzeuge gehört, beträgt 89 %.“), eine bestimmte Position relativ zum eigenen Fahrzeug („Die Wahrscheinlichkeit, dass sich Objekt X auf der linken Spur der Fahrbahn befindet, beträgt 96 %.“), eine spezifische Bewegungsrichtung und Geschwindigkeit („Die wahrscheinlichste Schätzung lässt vermuten, dass Objekt X sich mit 60 km/h auf das eigene Fahrzeug zubewegt.“) sowie weitere Parameter zugeordnet wurden, kann das System eine Vermutung darüber anstellen, ob ein Objekt bei der bisher vorgesehenen Fahrlinie ein Hindernis oder eine Gefährdung darstellen könnte – oder ob es vielleicht sogar zu einer Kollision kommen könnte. Sollte das der Fall sein, kann eine Neuberechnung der *Trajektorien* erfolgen: Es wird eine sichere Fahrlinie berechnet, deren Verfolgung in der aktuellen Situation fahrdynamisch und lenktechnisch möglich ist. Das Steuerungssystem versucht, dieser Linie unter Einbeziehung von Sensordaten, z. B. identifizierter Fahrbahnbegrenzungen, zu folgen. Währenddessen werden die Daten ständig aktualisiert und die ausgewählte Trajektorie wird adjustiert (vgl. auch Stiller, Kammel, Lucheva und Ziegler, 2008). Die enorme Komplexität und die erforderliche Rechenleistung nicht nur in der Informationsverarbeitung, sondern auch in der Umsetzung der Fahrzeugsteuerung verdeutlichen eindrucksvoll, welche Leistungen unser

Gehirn permanent vollbringt, ohne dass der Mensch sich dessen bewusst ist. Er nimmt es als geradezu selbstverständlich hin. Dass die Selbstverständlichkeit eben nicht Einfachheit bedeutet, darauf verweisen die Schwierigkeiten, die sich bei der Entwicklung von verkehrssicheren „sehenden Maschinen“ ergeben (für einen Überblick zu den Techniken der Bewegungsplanung und Kontrolle im Zusammenhang mit autonomen Fahrzeugen vgl. Paden, Cap, Yong, Yershov und Frazzoli, 2016).

2. Maschinelle und biologische Wahrnehmung

Bei einem Vergleich der Verarbeitung optischer Informationen durch autonome Automobile und der visuellen Informationsverarbeitung des Menschen zeigen sich Gemeinsamkeiten und Unterschiede. Ein Teil der Aufgaben ist identisch. Viele der beim menschlichen Sehen festgestellten Heuristiken haben sich bei der Entwicklung autonomer Automobile bewährt. Offensichtlich kommen in einigen Fällen dieselben Lösungen zum Tragen. Die Identifikation von Ecken und Kanten und die Analyse optischer Fließmuster mit ihren Invarianten spielen hier wie dort eine besondere Rolle. Texturunterschiede helfen bei der Abgrenzung von Oberflächen, Symmetrie und geometrische Grundbausteine bei der Objekterkennung. Die Disparität schließlich erleichtert eine dreidimensionale Repräsentation.

Wie biologische Systeme nutzen auch maschinelle Systeme die Probabilitäten der Außenwelt, um die Ambiguität der eintreffenden Signale zu reduzieren. Vor allem um die Rechenleistung zu verringern, aber auch um unabhängiger von Zufälligkeiten und Störungen zu werden, greifen sie auf die Prinzipien der Kontrastbildung und der Gruppierung zurück. In vielen Einzelleistungen sind sie bereits effektiver als die menschliche Wahrnehmung (z. B. ist die Messgenauigkeit höher). Weniger effektiv sind sie bislang in Bezug auf die Stabilität und Multiperspektivität. Das hängt im Wesentlichen damit zusammen, dass die optischen Informationen in autonomen Fahrzeugen überwiegend

seriell, also in aufeinanderfolgenden Auswertungsschritten verarbeitet werden. Es ist eine Bottom-up-Strategie. Zwar wird Vorwissen über die Umwelt in Form von Wahrscheinlichkeiten einbezogen, aber dieses Wissen ist bisher eher statisch und wird dem System vorgegeben.

In diesem Zusammenhang ist die Rolle der Aufmerksamkeit bedeutsam. Einen vergleichbaren Prozess gibt es in maschinellen Systemen zurzeit nur ansatzweise. Damit fehlt im Sinne einer Top-down-Strategie die Möglichkeit, gezielt Such- und Orientierungsreaktionen einzuleiten. Das visuelle System des Menschen ist wesentlich effektiver, wenn Mehrdeutigkeiten sich nicht unmittelbar auflösen lassen und Interpretationen miteinander konkurrieren. Durch gezielte Augenbewegungen und längere Fixationen lässt sich der Prozess der Informationsgewinnung optimieren. Den maschinellen Systemen fehlt bisher ebenfalls eine echte *Situationserkennung*. Diese setzt eine Objekterkennung und einen Vergleich mit anderen Verkehrssituationen voraus. Wahrscheinlichkeiten für bestimmte Objekte und Ereignisse ließen sich ableiten, die helfen, verbleibende Mehrdeutigkeiten zu reduzieren.

Zwischen den beiden Systemen ist der Effektivitätsunterschied bei der Ambiguitätsreduktion der augenscheinlichste. Ein weiterer ist in der Anpassungsfähigkeit bzw. in der Universalität des Modells auszumachen: Computerprogramme, die autonome Fahrzeuge steuern, sind für diese Aufgabe entwickelt worden. Das Siegerfahrzeug der Grand Challenge von 2005, die in unbewohntem Gelände stattfand, hätte ohne grundlegende Modifikationen nicht an der Urban Challenge von 2007 teilnehmen können. So komplex die autonomen Fahrzeuge auch scheinen, bei Änderungen bzw. Erweiterungen der Ausgangssituation sind sie bisher kaum in der Lage, sinnvoll zu reagieren. Ihre Programmierung ist (notwendigerweise) spezifisch.

Automobile fahren ist nicht die einzige Tätigkeit des Menschen. Andere Aufgaben der Umgebungswahrnehmung, wie etwa „Einkaufen“, „Handwerken“ oder „Kochen“, stellen andere Anforderungen an das visuelle System. Für all diese unterschiedlichen Aufgaben muss die Repräsentation der Welt adäquat sein. Sie muss universell sein, kann aber nicht in allen Aspekten gleich präzise sein. Steht nicht der Sicherheitsaspekt im Vordergrund, kann sie in der jeweiligen Situation durch Lernprozesse angepasst oder erweitert werden.

3. Modelle der visuellen Informationsverarbeitung

Die im Folgenden dargestellten drei Modelle haben je eigene Ziele, setzen deshalb verschiedene Schwerpunkte und wählen unterschiedliche Auflösungs-niveaus. Ihre Autoren sind David Marr, Steven Grossberg und Steven Palmer. Die Modelle sind geeignet, den Forschungsstand zu reflektieren und in die spezielle Art der Theoriebildung einzuführen. Sie liegen auf unterschiedlichen Beschreibungsebenen und konkurrieren nicht miteinander. Untereinander weisen sie viele Beziehungen auf.

Die Überlegungen von David Marr (1982) sind, historisch gesehen, vielleicht die einflussreichsten. Ausgehend von Graustufenbildern simuliert Marr den Prozess der visuellen Informationsverarbeitung als einen rein datengesteuerten Prozess. Als *computationales Modell* deckt es die grundlegenden Probleme auf, die bei der Konstruktion der visuellen Welt zu lösen sind. Die menschliche und die technische Informationsverarbeitung vergleichend lassen sich drei für beide Bereiche bedeutende Fragenkreise abgrenzen: 1. Welche Ziele und Aufgaben muss das System erfüllen? 2. Mit welcher Folge von Einzelschritten lassen sich die Aufgaben lösen? 3. Wie können Neurone oder elektronische Schaltkreise diese Abfolge ausführen? Die zweite Frage ist die nach dem angemessenen Algorithmus oder der richtigen Heuristik. Marr hat gerade auf dieser Ebene vielbeachtete Beiträge geliefert. Ein Algorithmus, der in einem Computerprogramm zu zufriedenstellenden Resultaten führt,

erlaubt nach Marr Rückschlüsse auf ähnliche Algorithmen oder Heuristiken bei der menschlichen Informationsverarbeitung. Prozesse am Rechner zu simulieren impliziert das vollständige Hinterfragen der Selbstverständlichkeit des Sehens. Jedes Problem muss erkannt und versuchsweise gelöst werden. Eine gefundene Lösung bestätigt zunächst, dass es eine Lösung gibt, nicht aber, ob sie beim menschlichen Sehen Verwendung findet.

David Marr verstarb im Alter von 35 Jahren. Sein Gesamtwerk blieb fragmentarisch. Er konnte die Komplexität scheinbar einfacher Aufgaben des Sehens aufzeigen und entwickelte vor allem für die ersten Stufen der Informationsverarbeitung eine Reihe von Algorithmen und Heuristiken, die eine gute Annäherung an diese Komplexität ermöglicht. Wegen der Tiefe seiner Analyse und der bemerkenswerten Fähigkeit seines Modells zur Integration physikalischen, physiologischen und psychologischen Wissens hat es die Forschung wie kein anderes beeinflusst.

Die FACADE Theorie von Steven Grossberg verbindet die computationale Sichtweise von David Marr mit neurophysiologischen und psychophysischen Erkenntnissen zur vorbewussten visuellen Wahrnehmung (vgl. z. B. Grossberg 1994, 1997) und steht damit stellvertretend für eine neuere Generation von Modellen. Sie ist eine *neuronale Netzwerktheorie*, die eine Reihe von Subsystemen postuliert und die in den letzten 25 Jahren Schritt für Schritt weiterentwickelt wurde. Grossberg geht es explizit um die Modellierung menschlichen Sehens auf der Basis der biologischen Grundlagen. Im Mittelpunkt steht das Problem der Ambiguitätsreduktion: Wie gelingt es dem visuellen System auf der Basis einer unvollständigen und vieldeutigen Stimulation ein eindeutiges Perzept zu generieren? FACADE steht für Form-And-Color-And-DEpth. Das Modell entwickelt Szenarien für das Zusammenspiel dieser drei Faktoren. Auch Grossberg wählt ein hohes Auflösungs-niveau. Seine Annahmen, wo und wie die visuellen Informationen verarbeitet werden, sind sehr spezifisch. Sie sollen klären, anhand welcher konkreten

Verschaltungen ein neuronales Netzwerk wie der visuelle Cortex die verschiedenen Verarbeitungsschritte effizient bewältigt und wie das genaue Zusammenspiel von Analyse- und Konstruktionsprozessen auf neuronaler Ebene aussieht.

Der Ansatz von Steven Palmer (1999) schließlich verfolgt ein anderes Ziel. Er stellt den Konsens dar, der in der Forschung seit den bahnbrechenden Arbeiten von David Marr erreicht wurde und schafft so einen begrifflichen Rahmen für den derzeitigen Forschungsstand. Ein solches *Rahmenmodell* ist selbst nicht empirisch prüfbar. Doch gerade wegen seines geringen Auflösungslevels ist es gut in der Lage, einen Überblick zu verschaffen. Es fasst das aktuelle Wissen in einem Bild zusammen.

Das Modell von Palmer (1999) ist geeignet, als Strukturierung eine gemeinsame Begrifflichkeit zu entwerfen. Es gliedert den Prozess der visuellen Informationsverarbeitung in vier Stufen. Jede Stufe umfasst eine ganze Reihe Einzelprozesse. Ausgangspunkt für die Analyse der Stimulation und die Konstruktion der visuellen Welt ist das retinale Bild. Zunächst werden lokale Inhomogenitäten identifiziert, durch Kontrastbildung verstärkt und schließlich gruppiert. Die wichtigsten Operationen sind das Erkennen von Kanten und Ecken auf der Basis von Helligkeitsunterschieden sowie die Identifizierung unterschiedlicher Regionen im Bild (*bildbasierte Verarbeitung*). Dieser Schritt ist grundlegend für alle weiteren Verarbeitungsstufen; auf dieser Ebene dürfte es viele algorithmische Lösungen geben. Heuristiken werden umso bedeutsamer, je weiter sich die Repräsentationen vom proximalen Reiz entfernen. In einem zweiten Schritt wird eine Repräsentation erstellt, die bereits die Oberflächen bzw. Texturen und ihre relative Neigung zueinander berücksichtigt (*oberflächenbasierte Verarbeitung*). Damit wird die Wiedergabe räumlicher Verhältnisse und Entfernungen möglich. Das Resultat ist eine stark perspektivenabhängige Repräsentation, die der retinalen Abbildung verpflichtet ist. Während bei der bildbasierten Verarbeitung Regionen im Bild identifiziert werden müssen, sind sie bei

der oberflächenbasierten Verarbeitung bereits der Ausgangspunkt der Analyse. Im dritten Schritt wird mithilfe von „Primitiven“ und „Elementarbausteinen“ ein dreidimensionales Objekt konstruiert. Dessen Repräsentation ist unabhängig vom aktuellen Standort (*objektbasierte Verarbeitung*). Ein Vergleich mit Vorstellungs- bzw. Gedächtnisbildern wird ermöglicht. Objekte können, auch wenn sie aus einer anderen Perspektive gezeigt werden, wiedererkannt werden. Im letzten Schritt erfolgt eine kategoriale Zuordnung. Sie ist notwendig, damit es zu einer Verbindung mit dem Wissen, aber auch den Wünschen und Zielen einer Person kommen kann. Hier stehen funktionale Aspekte im Vordergrund. Die Inhalte der Repräsentation werden „frei verfügbar“ (*kategorienbasierte Verarbeitung*).

Wichtig für das Verständnis des Modells ist, dass ein Prozess (z. B. die bildbasierte Verarbeitung) nicht abgeschlossen sein muss, bevor der nächste Prozess (z. B. die oberflächenbasierte Verarbeitung) beginnt und dass es neben den aufwärts gerichteten Prozessen immer auch Rückkoppelungsschleifen bzw. abwärts gerichtete Prozesse gibt. Auf diese Weise wird nicht nur die Sequenzialität der Informationsverarbeitung, sondern auch ihre Zirkularität berücksichtigt. Die Prozesse lassen sich als *induktives Schließen* beschreiben. Im Gegensatz zum deduktiven Schließen, bei dem, solange die Prämissen stimmen, die Schlussfolgerungen eindeutig sind, gibt es beim induktiven Schließen immer eine Unsicherheit. Das Modell geht von unvollständiger Information und von Wahrscheinlichkeiten aus.

Um zu verstehen, was auf den einzelnen Stufen der visuellen Informationsverarbeitung passiert, wird auf die Forschung von Marr zurückgegriffen und durch neuere Befunde ergänzt. Es zeigt sich: Das aktuelle Wissen über die frühen Stufen der Informationsverarbeitung ist deutlich detaillierter als über die späten Stufen. Nicht immer müssen alle

Diese machen für eine Simulation des Sehvorganges zumeist vereinfachende Annahmen: Die Retina wird als eine ebene Fläche mit der Fovea als Zentrum betrachtet; auf ihr wird eine gleichmäßige Verteilung der Rezeptoren unterstellt.

Nach heutigem Kenntnisstand bilden vor allem Helligkeitsunterschiede die Basis der Objektwahrnehmung. Im Prozess der Konstruktion von Formen auf der Grundlage dieser Helligkeitsunterschiede treten einige typische Probleme auf. Zum einen kann es vorkommen, dass zwei angrenzende oder überlappende Objekte sich nicht durch Helligkeitsunterschiede voneinander abgrenzen. Zum anderen liefern Helligkeitswerte oft mehrdeutige Informationen. Vier Faktoren können für die Helligkeitsunterschiede verantwortlich sein: 1. Geometrische Verhältnisse, 2. Lichtreflexionen der sichtbaren Oberflächen, 3. Beleuchtungsstärke der Szene und 4. Perspektive der Beobachterin oder des Beobachters. Diese vier Faktoren mischen sich im proximalen Reiz. Sie müssen im Laufe des Informationsverarbeitungsprozesses voneinander getrennt werden.

Helligkeitsunterschiede sind in der physikalischen Welt nicht willkürlich, sondern folgen festen Regeln. Sie können daher zur Identifizierung sichtbarer Oberflächen genutzt werden. Unter dem Aspekt der visuellen Informationsverarbeitung ist entscheidend, dass Oberflächen auf unterschiedlich feinem oder auch grobem Niveau strukturiert sind. Ein Beispiel dafür ist das Fell einer Katze. Auf einem sehr feinen Niveau besteht das Fell aus einzelnen Haaren, wobei jedes Haar auf eine charakteristische Weise Licht reflektiert. Die nächsthöhere Ebene ist die Musterung des Fells durch benachbarte parallele Haare. Die höchste Organisationsebene bildet die Fellzeichnung der Katze.

Abbildung 76 erläutert die Grundidee einer *Erstskizze* nach Marr. Auf der niedrigsten Ebene (Rohskizze) werden Bildelemente (einzelne Punkte oder Striche) hinsichtlich ihrer Helligkeitsveränderungen relativ zum

Hintergrund des Bildes analysiert. Diese Helligkeitsveränderungen werden in der Abbildung durch Grenzlinien dargestellt. Aus einem Punkt wird so ein Kreis, aus einem Strich eine elliptische Grenzlinie. Bei Strichen werden zudem noch Anfangs- und Endpunkte bestimmt (in Abbildung 76 sind diese durch fette Punkte markiert).

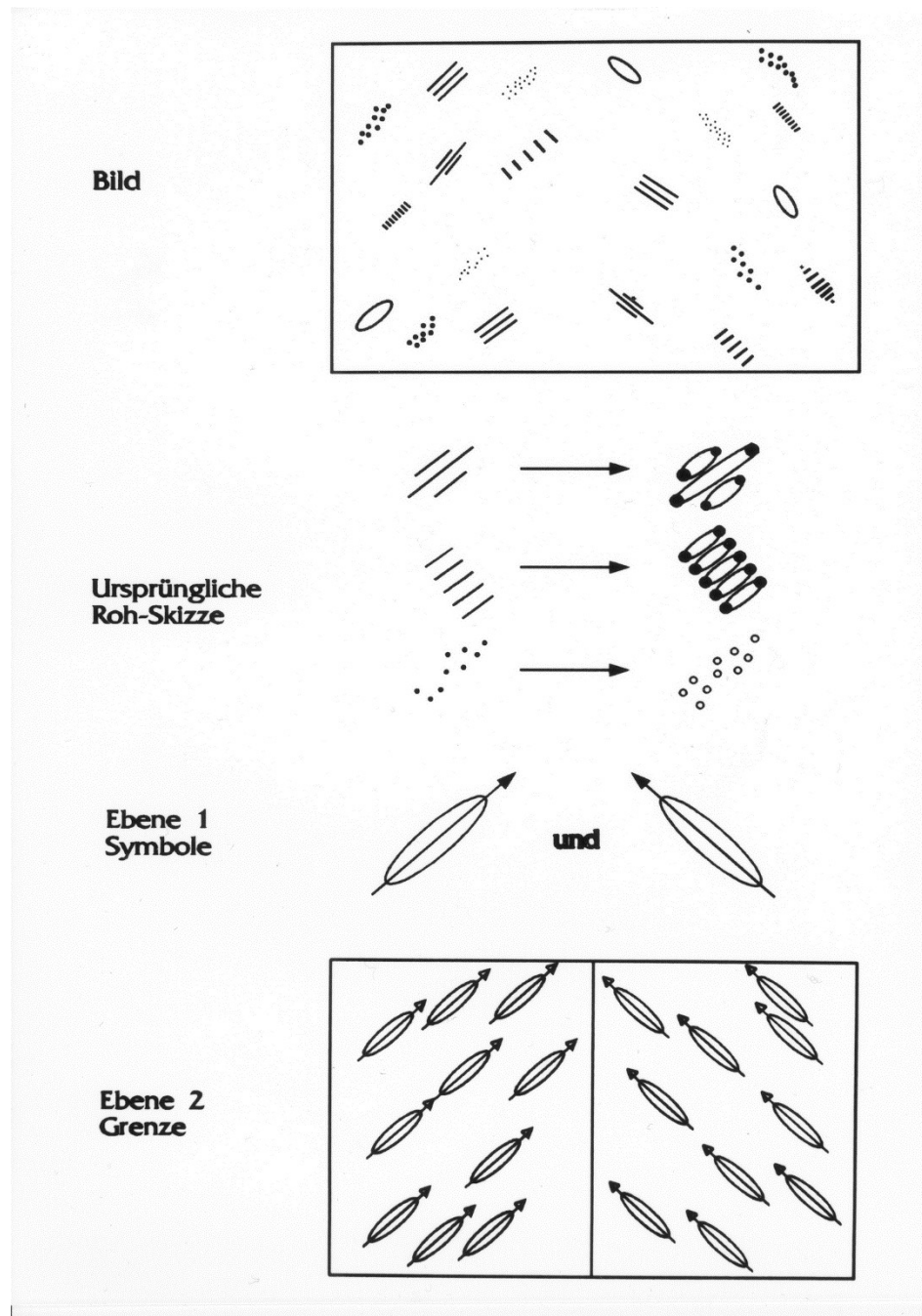


Abbildung 76: Schematische Darstellung der Verarbeitungsschritte bis zur Erstskeizze (Marr, 1982). Erläuterung im Text.

Im nächsten Schritt (Ebene 1) werden ähnliche Elemente gleicher Orientierung zu Zeichen höherer Ordnung zusammengefasst. Diese Symbole sind in der Abbildung 76 als Ellipsen dargestellt, deren Orientierung durch einen Vektor markiert wird. Schließlich (Ebene 2) wird eine Grenze zwischen den unterschiedlich orientierten Zeichen der linken und der rechten Bildhälfte konstruiert. Diese Grenze repräsentiert die Trennungslinie zwischen zwei Texturen, die sich auf der Dimension der Orientierung ihrer Elemente voneinander unterscheiden. Der Gruppierungsvorgang ist wesentlich umfangreicher und komplexer, als er in der Abbildung 76 dargestellt ist. Es gibt nicht nur zwei Ebenen der Gruppierung, sondern nahezu unbegrenzt viele, in denen immer wieder andere Ausgangselemente auf unterschiedlich feinem bzw. grobem Auflösungs-niveaus geordnet werden. Die Gruppierungsmechanismen suchen nach immer größeren Zeichen, die immer mehr Bildelemente vereinen. Dieser Prozess würde viel zu umfangreich und langwierig sein, wenn das visuelle System nicht auf die Probabilitäten der Außenwelt Bezug nehmen würde.

Der gerade beschriebene Prozess muss genauer betrachtet werden, denn er ist sowohl für computationale als auch für neuronale Modelle typisch. Die räumliche Verteilung der Helligkeitswerte lässt sich als eine mathematische Funktion darstellen. Von besonderer Bedeutung für die Beschreibung sind abrupte Änderungen der Helligkeitswerte; sie markieren Änderungen der räumlichen Tiefe oder Änderungen in der Orientierung von Oberflächen. Mathematisch lassen sie sich durch Extrem- und Wendestellen beschreiben. Die realistischerweise zu erwartenden Helligkeitsfunktionen enthalten verrauschte Informationen. Man muss die Kante glätten, bevor die Ableitungen gebildet werden können.

Um dieses Problem zu lösen, entwickelte Marr einen mathematischen Algorithmus, der das „Rauschen“ aus den Helligkeitsfunktionen herausfiltert und lokale Helligkeitsunterschiede durch eine Verstärkung

des Kontrasts zusätzlich akzentuiert. Der leitende Gedanke besteht darin, dass nicht Helligkeitsunterschiede zwischen zwei benachbarten Bildpixeln, sondern Helligkeitsunterschiede innerhalb einer kreisförmigen Region analysiert werden. Jeder Bildpixel innerhalb dieser Region wird durch das gewichtete Mittel der benachbarten Bildpixel ersetzt. Die einzelnen Pixel gehen mit unterschiedlicher Gewichtung in die Berechnung ein. Die Gewichtungsfaktoren richten sich nach einer Funktion, die Marr mit einem mexikanischen Hut verglichen hat und als *Zentrum-Umkreis-Filter* bezeichnet (vgl. Abbildung 77). Bildpixel innerhalb eines zentralen Kreises haben einen positiven Gewichtungsfaktor, der umso größer wird, je näher sie am Zentrum liegen. Der Kreis wird von einem Kreisring umgeben. Bildpixel innerhalb des Kreisringes erhalten negative Gewichtungsfaktoren. Der Wechsel der Polung führt zu einer Akzentuierung.

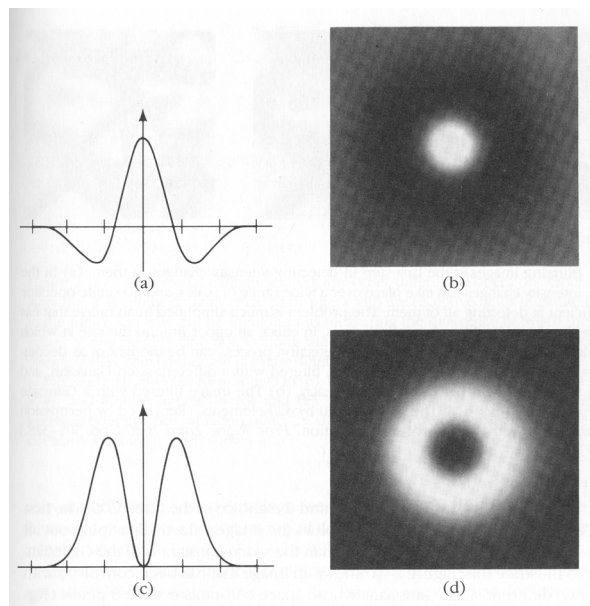


Abbildung 77: Wirkungsweise des mathematischen Filters in ein- und zweidimensionaler Darstellung (Marr, 1982); (a) und (b) veranschaulichen einen Filter, der im Zentrum aktivierend ($y > 0$) und im Umfeld hemmend ($y < 0$) wirkt, während bei (c) und (d) ein Filter gezeigt wird, der im Umfeld aktivierend ($y > 0$) und im Zentrum weder aktivierend noch hemmend wirkt ($y = 0$). Die Erregungsmuster ergeben sich jeweils durch einheitliche Beleuchtung der entsprechenden rezeptiven Felder.

Um eine Analyse der Bildstruktur auf unterschiedlichen Ebenen zu erreichen, werden Helligkeitssprünge mit Filtern berechnet, die sich hinsichtlich des Durchmessers der analysierten Bildregion unterscheiden. Dadurch können Helligkeitsveränderungen unterschiedlicher Ausdehnung registriert werden. Abbildung 78 zeigt die Analyse des Fotos einer Plastik von Henry Moore (a), mit von (b) bis (d) zunehmend größeren mathematischen Filtern. Sie entsprechen der aufwärtsgerichteten Informationsverarbeitung, bei der u. a. immer größere Teile des visuellen Feldes durch spezifischere Neuronenverbände analysiert werden. Es entstehen Kanten, Streifen, Ecken, Flecken und Konturen. Die Grundelemente zeichnen sich durch typische Intensitätsänderungen in ihren mathematisch gefilterten Helligkeitsfunktionen aus. Jedes einzelne der Elemente kann auf fünf Dimensionen beschrieben werden: Orientierung, Kontrast, Länge, Breite und Position.

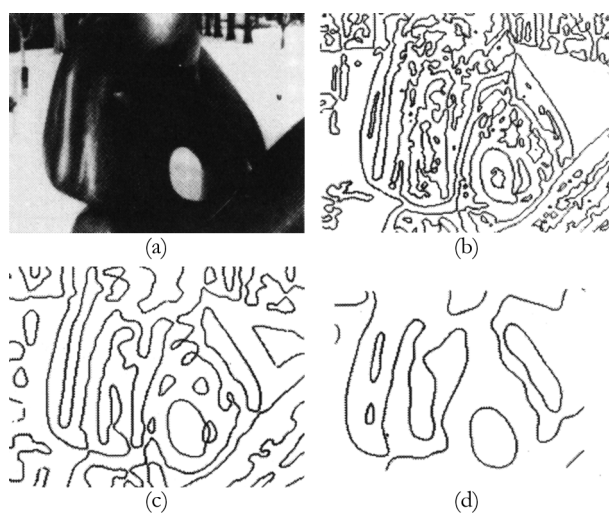


Abbildung 78: Bildzerlegung durch mathematische Filter (Marr, 1982).

Bei der Weiterverarbeitung werden zwei Ziele verfolgt: 1. die Bildung von Zeichen, die die Struktur der Oberfläche erfassen und 2. das Auffinden von Änderungen in den Parametern der Zeichen, die auf Änderungen der Tiefe oder Orientierung sichtbarer Oberflächen schließen lassen. Vereinfacht geht es darum, Zeichen und Grenzlinien dadurch zu bilden, dass die

Grundelemente zu *Zeichen höherer Ordnung* zusammengefasst werden. Bei dieser Gruppierung werden die Grundelemente nach dem Ähnlichkeitskriterium zu Linien, Kurven, größeren Flecken usw. geordnet. Die neu gebildeten Zeichen werden dann mit den gleichen fünf Parametern wie die Grundelemente (Orientierung, Kontrast, Länge, Breite und Position) beschrieben. Die resultierenden Zeichen werden zu höheren Zeichen zusammengefasst, bis sich keine höheren Zeichen mehr bilden lassen.

Neben der Bildung von Zeichen stellt die Bestimmung von *Grenzzlinien* das zweite grundlegende Problem auf dieser Stufe der Informationsverarbeitung dar. Hierzu können sechs Bildeigenschaften verwendet werden. Drei dieser Eigenschaften beziehen sich direkt auf die Zeichen: die durchschnittliche Orientierung, die Helligkeit und die Größe. Daneben werden drei Parameter eingesetzt, die sich auf die räumliche Anordnung beziehen: die lokale Dichte der Zeichen, die lokale Entfernung benachbarter, ähnlicher Zeichen und das aus der Orientierung benachbarter Zeichen entstehende Muster. Mittels dieser sechs Parameter lässt sich die Struktur von Oberflächen weiter aufschlüsseln. Abbildung 79 zeigt Beispiele, wie mit den Parametern Oberflächen strukturiert und Grenzzlinien gebildet werden können. Abbildung 79 (a) zeigt eine Grenzzlinie, die durch eine Änderung der Punktdichte erzeugt wird. In Abbildung 79 (b) wird sie durch eine Änderung der durchschnittlichen Größe der Flecken hervorgerufen. Die Grenze in Abbildung 79 (c) geht auf eine 45° Rotation der Zeichen, die in Abbildung 79 (d) und (e) auf mehrere Faktoren gleichzeitig zurück, (f) schließlich zeigt als Vergleich eine Grenze, die durch eine Linie gebildet wird.

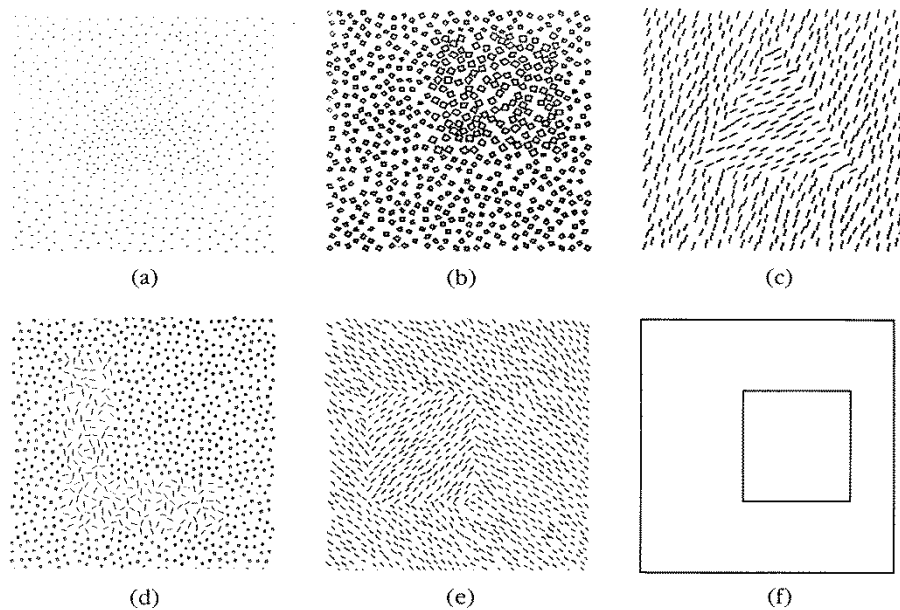


Abbildung 79: Konstruktion von Bildgrenzen durch Texturmerkmale (Marr, 1982).

Der Prozess des Erstellens der Erstskeizze am Beispiel des Katzenfells: Zunächst wird die Struktur einzelner Haare erfasst (Ebene einzelner Zeichen), dann die Zeichnung einzelner Teile des Fells (Ebene der Textur niedriger Ordnung) und zuletzt das Streifenmuster des Fells (Ebene der Textur höherer Ordnung; Repräsentation von Grenzlinien). Die Informationen aus allen Analyseebenen gehen in die bildbasierte Verarbeitung ein.

5. Oberflächenbasierte Verarbeitung

Die oberflächenbasierte Verarbeitung bildet die Schnittstelle zwischen Wahrnehmung und Kognition. Sie enthält das Maximum an Information, das durch frühe visuelle Verarbeitung zu gewinnen ist. Sie beinhaltet u. a. Informationen über die Orientierung und ungefähre Ausdehnung sichtbarer Oberflächen sowie eine symbolhafte Repräsentation der Übergänge zwischen Oberflächen unterschiedlicher Orientierung und Tiefe. Bei der computationalen Umsetzung nach Marr werden die einzelnen Bildpunkte in ein Koordinatensystem eingetragen, das aus den Sichtlinien der beiden Augen gebildet wird (2,5-D-Skizze). Abbildung 80

zeigt ein Beispiel dieser Repräsentationsform. An jedem Punkt des Koordinatensystems zeigt ein Vektor die räumliche Orientierung des Bildelementes an der betreffenden Stelle an. Tiefensprünge der sichtbaren Oberflächen werden durch durchgezogene Linien, Änderungen der Orientierung von Oberflächen durch gepunktete Linien dargestellt.

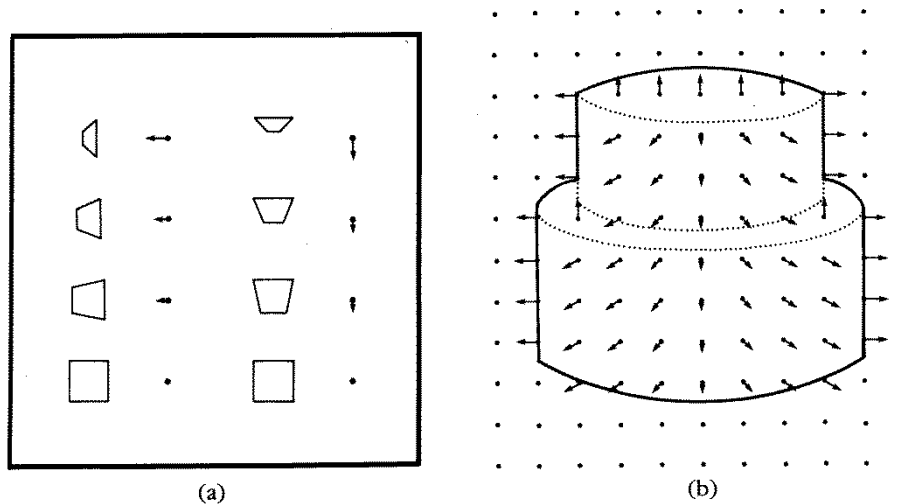


Abbildung 80: Illustration der 2,5-D-Skizze (Marr, 1982); (a) erläutert, wie verschiedene Neigungswinkel der Oberflächen in (b) mit Pfeilen dargestellt werden.

Bei der oberflächenbasierten Verarbeitung werden Informationen wie Schattierung, Texturgradienten, Bewegung und binokulare Disparität genutzt, die bei der bildbasierten Verarbeitung noch unberücksichtigt blieben. Insbesondere die Verarbeitung von Informationen aus der binokularen Disparität sowie aus der Bewegung stellen neue Probleme dar, die sowohl von einem künstlichen wie von einem biologischen System zu lösen sind.

a) Tiefenwahrnehmung

Beide Augen betrachten die sichtbare Welt aus leicht unterschiedlichen Blickwinkeln. Fixieren sie einen bestimmten Punkt (A) des visuellen Feldes, wird er auf das Zentrum der jeweiligen Retina (der Fovea)

abgebildet. Ein benachbarter Punkt (B) wird auf Netzhautregionen abgebildet, die sich in einem bestimmten Abstand zur Fovea befinden. Dieser Abstand ist bei beiden Augen verschieden, da sie den Punkt (B) aus unterschiedlichen Blickwinkeln erfassen. Aus der Differenz der beiden Abstände, der *binokularen Disparität*, lässt sich über einfache trigonometrische Berechnungen die Entfernung des Punktes (B) zur Beobachterin oder zum Beobachter berechnen. Infolgedessen ist die binokulare Disparität eine der wichtigsten Tiefeninformationen. Die einem Punkt des visuellen Feldes entsprechenden retinalen Abbildungen beider Augen werden als *korrespondierende Netzhautstellen* bezeichnet. Bei der Tiefenwahrnehmung liegt das zentrale Problem der Informationsverarbeitung darin, korrespondierende Netzhautstellen als solche zu erkennen. Jene Elemente in den beiden Bildern, die derselben Stelle im Raum entsprechen, müssen einander zugeordnet werden, ohne dass die Objekte oder Teile davon bereits erkannt sind. Wie werden korrespondierende Netzhautstellen aufgefunden?

Marr und Poggio (1976) haben Heuristiken für die binokulare Kombination von Bildpunkten aufgestellt, die das Korrespondenzproblem für Computerberechnungen handhabbar machen: 1. Vereinbarkeit: Korrespondierende Bildpunkte sind einander in ihren Parametern (lokale Dichte, lokale Entfernung, Orientierung) ähnlich. 2. Eindeutigkeit: Jeder Punkt eines Bildes entspricht genau einem Punkt des anderen Bildes. 3. Kontinuität: Zwei Bildpunkte werden dann bevorzugt miteinander kombiniert, wenn sie eine ähnliche Disparität aufweisen wie benachbarte Bildpunkte, die zuvor kombiniert wurden. Diese Heuristik trägt der Annahme Rechnung, dass sich die Tiefe stetig ändert und nur an den Objekträndern Sprünge auftreten.

b) Bewegungswahrnehmung

Viele Forscher gehen heute davon aus, dass bei der Verarbeitung von Bewegungsinformationen zwei Prozesse beteiligt sind. Bewegung wird

zum einen durch Bewegungsdetektoren und zum anderen durch komplexere Informationsverarbeitungsprozesse identifiziert. Auf der Detektorebene werden Bewegungen in sehr kleinen Zeitintervallen analysiert. Die Detektoren geben schnell verfügbare Hinweise darüber, wo im Gesichtsfeld eine Bewegung stattfindet und in welcher ungefähren Richtung sie verläuft. Sie liefern zweidimensionale Informationen über die Struktur bewegter Objekte, indem sie auf Diskontinuitäten im visuellen Feld reagieren, also Übergänge zwischen bewegten und unbewegten Oberflächen anzeigen. Aus diesen Diskontinuitäten der Bewegung lassen sich Objektgrenzen bestimmen. Eine komplexere Bewegungswahrnehmung steht immer noch vor zwei Problemen, die sich auf der Detektorebene kaum lösen lassen: 1. dem Korrespondenzproblem und 2. dem Problem, aus Bewegungen räumliche Strukturen zu ermitteln.

Das Korrespondenzproblem beschäftigt sich mit der Frage, welches Element des Bildes zum Zeitpunkt t_1 welchem Element zum Zeitpunkt t_2 entspricht. Das zweite Problem besteht darin, aus korrespondierenden Bildpunkten die Bewegungsvektoren und die dreidimensionale Struktur des Bildes zu bestimmen. Einen bekannten Versuchsaufbau hierzu hat Ullmann (1979) vorgelegt (vgl. Abbildung 81).

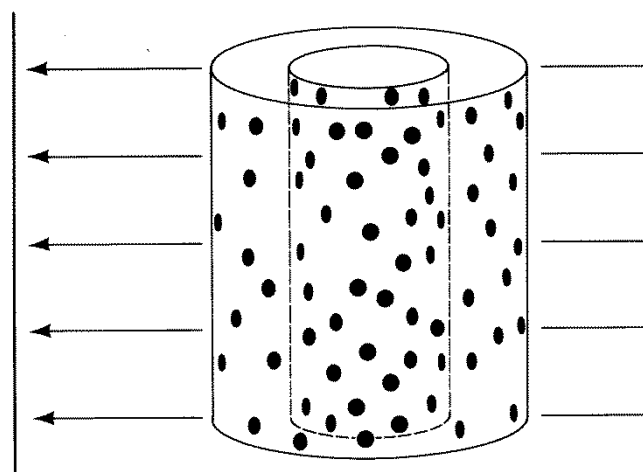


Abbildung 81: Gezeigt ist ein Schema der von Ullmann benutzten Versuchsaufbau.

Das Punktmuster zweier gegeneinander rotierender Zylinder wird auf eine Leinwand projiziert. Der Beobachterin oder dem Beobachter wird keine kontinuierliche Bewegung, sondern eine Abfolge von Einzelbildern gezeigt. Jedes einzelne Punktmuster für sich ist nicht interpretierbar. Aus einer Abfolge mehrerer Einzelbilder lässt sich die Bewegung und die räumliche Struktur der bewegten Objekte rekonstruieren (Marr, 1982).

Die Erklärung dieses Bewegungsphänomens vereinfacht sich durch die sogenannte Starrheitsannahme. Die meisten Strukturen der visuellen Welt sind näherungsweise starr. Die Form einer Struktur ändert sich während einer Bewegung nicht, es ändert sich lediglich die Position der Struktur im Raum. Ein Gegenbeispiel wäre ein fliegender Vogel, dessen Gestalt sich während des Fluges ändert. Für den Fall der Bewegung starrer Körper reduziert sich das Korrespondenzproblem der Bewegungswahrnehmung auf das bekannte Korrespondenzproblem der Tiefenwahrnehmung: Aus der veränderten Position kleiner Gruppen von Punkten werden Bewegungsvektoren berechnet, die die genaue räumliche Verschiebung dieser Punktgruppen wiedergeben. Auf diese Art und Weise erhält man ein Feld von Bewegungsvektoren, aus dem sich räumliche Strukturen ableiten lassen. Selbstverständlich werden im Gehirn keine Bewegungsvektoren berechnet. Es handelt sich wieder um ein Modell. Aber es ist davon auszugehen, dass vergleichbare Prozesse am Werk sind, die es ermöglichen, Objekte trotz Eigen- oder Fremdbewegung zuverlässig zu identifizieren.

6. Objektbasierte Verarbeitung

Ausgangspunkt für die Beschreibung der objektbasierten Verarbeitung ist die Frage nach der Repräsentation visueller Informationen, damit sie mit Gedächtnisinhalten verglichen werden können. Die oberflächenbasierte Verarbeitung liefert eine unzureichende Grundlage, da sie keine Anhaltspunkte über verborgene (z. B. verdeckte) Oberflächen enthält. Außerdem ist sie standortzentriert, liefert also eine von der Perspektive der

Beobachterin oder des Beobachters abhängige Beschreibung des Objekts. Zu jedem Objekt existieren unendlich viele oberflächenbasierte Beschreibungen, was eine Objekterkennung, also eine Übereinstimmung mit Gedächtnisinhalten, unwahrscheinlich macht (vgl. Problem der inversen Projektion). Eine geeignete objektbasierte Repräsentation sollte folgenden Kriterien genügen: 1. Sie muss unabhängig von der Perspektive sein (diese Anforderung erfüllt z. B. eine dreidimensionale Computergraphik, die sich auf dem Bildschirm um alle Achsen beliebig rotieren lässt). 2. Die Repräsentation sollte hierarchisch organisiert sein: Für das gesamte Objekt und für seine Teile sollten unabhängige Beschreibungen vorliegen. Es ist zu bedenken, dass Teile von Objekten wiederum Objekte sein können, je nachdem, wie nah sich die Beobachterin oder der Beobachter am Objekt befindet und auf welche Bereiche die Aufmerksamkeit gerichtet ist.

Die Gestalt eines Menschen wäre schwer wiederzuerkennen, wenn nur eine Beschreibung für den gesamten Körper vorläge. Ein sitzender Mensch sieht anders aus als ein stehender. Trotzdem werden beide Gestalten als „Mensch“ erkannt. Eine Möglichkeit, dieses Problem zu lösen, besteht darin, dass neben einer Repräsentation des ganzen Körpers auch Beschreibungen einzelner Körperteile und ihrer Lage zueinander vorliegen. Es wäre eine Repräsentation eines einzelnen Armes denkbar, die die Elemente Ober- und Unterarm enthält. Daneben mag eine Repräsentation eines Unterarmes bestehen, die die Elemente Unterarm und Hand trennt. Eine weitere Repräsentation unterscheidet Handteller und Finger etc. (vgl. Abbildung 82).

Ein sitzender Mensch unterscheidet sich von einem stehenden Menschen nunmehr durch eine unterschiedliche Orientierung der Beine zum Rumpf sowie in der relativen Orientierung von Ober- und Unterschenkel. In den übrigen Repräsentationen bestünden keine Unterschiede. Eine hierarchische Organisation verschiedener Repräsentationen erlaubt es, variable Formen zu erkennen.

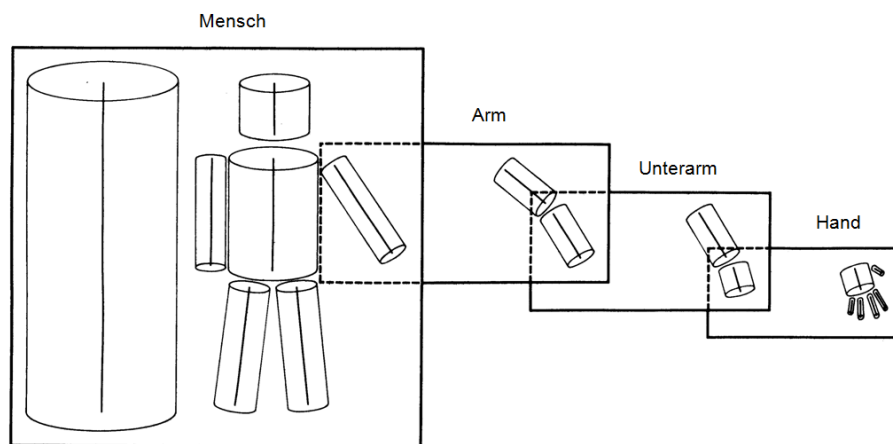


Abbildung 82: Hierarchische Repräsentation einer Form in der dreidimensionalen Beschreibung (Marr, 1982).

Nach Marr sollte ein Modell der Formrepräsentation vor allem folgende drei Aspekte beinhalten: 1. das Koordinatensystem, in dem die Form dargestellt wird, 2. die Basiselemente, mit denen eine Form beschrieben werden kann und 3. die Organisationsform der Repräsentation, also die Kombination der Basiselemente und ihre Relationen zueinander. Eine einfache Illustration dieser Grundidee zeigt Abbildung 83. Die abgebildeten Pfeifenreiniger-Figuren können identifiziert oder zumindest klassifiziert werden, obwohl nur die Achsen wesentlicher Teile repräsentiert sind. Stellt man sich um die Achsen zylindrische Körper vor, gelangt man zu der Repräsentationsform, wie sie bereits in Abbildung 82 gezeigt wurde.

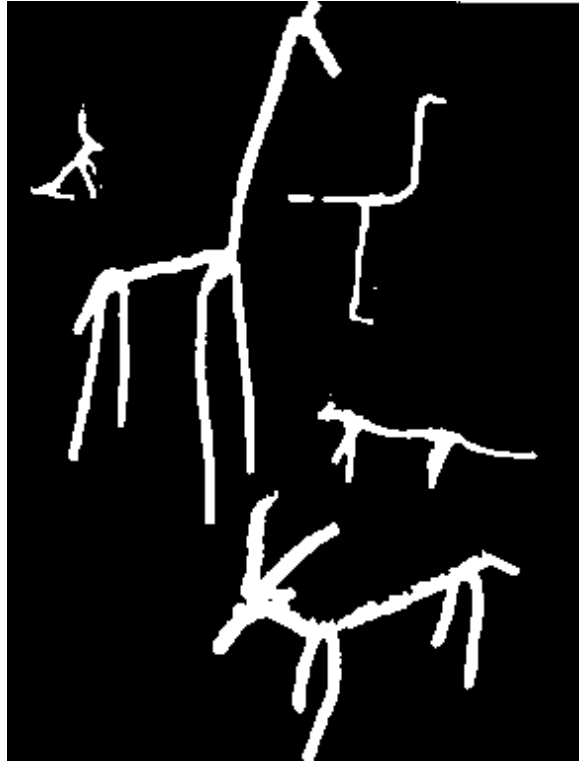


Abbildung 83: Pfeifenreiniger-Figuren (Marr, 1982). Erläuterung im Text.

Nach Marr sind zum Erkennen von Formen keine genau analysierten Oberflächenstrukturen von Bedeutung. Formen werden direkt aus der relativen Orientierung stilisierter Untereinheiten zueinander erkannt. Diese bilden die Basiselemente der Form (2. Aspekt). Die Organisation der dreidimensionalen Repräsentation wurde am Beispiel der Abbildung eines menschlichen Körpers bereits angedeutet. Formen werden demnach hierarchisch organisiert. Auf mehreren, unterschiedlich feinen Analyseebenen liegen Repräsentationen des Ganzen und seiner Untereinheiten nebeneinander vor (3. Aspekt). Das Koordinatensystem, anhand dessen die räumlichen Relationen innerhalb einer Form berechnet werden, wird nun nicht mehr durch die Sichtlinien konstituiert, sondern durch die Hauptachsen der Form (Linien in Abbildung 82). Auf der Analyseebene „Mensch“ bildet die Achse des links abgebildeten Zylinders die Hauptachse der Form. Sie liefert das Bezugssystem für die Achsen der Untereinheiten Kopf, Rumpf und Gliedmaßen (Figur rechts neben dem Zylinder), deren Orientierung zur Hauptachse berechnet wird (1. Aspekt).

Die dreidimensionale Repräsentation besteht aus hierarchisch organisierten Untereinheiten einer Form, die auf jeder Analyseebene in ein eigenständiges Koordinatensystem eingetragen werden, das sich nach der jeweiligen Form richtet. Anhand der dreidimensionalen Darstellung sind Vergleiche mit in gleicher Weise repräsentiertem Gedächtnismaterial möglich.

7. Kategorienbasierte Verarbeitung

Wenn das Ziel eine auf die Bedürfnisse, Ziele und Wünsche eines Organismus abgestimmte Repräsentation ist, dann reicht die objektbasierte Repräsentation nicht aus. Notwendig ist eine Darstellung mit funktionalen Aspekten. Erzielt wird sie über einen Prozess der Kategorisierung. Marr hat für diese Stufe keine Annahmen formuliert. Palmer (1999, S. 91) geht davon aus, dass zwei eigenständige Schritte beteiligt sind. Zunächst identifiziert das visuelle System ein Objekt aufgrund seiner sichtbaren Eigenschaften wie Form, Größe, Farbe und dem Ort, an dem es sich befindet, als Mitglied einer Kategorie. In einem zweiten Schritt erhält es aufgrund dieser Klassifikation den Zugang zu einer großen Menge gespeicherter Informationen über typische Eigenschaften der Objekte, die zu dieser Kategorie gehören. Das beinhaltet auch Informationen zur Funktion und zu Erwartungen in Bezug auf zukünftiges Verhalten. Ein Beispiel: Wenn das Objekt in Abbildung 75 als Becher klassifiziert wurde, ist klar, dass es geeignet ist, a) Flüssigkeit aufzubewahren und b) daraus zu trinken. Zwei getrennte Operationen zu postulieren hat den Vorteil, dass jede beliebige funktionale Eigenschaft mit jedem beliebigen Objekt assoziiert werden kann. Die Relation zwischen der Form eines Objektes und der Information, die über seine Funktion gespeichert ist, kann zufällig sein.

Neben der Idee einer indirekten Wahrnehmung der Funktion ist die Vorstellung einer direkten Wahrnehmung ohne vorherige Klassifikation verbreitet. Gibson (1979) hat das Konzept der „affordances“

vorgeschlagen. Demnach ist einer Oberfläche und auch einem Objekt anzusehen, welche Handlungen möglich sind. Man sagt auch, ein Objekt habe Aufforderungscharakter. Die Oberfläche enthält Angebote, welche Bewegungen und Handlungen möglich sind (vgl. Abbildung 84). Schlussfolgernde kognitive Prozesse oder ein Vergleich mit Gedächtnisinhalten sind nicht notwendig.

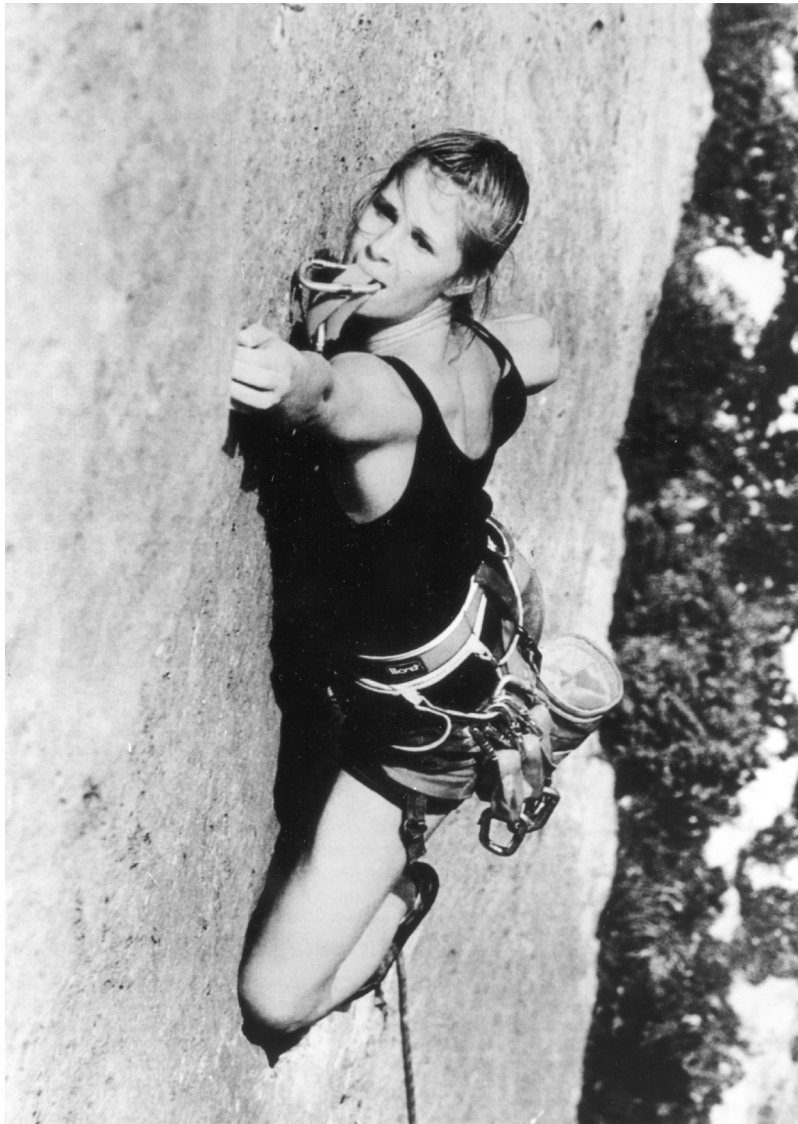


Abbildung 84: Angebote sind auf die Motive, Ziele und Bedürfnisse eines Individuums bezogen. Das Bild zeigt eine Extremkletterin. Für sie stellt die Felswand ein „Angebot“ zum Klettern dar (Foto von Gerard Kosicki, aus Kaminski und Immelmann, 1987).

Schließlich könnte es auch sein, dass beide Arten (direkte und indirekte Wahrnehmung) der Funktion eine Rolle spielen bzw. sich ergänzen. Bei einigen Objekten, wie einem Stuhl oder einer Tasse, sind die sichtbare Form und die Funktion stark aufeinander bezogen, sodass sie vielleicht nicht kategorisiert werden müssen, bevor sie benutzt werden können. Bei anderen Objekten (z. B. einer Fernbedienung) ist die Nutzung derart weit entfernt von den visuellen Eigenschaften, dass für ihre richtige Verwendung nicht einmal eine Klassifikation ausreicht.

Suchbild und Bildtitel

Der Einfluss von Wissen und Erfahrung auf die Wahrnehmung ist im Regelfall nicht zu beobachten, da diese Prozesse ähnlich schnell ablaufen wie die Bottom-up-Prozesse und ebenfalls nicht bewusst sind. Es ist daher notwendig, noch einmal „Bruchstellen“ aufzusuchen, um das Zusammenspiel näher zu bestimmen. Zwei kritische Ereignisse sind Such- und Vexierbilder (vgl. Abbildungen 85 und 86; bitte betrachten Sie die Abbildungen zunächst, ohne die Bildunterschrift zu lesen, und fahren erst dann im Text fort). In Suchbildern ist für die Betrachterin oder den Betrachter zunächst keine einheitliche Interpretation zu gewinnen. Es bleibt unsicher, was auf dem Bild dargestellt ist. Wird verbal eine ergänzende Information gegeben (z. B.: „In der Abbildung 85 ist ein Dalmatiner zu sehen.“), setzen aufgrund des Hinweises neue Suchprozesse ein. Hypothesengeleitet testet die Betrachterin oder der Betrachter, welche Gruppierung der Flecken einen Dalmatiner ergeben könnte. Die Suche ist nicht einfach, weil der Hund nicht in einer kanonischen Ansicht gezeigt wird. Ist die Figur entdeckt, kann sie jederzeit ohne Schwierigkeiten wiedergefunden werden. Bei den Vexierbildern ist die Ausgangssituation etwas anders. Hier existiert nach kurzer Betrachtung eine in sich schlüssige Interpretation (z. B.: „Die Abbildung 86 stellt eine Küstenlandschaft mit Bäumen dar; in der Ferne ist ein Segelschiff zu sehen.“). Der Suchprozess wird abgebrochen. Erhält die Betrachterin oder der Betrachter einen Hinweis auf den Bildtitel („Napoleon auf Elba“), beginnt die Suche erneut.

Die Hypothese lautet: Wenn dies ein Bild von Napoleon ist, dann muss sich eine Darstellung Napoleons auf dem Bild befinden. Folglich generiert die Betrachterin oder der Betrachter zunächst sein Vorstellungsbild von Napoleon und sucht dann den distalen Stimulus erneut ab. Häufig wird erkannt, dass sich eine typische Profilansicht Napoleons zwischen den beiden Bäumen im Vordergrund befindet. Eigentlich ist die Darstellung genügend groß und zentral positioniert, um unmittelbar erkannt zu werden.



Abbildung 85: Suchbild. Gesucht wird ein Dalmatiner.



Abbildung 86: Vexierbild. Gesucht wird die Gestalt Napoleons.

Der Wechsel von der einen Interpretation zur anderen erfolgt ohne Übergang. Offensichtlich wird dieselbe Stimulation auf einmal neu geordnet. Ähnlich wie bei einem Necker-Würfel entsteht Multistabilität.

Wird das Vorgehen auf der theoretischen Ebene rekonstruiert, ergibt sich folgender Ablauf: Die verbale Information aktiviert im semantischen Gedächtnis das Wissen über Napoleon und hierüber ein Vorstellungsbild mit den charakteristischen Eigenschaften seiner Figur. Das Vorstellungsbild ist in der Lage, im Sinne der Top-down-Verarbeitung über die Aufmerksamkeit die Blickbewegungen zu steuern. Hypothesengeleitet wird der distale Stimulus erneut abgesucht. Es entstehen wie auch beim ersten Mal Gruppierungen der Helligkeitssprünge und Intensitätsverläufe entsprechend den Gestaltgesetzen. Allerdings wird jetzt der Schwellenwert für eine einzelne Figurbildung gegenüber allen anderen herabgesetzt (Napoleon). Immer dann, wenn der distale Stimulus es zulässt, wird eine Gruppierung bevorzugt, die Ähnlichkeit mit dem Vorstellungsbild von Napoleon hat. Der Suchprozess wird abgebrochen, wenn die Hypothese verifiziert ist. Kann sie nicht bestätigt werden, bleibt ein Spannungssystem, ähnlich einer unerledigten Aufgabe, für längere Zeit bestehen.

Die Beispiele können die Funktion eines Bildtitels verdeutlichen und klären, warum zum Beispiel Museumsbesucher*innen die Beschriftung lesen, bevor sie das Bild betrachten. Der Titel suggeriert den Anspruch, das Bild unter einer Interpretation zusammenzufassen. Er reduziert die Unsicherheit der Betrachterin oder des Betrachters und liefert eine Hypothese für die einheitliche und sinnvolle Gruppierung aller Inhomogenitäten im Bild. Er soll helfen, den Suchprozess abzukürzen oder, bei einer bereits länger andauernden Suche, das Spannungssystem zu reduzieren. Dazu werden über den Titel Vorstellungsbilder generiert, die mit dem aktuellen Perzept verglichen werden. Das Bedürfnis nach einem Bildtitel entspricht dem Ökonomieprinzip der Umgebungswahrnehmung.

Hier ist die Beobachterin oder der Beobachter gewohnt, dass zeitgleich mit der Wahrnehmung auch die Bedeutung präsent ist. Der Bildtitel verweist auf außerbildliche Zusammenhänge. Er kann auch eingesetzt werden, um die Betrachterin oder den Betrachter zu irritieren. Die Abbildung 87 (a) zeigt erneut ein Suchbild. Solange nicht bekannt ist, dass es sich um eine Jesusdarstellung handelt, gibt es kaum eine Chance, die unterschiedlichen Flecken im Bild zu einer Hypothese zusammenzufassen. Würde man die Titel beider Suchbilder ändern, führte die Suche in beiden Fällen ins Leere. Eindeutigkeit wird nicht erreicht, die Suche kann nur abgebrochen werden.



(a) (b)

Abbildung 87: Suchbild. Jesusdarstellung (a) und seine Rotation (b).

Eine Rotation der Abbildung 87 (a) um 180° macht es auch bei Kenntnis des richtigen Titels unmöglich, den Jesuskopf wiederzuerkennen (vgl. Abbildung 87 b). Wissen kann also bei der Ambiguitätsreduktion helfen, das Potential ist aber begrenzt. Der Lösungsraum wird durch die autochthonen Gruppierungen vorgegeben. In diesem Sinne hilft die Erfahrung bei der Interpretation eines Reizmusters, aber sie kann keine Eindeutigkeit erzwingen. Zumindest in dieser Situation gibt es keine Priorität des Wissens vor dem Sehen.

Zum umgekehrten Fall: Kann ein distaler Stimulus etwas zeigen, das noch nicht gesehen wurde, etwas, das der bisherigen Erfahrung widerspricht? Die meisten Betrachter*innen berichten, wenn sie nach einem naiv-phänomenalen Urteil gefragt werden, dass Abbildung 88 ein transparentes Messer vor einem Glas zeige. Kaum eine Betrachterin oder ein Betrachter dürfte jemals ein durchsichtiges Messer gesehen haben. Kritisch-phänomenal wissen wir: Messer sind normalerweise aus Metall und Metalle sind undurchsichtig. Wir sehen in diesem Fall etwas gegen besseres Wissen.

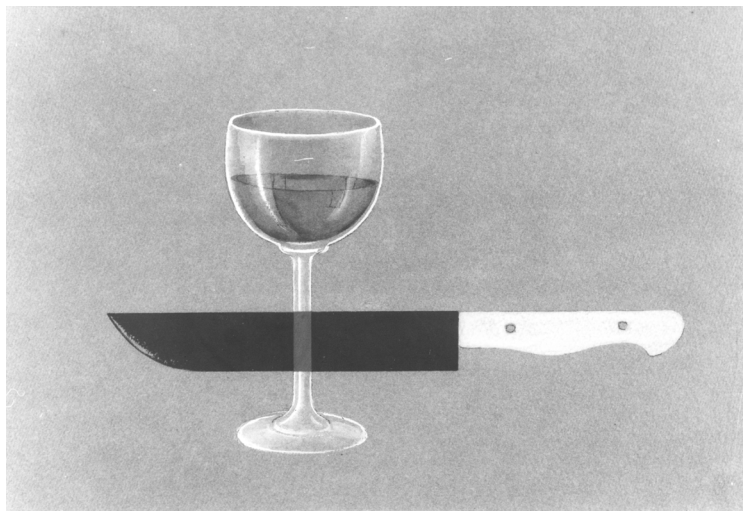


Abbildung 88: Durchsichtiges Messer (Kanizsa, 1979).

Die Heuristiken, die im Zusammenhang mit der Helligkeitswahrnehmung, der Figur-Grund-Unterscheidung und der Tiefenwahrnehmung wirksam sind, führen dazu, dass die Betrachterin oder der Betrachter keine Wahl hat. Für die Meisten ist dieses Bild nicht einmal mehrdeutig. Sollte das Prinzip der Probabilität die Wahrnehmung eines durchsichtigen Messer nicht verhindern? Denn schließlich steht dieses Perzept doch im Gegensatz zu den empirischen Wahrscheinlichkeiten für ein solches Ereignis. Aber 1. führen Heuristiken nicht immer zu einer richtigen Lösung und 2. können verschiedene Heuristiken in Konflikt zueinander stehen. Letzteres ist hier der Fall.

Verlegt in eine Situation der Umgebungswahrnehmung und nicht der Bildwahrnehmung, hätte eine Beobachterin oder ein Beobachter die Möglichkeit, die Hypothese, dass sich das Messer vor dem Glas befindet und deshalb durchsichtig sein muss, durch eine Veränderung des Standortes zu überprüfen. Die Zirkularität würde Sicherheit schaffen. Gegebenenfalls müsste das Wissen auf der Basis einer neuen Erkenntnis erweitert werden: Messer können aus einem Material bestehen, das lichtdurchlässig ist. Bei der Bildwahrnehmung aber liefert uns der Wechsel der Perspektive keine neuen Informationen.

8. Ambiguitätsreduktion durch Ergänzung

Die FACADE-Theorie von Stephen Grossberg ist gleich aus mehreren Gründen ein typischer Vertreter für die Modelle zur visuellen Informationsverarbeitung: 1. Sie wurde schrittweise und in engem Kontakt zu den empirischen Befunden entwickelt und revidiert. 2. Sie ist bemüht, sich der Komplexität der zugrundeliegenden Prozesse zu nähern und die neuronalen Strukturen zu benennen, die beteiligt sind. 3. Sie verwendet eine eigene, künstliche Begrifflichkeit.

Die Mehrdeutigkeit der Stimulation wird in einem hierarchischen Prozess mithilfe von Heuristiken reduziert (Grossberg 1994, 1997; Kelly und Grossberg, 2000). Genutzt wird vor allem das Prinzip der Ergänzung. Die Theorie nimmt an, dass auf den Verarbeitungsstufen unterschiedliche Ergänzungsprozesse ablaufen. Im Mittelpunkt stehen Prozesse der Komplettierung, die auf den Stufen der bildbasierten- und der oberflächenbasierten Verarbeitung stattfinden. Mit Fortschreiten des Analyse- und Konstruktionsprozesses findet keine kontinuierliche Reduktion von Unsicherheit statt, sondern die Auflösung der Ambiguität kann neue Mehrdeutigkeiten hervorbringen, die in einem späteren Schritt wieder reduziert werden müssen. Einen wichtigen Grundsatz zur Überwindung dieser Mehrdeutigkeiten sieht die FACADE-Theorie in der

zunächst separaten Analyse von Kanten, Oberflächen und ihren Merkmalen. Ein *Boundary Completion System* (BCS) für die Erkennung von Kanten und Konturen steht einem *Feature Completion System* (FCS) für die Analyse von Oberflächen und ihren Merkmalen gegenüber. Entscheidend ist, dass die beiden Systeme *komplementäre* Eigenschaften der Stimulation analysieren und vervollständigen: Ein Bereich, der eine einheitliche Oberfläche darstellt, kann keine Kanten oder Konturen beinhalten. Gleichzeitig grenzen Kanten verschiedene Oberflächen voneinander ab. Das visuelle System muss entscheiden, welche der im distalen Reiz vorhandenen Kontraste Teil einer Oberfläche und welche Kontraste Teil einer für die Formerkennung bedeutsamen Kontur sind. Wurden entsprechende Konturen identifiziert, können Oberflächeneigenschaften wie Farbe oder Helligkeit im Zuge von Ausbreitungsprozessen (*Filling-In*) durch das FCS in den umgrenzten Bereichen eingefügt werden.

Auf neuronaler Ebene werden beide Systeme im ventralen (okzipital-temporalen) Verarbeitungspfad verortet. Für beide werden neuronale Strukturen im Corpus Geniculatum Laterale (CGL, befindlich im Thalamus) sowie in verschiedenen Bereichen des visuellen Cortex angenommen. Prozesse der abschließenden Synthese werden in V4 vermutet. Entscheidend ist, dass die FACADE-Theorie – im Gegensatz zu vielen früheren Ansätzen – keine Unabhängigkeit der verschiedenen Verarbeitungspfade annimmt, sondern auf allen Verarbeitungsebenen eine enge Kooperation zwischen BCS und FCS sowie dem dorsalen Pfad voraussetzt. Wie sich bei der Erläuterung der einzelnen Ebenen von BCS und FCS zeigen wird, gewährleistet nur die Interaktion beider Systeme eine effektive Reduktion der Ambiguität.

Boundary Completion System (BCS)

In Bezug auf die Bedeutung von Kanten und Ecken unterscheiden sich die phänomenale und die physiologische Ebene. Während im Perzept die

Objekte und ihre Oberflächen dominieren, ist im Informationsverarbeitungsprozess die Identifikation von Grenzen die zentrale Voraussetzung für alle weiteren Analyseschritte. Nur wenn sie erfolgreich ist, können Figuren und damit auch Objekte erkannt werden. Als Basis für die Kantenanalyse werden *einfache Zellen* in V1 angesehen, die auf Kontraste mit einer bestimmten Orientierung in einem eng umgrenzten Bereich der Retina reagieren. Entsprechend werden sie als *orientierungssensitive lokale Kontrastdetektoren* bezeichnet. Die Existenz dieser Zellen ist seit langem gut belegt. Sie integrieren Informationen einer Vielzahl rezeptiver Felder der Retina und des CGL. Durch die Zusammenfassung mehrerer hochauflösender, einzelner Signale werden Informationen über die Ausrichtung längerer Kantenbereiche gewonnen. Die Unsicherheit über den exakten Anfang und das genaue Ende dieser Kante wird währenddessen verstärkt, da sich 1. die rezeptiven Felder der einfachen Zellen über einen größeren Bereich als die tatsächliche Kante erstrecken können und sich 2. die rezeptiven Felder überschneiden können.

In V1 existieren in einer retinotopen Organisation für jeden Bereich der Netzhaut Neurone, die auf lokale Kanten festgelegter Orientierung reagieren. Außerdem finden sich hier Sensitivitäten für Kontraststärken und die Kontrastpolarität (Richtung des Kontrastes). Die lokalen Kontrastdetektoren der einfachen Zellen trennen abrupte Helligkeitsunterschiede, die Teile einer Kontur sein können, von kontinuierlichen Helligkeitsverläufen, die einen Teil einer Oberfläche darstellen. Informationen über Orientierung und Kontraststärke der Kanten werden zur weiteren Verarbeitung zunächst an *komplexe Zellen* des V1 weitergeleitet. Sie haben größere rezeptive Felder und reagieren ebenfalls auf Linien oder Kanten bestimmter Orientierung. Im Unterschied zu den einfachen Zellen spielt es keine Rolle, wo im rezeptiven Feld sich die Linie oder Kante befindet, allein ihr Vorhandensein ist wichtig.

Das Auftreten von Kanten ist nicht hinreichend, um zuverlässig auf das Vorhandensein einer bedeutsamen Grenze für die Figurerkennung zu schließen. Auch Oberflächen können, z. B. als Bestandteil einer Textur, Kanten enthalten. Die zugehörige Heuristik lautet: Nur wenn sich die Summe der lokalen Kontraste zu einer geschlossenen Kontur verbinden lässt, kann eine Figur umgrenzt und später durch Filling-In-Prozesse vervollständigt werden. Da in der Umgebungswahrnehmung Teile der Kanten eines Objektes verdeckt sind, muss das visuelle System in der Lage sein, diese Teile zu ergänzen. Um längere, über die lokalen rezeptiven Felder einzelner komplexer Zellen hinausgehende Kanten zu erkennen, müssen die Teile vervollständigt werden. An den Enden der Kanten werden sogenannte *End Cuts* generiert, die das Ende einer Kante anzeigen und so die Kante über Ecken hinweg mit Kanten einer anderen Orientierung verbinden.

Doch wie entstehen End Cuts und was ist ihre Funktion? Möglich wird die Bildung von End Cuts durch die Vernetzung *hyperkomplexer Zellen* im V1; sie erhalten exzitatorische (erregende) Signale von komplexen Zellen mit entsprechender Orientierung. Auf Ebene der hyperkomplexen Zellen finden zwei kompetitive Prozesse statt: Die hyperkomplexen Zellen inhibieren (hemmen) zunächst benachbarte hyperkomplexe Zellen gleicher Orientierung. In einem nächsten Verarbeitungsschritt werden hyperkomplexe Zellen höherer Ordnung auf der gleichen Position, jedoch anderer Orientierung durch entsprechende Zellen gegensätzlicher Orientierung gehemmt: Am Ende von Kanten entsteht eine besonders starke Disinhibition (Erregung durch wegfallende Inhibition) von Kanten gegensätzlicher Orientierung. Auf diese Weise wird die Unsicherheit hinsichtlich des Endes einer Kante reduziert und eine Verbindung der Kantenenden mit Kanten anderer Orientierungen ermöglicht. Die Generierung von End Cuts auf Ebene der hyperkomplexen Zellen des V1 wird durch das Perzept, das bei einer Betrachtung der Ehrenstein-Figur in Abbildung 89 entsteht, veranschaulicht: Die Linienenden induzieren End

Cuts und das BCS generiert – bedingt durch die Disinhibition gegensätzlicher Winkel, die miteinander verbunden werden – das Perzept eines Kreises entlang der Linienenden.

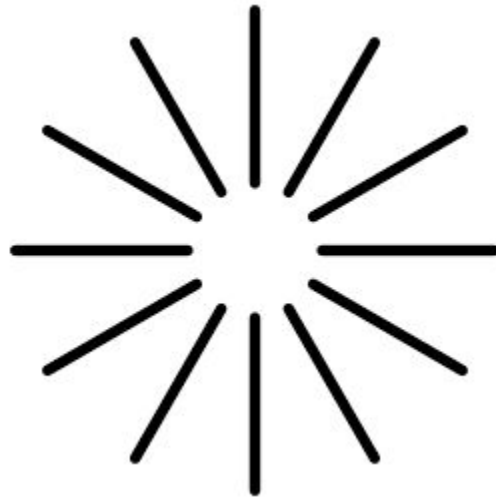


Abbildung 89: Ehrenstein-Figur. Es wird ein weißer Kreis gesehen, obwohl alle weißen Flächen dieselbe Leuchtdichte aufweisen.

Wegen seiner großen Bedeutung wird der Prozess der Kantenergänzung durch das BCS genauer betrachtet. Am Beispiel der Ehrenstein-Figur stellt sich die Frage, warum der weiße Kreis auf schwarzen Linien zu liegen scheint, die meist als hinter dem Kreis liegend und als durchgehend beschrieben werden? Warum werden unterbrochene Linien zu einer perzeptuellen Einheit verbunden? Die FACADE-Theorie nimmt an, dass die Ergänzungen durch bipolare Zellen geschehen, die miteinander und mit den hyperkomplexen Zellen des V1 vernetzt sind. Diese Interaktion findet in Form eines *CC-Loops* (Comperative-Cooperative Loop) statt und ist in Abbildung 90 dargestellt: Zwei ausreichend aktivierte hyperkomplexe Zellen mit rezeptiven Feldern gleicher Orientierung wirken exzitatorisch auf zwei bipolare Zellen gleicher Position und Orientierung (1). Diese beiden hyperkomplexen Zellen stellen die sich gegenüberliegenden Enden der bisher erfassten, sichtbaren Kanten dar. Die parallele Aktivierung führt zu einer Exzitation der dazwischenliegenden hyperkomplexen Zelle gleicher Orientierung (2).

Auf Ebene der hyperkomplexen Zellen kommt es somit zu einer kompetitiven Aktivierung durch bipolare Zellen, komplexe Zellen niedrigerer Verarbeitungsstufen und durch Interaktionen mit anderen hyperkomplexen Zellen anderer Orientierungen. Wird die entsprechende hyperkomplexe Zelle ausreichend aktiviert, führt dies zu einer Exzitation der zugehörigen bipolaren Zelle (3), die in Verbindung mit anderen bipolaren Zellen neue hyperkomplexe Zellen aktiviert (4). Die Kante wird schließlich durch die gleichzeitige Aktivierung hyperkomplexer Zellen einer bestimmten Orientierung entsprechend einer Feedbackschleife vervollständigt.

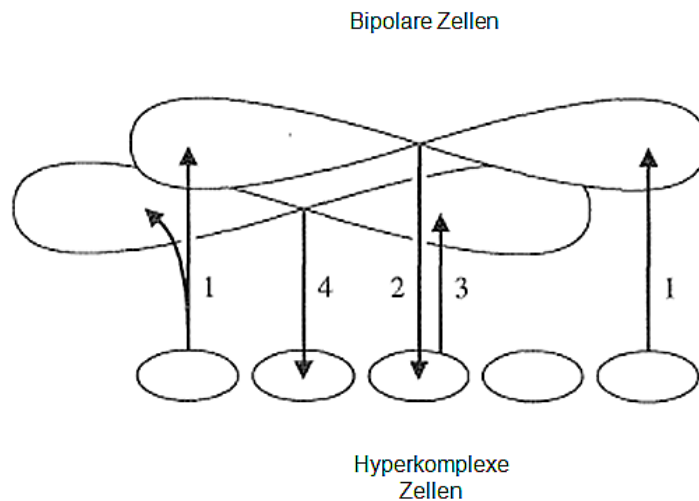


Abbildung 90: Beispiel eines CC-Loops (Grossberg, 1994). Die beiden äußeren hyperkomplexen Zellen korrespondieren mit den beiden Enden einer Linie bzw. einer Kante. Weitere Erläuterung im Text.

Vor allem die kompetitive Verarbeitung trägt dazu bei, die statistisch wahrscheinlichste Lösung der Kantenstruktur zu ermitteln. Die kompetitive Verarbeitung führt, wie in ähnlicher Weise bereits bei den Zellen der Retina gezeigt wurde, zu einer Kontrastverstärkung, die es ermöglicht, eine Kante viel stärker als Trennung zweier Oberflächen wahrzunehmen, als dies möglich wäre, wenn nur die Stimulation betrachtet würde.

Feature Completion System (FCS)

Nach der FACADE-Theorie werden die Oberflächen im Zuge eines Ausbreitungsprozesses in die Konturen eingefügt. Entsprechend geschlossene Konturen werden als *Filling-In-DOMain (FIDO)* bezeichnet. In einem ersten Schritt muss der Einfluss der Umgebungsbeleuchtung eliminiert werden, da die eingefügten Oberflächenmerkmale ausschließlich die Reflektanzeigenschaften der beleuchteten Objekte wiedergeben sollen (vgl. Helligkeits- und Farbkonstanz); nur auf diese Weise ist eine zuverlässige Objekterkennung zu gewährleisten. Die FACADE-Theorie nimmt an, dass das visuelle System dies erreicht, indem die Farbe oder die Helligkeit einer Oberfläche relativ zur Farbe oder Helligkeit anderer Oberflächen im Gesichtsfeld bestimmt wird. Dieses Verhältnis bleibt unter den meisten Beleuchtungsbedingungen konstant und stellt einen guten Indikator für die tatsächliche Reflektanz von Oberflächen dar. Die Unsicherheit kann weiter reduziert werden, indem das FCS Bereiche in der Nähe von Kanten, die häufig stark schwankenden Beleuchtungsverhältnissen ausgesetzt sind, in Interaktion mit dem BCS systematisch verzerrt. Für die Ermittlung der Oberflächeneigenschaften kommt diesen Bereichen dann eine geringere Bedeutung zu. Der Filling-In-Prozess gleicht einem Diffusionsprozess: Farbe und Helligkeit der Oberfläche verteilen sich gleichmäßig in der FIDO, begrenzt durch die Kanten, die das BCS identifiziert hat.

Ein berühmtes Experiment von Yarbus (1967) veranschaulicht diesen unter normalen Bedingungen nicht wahrnehmbaren Prozess. Wird die Kante einer Figur durch Einsatz technischer Mittel auf einen Bereich der Retina fixiert und werden die durch die Augenbewegungen verursachten Verschiebungen damit eliminiert, geschieht auf der phänomenalen Ebene folgendes: Nach kurzer Zeit verblasst nicht nur die Kante, sondern die Farbe, die die Figur umgibt, scheint in die Figur hineinzufließen. Die Erklärung der FACADE-Theorie für diese Beobachtung ist, dass nach dem Wegfall der Kanten das FCS die Oberfläche über den gesamten Bereich der

Stimulation ausdehnt und über einen Diffusionsprozess ein einheitliches Filling-In dieser neuen Fläche erreicht.

Ein abschließender Vergleich zwischen BCS und FCS verdeutlicht ihre komplementären Funktionen (vgl. Abbildung 91): Die Aufgabe des BCS ist es, eine Repräsentation der Grenzen zu erstellen. Die Kantenbestimmung ist nach „innen“ gerichtet. Die Neurone des BCS sind sensitiv für die Orientierung, aber nicht für die Kontrastpolarität. Der Prozess insgesamt wird von Grossberg als ein Kooperationsprozess beschrieben. Das FCS hat die Aufgabe, eine Repräsentation der Flächen zu erstellen. Der Filling-In-Prozess ist nach „außen“ gerichtet. Die Neurone des BCS sind sensitiv für die Kontrastpolarität, aber nicht für die Orientierung. Beide Systeme interagieren.

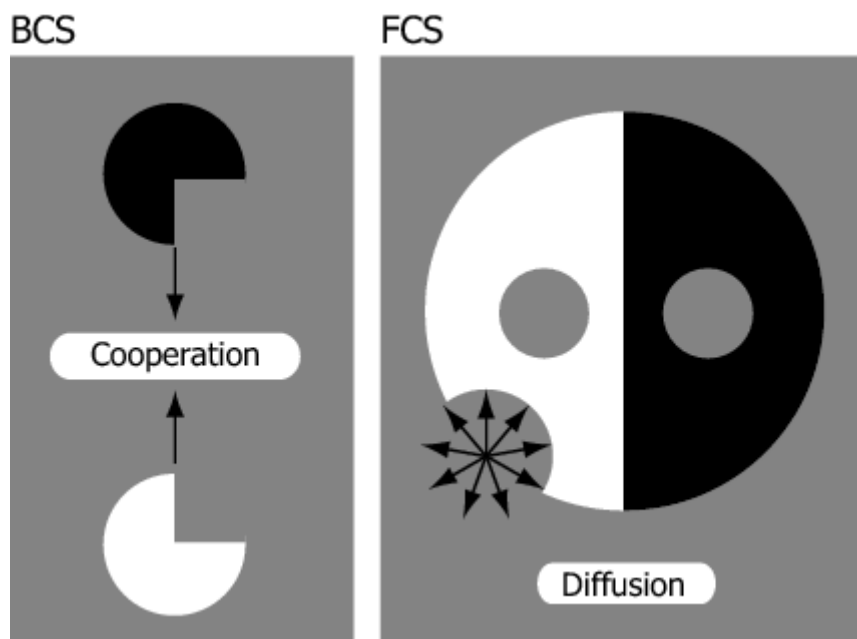


Abbildung 91: Vergleich von BCS und FCS (Bruchmann, 2007; modifizierte Version von Grossberg, 1997). Komplementäre Eigenschaften vom Boundary Contour System (BCS) und Feature Contour System (FCS).

Binokulare Verarbeitung von BCS- und FCS-Informationen

Bisher wurde von der monokularen Analyse einer Szene ausgegangen. Obwohl die Erstellung monokularer FIDO einen wichtigen Zwischenschritt der Objektwahrnehmung entsprechend der FACADE-Theorie darstellt, ergibt sich das Perzept aus einer Anzahl binokularer FIDO, die Informationen beider Augen beinhalten. Abbildung 92 zeigt eine Übersicht der binokularen Interaktionen auf den verschiedenen Analyseebenen und gibt gleichzeitig einen Überblick über die *funktionale Architektur* des FACADE-Modells: Während die Informationsverarbeitung im BCS und im FCS anfangs monokular stattfindet, verarbeiten im BCS bereits die komplexen Zellen des V1 Signale beider Augen. Ein Teil der Neurone ist besonders sensibel für binokulare Disparitäten von Kanten im Bereich ihrer rezeptiven Felder. Es ergibt sich eine *Größen-Disparitäts-Korrelation*: Um die monokularen Informationen über Kanten mit einer großen Disparität – also einer größeren Differenz zwischen deren Positionen auf den beiden Retinae – sinnvoll integrieren zu können, ist ein größeres rezeptives Feld nötig als bei Kanten mit geringeren Disparitäten. Eine solche Korrelation findet sich auch bei den proximalen Stimuli, bei denen nahe Objekte mit großen Disparitäten eine deutlich größere Projektion auf den Retinae erzeugen als weit entfernte Objekte mit geringeren Disparitäten. Zwischen den komplexen Zellen, die unterschiedliche Disparitäten in rezeptiven Feldern gleicher Größe und Position kodieren, findet eine Interaktion statt. Zellen, die größere Disparitäten kodieren, inhibieren diejenigen Zellen, die geringere Disparitäten kodieren. Ein solcher Inhibitionsprozess ergibt insofern Sinn, als dass sich Objekte mit größeren Disparitäten normalerweise näher an der Beobachterin oder am Beobachter befinden und häufig weiter entfernte Objekte verdecken.

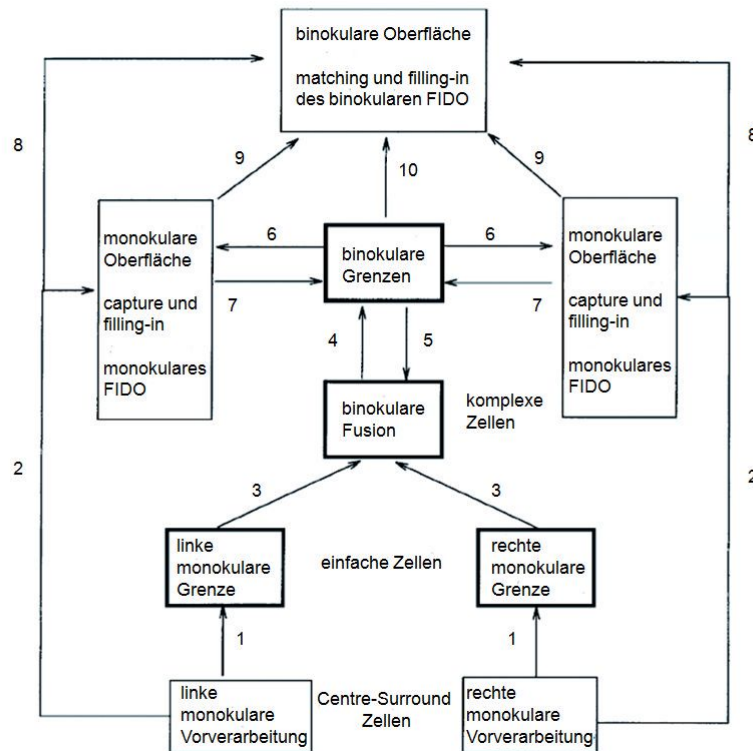


Abbildung 92: Funktionale Architektur des FACADE-Modells (Kelly und Grossberg, 2000). Erläuterung im Text.

Zunächst erfolgt für das linke und das rechte Auge getrennt eine Vorverarbeitung mithilfe von *Center-Surround-Zellen*. Sie sind nicht orientierungssensitiv, stellen aber den Input für die orientierungssensitiven einfachen Zellen zur Verfügung (Pfad 1). Die neuronale Aktivität der einfachen Zellen bildet die Basis für die binokulare Fusion mithilfe komplexer Zellen. Erreicht wird eine binokulare Repräsentation der Grenzen (Pfad 3 und 4). Über Feedbackschleifen findet eine Rückkopplung zu den komplexen Zellen statt (Pfad 5). Ergänzend senden die Center-Surround-Zellen ein Signal zum monokularen FIDO (Pfad 2). Hier bewirken sie einen Diffusionsprozess. Dieser Ausbreitungsprozess geht weiter in den verschiedenen Abteilungen des binokularen BCS (Pfad 6). Die resultierende Oberflächenrepräsentation wird zurückgemeldet, um das BSC-Signal zu unterstützen (Pfad 7). Die übrigen Pfade (Pfade 8, 9 und 10) sind bedeutsam, um eine dreidimensionale Repräsentation mithilfe von Filling-In-Prozessen zu

erklären. Grundlage des Perzeptes ist somit letztendlich eine Reihe von binokularen FIDO (aus Kelly und Grossberg, 2000).

Um die weitere Verarbeitung binokularer Informationen zu erläutern, muss auch die Verarbeitungsebene der hyperkomplexen Zellen um den Faktor der Disparität ergänzt werden: Nur Kanten gleicher oder ähnlicher Disparität bilden – unter Anwendung von CC-Loops – eine einheitliche, geschlossene Kontur. Die Interaktionen (Pfad 4 und 5) verhindern, dass die Konturen weit voneinander entfernter Objekte zu einer gemeinsamen Kontur zusammengefügt werden. Von großer Bedeutung für die Vervollständigung der Konturen sind auch *Null-Disparitäts-Zellen*, die durch Kanten aktiviert werden, die keine Disparität aufweisen. Es handelt sich vor allem um waagerechte Kanten und um Kanten, die aufgrund von Verdeckungen lediglich von einem Auge wahrgenommen werden. Ohne die Signale der Null-Disparitäts-Zellen könnte letztendlich kaum eine Kontur ergänzt werden.

Die vollständigen dreidimensionalen Konturen einer bestimmten Disparität werden als *BCS-Kopie* bezeichnet. In einem Interaktionsprozess (Pfad 6 und 7) mit den monokularen FIDO werden die Helligkeits- und Farbwerte derjenigen FIDO übernommen, deren Konturenstruktur mit denen der zugehörigen BCS-Kopie übereinstimmt. FCS-Signale von FIDO, die nicht mit der BCS-Kopie übereinstimmen, werden unterdrückt. In einem umgekehrt wirkenden Prozess beeinflussen die Kanten monokularer FIDO zudem die Konturen der BCS-Kopie: Kanten gleicher Disparität werden verstärkt, Kanten geringer Disparität an der gleichen Position inhibiert. Die Erkennung verschiedener, sich verdeckender Figuren wird erleichtert.

Die sich aus dieser *kompetitiven Interaktion* ergebenden binokularen Konturen stellen die Basis für die Extraktion binokularer FIDO dar: Informationen aus den monokularen FIDO werden zusammengeführt und durch Filling-In-Prozesse vervollständigt (Pfad 9 und 10). Es wird

angenommen, dass diese abschließende Konstruktion binokularer FIDO im V4 des visuellen Cortex stattfindet. Wie aus den beschriebenen Auswertungsschritten hervorgeht, beschränkt sich die Konstruktion einzelner binokularer FIDO jeweils auf einen bestimmten Disparitätsbereich. Das vollständige Perzept ergibt sich aus einer Anzahl binokularer FIDO, die Objekte in unterschiedlichen Entfernungen repräsentieren. Bereiche, die nicht durch eine Kontur umgrenzt sind, bilden durch unbegrenztes Filling-In den Hintergrund. Kommt es im Laufe des Analyse- und Konstruktionsprozesses bei besonderen Kanten- und Oberflächenkonstellationen zum Filling-In einer Figur niedrigerer Disparität hinter einer Figur höherer Disparität, wird die Oberfläche der vorderen Figur als transparent wahrgenommen.

Die FACADE-Theorie ist ein Beispiel für die komplexe Theoriebildung in der heutigen Wahrnehmungsforschung. Empirisch prüfen lassen sich lediglich einzelne Aussagen, nicht die Theorie als Ganze. Welche Vorhersagen der Theorie gegenwärtig testbar sind, hängt von den Forschungsmethoden ab. Theorie und Methode sind insofern nicht unabhängig voneinander. Gerade im Hinblick auf die Methoden hat sich in den letzten Jahrzehnten viel verändert.

Traditionell arbeitet die Wahrnehmungsforschung mit einer sorgfältigen Beschreibung der Phänomene durch die Probandin oder den Probanden im Rahmen psychophysischer Experimente. Unter kontrollierten Laborbedingungen werden die Merkmale des Reizes oder die Reizdarbietung systematisch variiert. Die zu prüfende Theorie legt bedeutsame Variationen fest. Hinzu gekommen sind Computersimulationen zu den Prozessen der visuellen Informationsverarbeitung, Methoden der funktionellen Neuro-Bildgebung sowie Einzelzelleableitungen in Tierexperimenten. Ein zentrales Problem besteht darin, die Forschungsergebnisse dieser vier sehr unterschiedlichen Stränge zusammenzuführen.

Die zu Beginn des Buches vorgestellten und an vielen Beispielen erläuterten Prinzipien der visuellen Informationsverarbeitung könnten Bausteine einer universellen Theorie des Sehens sein. Die Heuristiken, mit denen diese Ziele erreicht werden, sind in manchen Teilbereichen der Forschung gut untersucht (vgl. Kap. IV Gliederungs- und Gruppierungsprozesse). In anderen Bereichen steht die Forschung erst am Anfang (z. B. beim Zusammenspiel der visuellen Sensorik mit der Motorik). Eine universelle Theorie wird vor allem prüfbare Annahmen zum Verhältnis von daten- und hypothesengesteuerter Verarbeitung formulieren müssen. Ein Leitgedanke könnte sein, dass die hypothesengesteuerte Verarbeitung immer dann an Bedeutung gewinnt, wenn die datengesteuerte Verarbeitung über die Gliederungs- und Gruppierungsmechanismen keine Eindeutigkeit herstellen kann.

Zusammenfassend lässt sich der Prozess des Sehens so beschreiben: Ausgangspunkt sind die mehrdeutigen und sich ständig verändernden Stimulationen auf den Retinae der beiden Augen. Innerhalb von Millisekunden werden diese Ambiguitäten reduziert und ein eindeutiges und stabiles Perzept erstellt. Dazu analysiert das Gehirn die Stimulation. Es akzentuiert die Unterschiede, gruppiert die Inhomogenitäten und ergänzt die Lücken. Es berücksichtigt die Wahrscheinlichkeiten, mit denen Ereignisse in der Außenwelt auftreten. Die visuelle Welt ist keine Kopie der Außenwelt, sondern ein Modell, das in fast allen Situationen schnelle und zielgenaue Reaktionen erlaubt. Es ist die Grundlage für unsere Handlungen, Entscheidungen und Vorhersagen. Da das Modell multiperspektiv ist, ist auch eine zuverlässige Identifikation und Lokalisation von Objekten möglich, die uns aktuell in einer anderen Perspektive begegnen. Kann aufgrund der Stimulation und der nachfolgenden Inferenzprozesse kein eindeutiges Perzept erstellt werden, hilft die Zirkularität des Sehvorgangs. Durch Augenbewegungen, eine Veränderung des Standortes oder die Einbeziehung anderer Sinnesmodalitäten werden zusätzliche Informationen gewonnen. Eine

scharfe Trennung zwischen Wahrnehmen, Erinnern und Denken ist Fiktion. Letztlich sind alle kognitiven Prozesse am Aufbau unserer visuellen Welt beteiligt.

Kapitel VIII: Ästhetische Wahrnehmung

Ästhetische Wahrnehmungen sind ein wichtiger Bestandteil der visuellen Welt, aber selten werden die zugehörigen Forschungsfragen aus einer naturwissenschaftlichen Perspektive reflektiert. Auch lassen sie sich keinem speziellen Bereich wie der Farb- oder Objektwahrnehmung zuordnen. Vielmehr können alle Reize hinsichtlich ihrer ästhetischen Eigenschaften beurteilt werden. Die Suche nach Schönheit lässt sich durch alle Kulturen und Epochen verfolgen. Umberto Eco hat mit den von ihm herausgegeben Bänden „Die Geschichte der Schönheit“ und „Die Geschichte der Häßlichkeit“ gezeigt, wie sich die Vorstellungen von dem, was als „schön“ beurteilt wird, immer wieder verändert haben. Dieses abschließende Kapitel geht nicht primär der Frage nach, was als schön empfunden wird. Der Blick richtet sich nicht auf die Wandlungen der Schönheitsvorstellungen. Im Zentrum der Aufmerksamkeit steht vielmehr die Frage: Warum wird etwas als schön wahrgenommen?

Es ist die Frage nach den grundlegenden Gesetzmäßigkeiten des ästhetischen Erlebens. Die Wahrnehmenden, die Subjekte der ästhetischen Erfahrung, stehen im Fokus. Das zugehörige Forschungsprogramm ist die *Experimentelle Ästhetik*. Sie lässt sich als der Teil der empirischen Ästhetik auffassen, der theoriegeleitet Kausalhypothesen mithilfe von Experimenten untersucht. Ästhetische Wahrnehmungen sind nicht auf die visuelle Modalität beschränkt, selbst wenn viele Untersuchungen aus diesem Bereich stammen. Seit den grundlegenden Arbeiten von Gustav Theodor Fechner sind etwa 150 Jahre vergangen (zur viel längeren Geschichte der „psychologischen Ästhetik“ vgl. umfassend Allesch, 1987; einführend Allesch, 2006). In dieser Zeit hat die Experimentelle Ästhetik keine kontinuierliche Entwicklung genommen. Vor allem in den letzten 30 Jahren jedoch findet sich ein deutlich gestiegenes Forschungsinteresse (vgl. Kebeck und Schroll, 2011). Dies hängt unmittelbar mit der

Entwicklung der Neurowissenschaften und ihrer Methoden wie EEG und MRT zusammen.

Fechner nahm an, dass durch eine systematische Variation von Merkmalen der Reize die Veränderung der zugehörigen ästhetischen Empfindungsstärke zuverlässig gemessen werden kann. Das Ziel war die allgemeinen „Principe des aesthetischen Gefallens“ zu bestimmen. Bereits 1865 hatte Fechner auf diese Weise Untersuchungen zum goldenen Schnitt durchgeführt. Seine Studie „Zur experimentalen Aesthetik“ erschien 1871 und eine systematische Darstellung des Ansatzes, einschließlich seiner methodologischen Vorstellungen, findet sich in der „Vorschule der Aesthetik“ (1876). Im Zentrum von Fechners Überlegungen steht das unmittelbare und spontane ästhetische Urteil als Ausdruck einer ästhetischen Empfindung.

Die Suche nach einem geeigneten „Mass für die Wohlgefälligkeit“ führte Fechner zu folgender Überlegung: „Direct kann man den Grad der Wohlgefälligkeit oder Missfälligkeit, unter dem ein Verhältnis diesem oder jenem unter diesen oder jenen Umständen erscheint, nicht messen; dazu gälte es ein Mass der Einzel-Lust und Unlust zu haben, was wir noch nicht haben; aber man kann das Mass andershin übertragen, indem man die Personen zählt, die dem einen und die dem anderen Verhältnisse bei gleichzeitiger Vorlage derselben oder überhaupt gleicher Möglichkeit der Wahl den Vorzug geben“ (Fechner 1871, S. 598). Dieses Vorgehen bezeichnet er als „extensives Mass der Wohlgefälligkeit“, das von dem erst noch zu findenden eigentlichen Ziel des „intensiven Masses der Wohlgefälligkeit“ abzugrenzen ist. Das Bild der Experimentellen Ästhetik, das Fechner entwarf und das bestimmend wurde für die weitere Entwicklung dieses Forschungsprogramms, geht von folgenden Annahmen aus: In der Ästhetik gibt es elementare Gesetzmäßigkeiten (z. B. Gliederungs- und Gruppierungsverhältnisse), die gleichzeitig universell sind. Diese beziehen sich auf den *direkten Eindruck* und lassen

sich unabhängig von dem assoziativen Gehalt der Objekte über subjektive Urteile im Rahmen naturwissenschaftlicher Methoden erheben.

In der Nachfolge von Fechner dominierte lange Zeit die Überzeugung, dass es, ausgehend von den Objekteigenschaften, möglich sein sollte, die ästhetische Wirkung von Gegenständen losgelöst von kulturellen Einflüssen oder individuellen Charakteristika vorherzusagen, eine *Objektästhetik* zu formulieren. Das Ziel war die Entwicklung einer Ästhetikformel, die beschreibt, wie die Stärke einer ästhetischen Empfindung mit den Eigenschaften des Gegenstandes variiert. Zwei Größen haben in unterschiedlichen Ansätzen eine zentrale Rolle gespielt: Ordnung und Komplexität.

1. Ordnung und Komplexität

Der Mathematiker George David Birkhoff geht davon aus, dass das ästhetische Maß (M) eines beliebigen Gegenstandes mit seinem Ordnungsgrad (O) zunimmt und mit seiner Komplexität (C) abnimmt ($M = O / C$). O wird additiv bestimmt über Maßzahlen für Spiegel- und Radialsymmetrien, Gleichgewicht und Orthogonalität, C über die notwendige Zahl von Geraden für ein Gitter, in dem die Figur vollständig enthalten ist. Die Anwendung dieser Formel erläutert Birkhoff an Polygonen (Birkhoff, 1933). Die Abbildung 93 zeigt vier Beispiele.

Gemäß dieses Ansatzes kommt es immer dann zu einem vergleichbaren M , wenn Ordnungsgrad und Komplexität in gleichem Verhältnis zueinander stehen. Abgeleitet wird die Bedeutung der beiden Größen O und C aus den Phasen einer ästhetischen Wahrnehmung. Danach erzeugt ein ästhetisches Objekt bei der Betrachterin oder beim Betrachter ein intuitives und bewertendes Gefühl, wobei sich folgende Schritte abgrenzen lassen: Zunächst werden die notwendigen Aufmerksamkeitsressourcen bereitgestellt, die bestimmt sind durch die Komplexität C eines

Gegenstandes. Als „Belohnung“ für den Aufmerksamkeits- und Wahrnehmungsaufwand resultiert eine Empfindung, die dem ästhetischen Wert M entspricht. Der dritte Teilprozess lässt sich als Erkennen der ordnungsstiftenden Merkmale O beschreiben, wozu Symmetrie, Wiederholung, Kontrast, Gleichheit und Ähnlichkeit gehören. Birkhoff hatte diesen Ansatz zunächst 1928 auf einem Mathematikerkongress in Bologna vorgestellt, bevor er 1933 sein berühmt gewordenes Buch „Aesthetic measure“ publizierte.

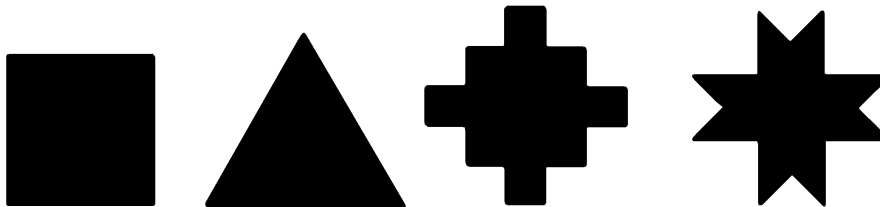


Abbildung 93: Beispiele für Birkhoff-Polygone. Für das ästhetische Maß $M = O / C$ ergeben sich für die vier Polygone von links nach rechts folgende Werte: 1,5; 1,2; 0,5; 0,5 (Kobbert, 1986).

„Man kann feststellen, dass die ästhetische Erfahrung drei sukzessive Momente enthält: 1. Eine vorgängige Anstrengung, die notwendig ist, um das Objekt richtig zu erfassen, und die die Proportionale der Komplexität C des Objektes ist; 2. Das Gefühl des Vergnügens oder ästhetischen Maßes, das diese vorgängige Anstrengung belohnt; 3. Dann die bewusste Wahrnehmung, dass sich das Objekt einer gewissen Harmonie oder Symmetrie oder Ordnung erfreut, die mehr oder weniger verborgen ist und eine notwendige, wenn nicht ausreichende Bedingung für die ästhetische Erfahrung selbst zu sein scheint.“ (zitiert nach Klütsch, 2007, S. 51)

Während die Formel von Birkhoff später häufig Gegenstand empirischer Prüfung war, wurde seine Theorie kognitiver Prozesse kaum rezipiert. Als kunstgeschichtliches Beispiel dienten Birkhoff chinesische Vasen

(Abbildung 94). Das von Birkhoff (1928) verwendete Gitter erleichtert die Beschreibung der Symmetrien. Er kommt zu dem Schluss: „Die Vasen, die den besten ästhetischen Eindruck geben, sind genau diejenigen, für die das Maß M das höchste ist“. Schließlich geht er noch einen Schritt weiter und schlägt vor, eine Vase zu konstruieren, die ein Maß $M = 1$ hat: „[...] besitzt zum Beispiel ein Maß $M = 1$ und ist – zumindest aufgrund unserer Formel – hinsichtlich ihrer Form schöner als die chinesischen Vasen [...] es wäre interessant, eine Vase mit dieser Form zu konstruieren“.

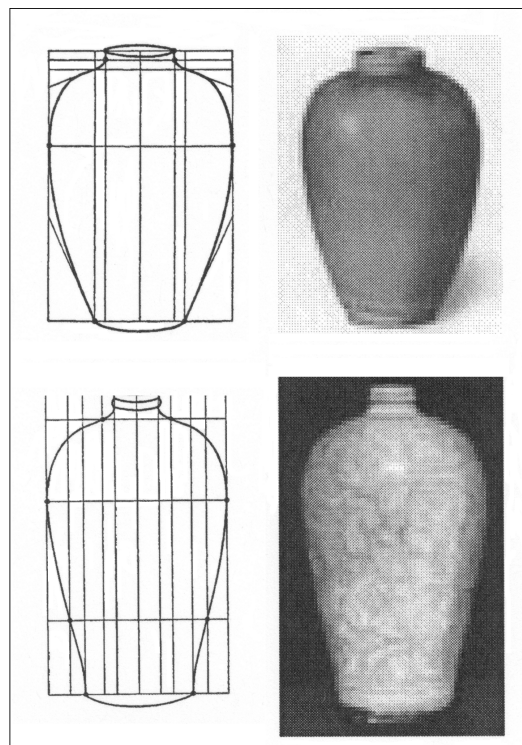


Abbildung 94: Chinesische Vasen und Umrisszeichnungen mit Gitter. Fotografien von Hobson, 1927, Zeichnungen von Birkhoff, 1928 (Klütsch, 2007).

Birkhoff selbst jedoch hat weder Konstruktionsregeln hierfür angegeben noch ein Beispiel erstellt. Hans Jürgen Eysenck postuliert aufgrund eigener empirischer Untersuchungen eine Produktformel, die enger mit den Vorstellungen Fechners und v. Ehrenfels verwandt ist (Eysenck, 1942).

Demnach wächst das ästhetische Maß sowohl proportional zum Ordnungsgrad als auch zur Komplexität ($M = O \times C$).

Von ähnlichen Überlegungen wie Birkhoff ausgehend hat sich in den 1950er und 1960er Jahren die *Informationsästhetik* entwickelt. Ihr Kerngedanke ist die Quantifizierung des Informationsgehaltes einer Nachricht. Eine Nachricht kann jeder beliebige Stimulus sein, auch ein ästhetischer Gegenstand (ein Bild, ein Musikstück, ein Gedicht). Die grundlegende Maßeinheit für Information („bit“) drückt die Anzahl der Ja/Nein-Entscheidungen aus, die erforderlich sind, um eine beliebige Reizkonstellation eindeutig zu beschreiben. Eine Wiederholung führt zu *Redundanz*. Sie gibt an, wie viel Information im Mittel pro Zeicheneinheit in einer Informationsquelle mehrfach vorhanden ist. Eine Informationseinheit ist demnach dann redundant, wenn sie ohne Informationsverlust weggelassen werden kann. Redundanz wird in technischen Systemen vor allem dazu genutzt, um Defekte bei der Nachrichtenübertragung ausgleichen zu können. Sie steht in einem engen Zusammenhang zur Fehlertoleranz. Sie erlaubt es dem Empfänger einer Information, verlorengegangene oder verfälschte Teilinformationen aus ihrem Kontext zu rekonstruieren. In Bezug auf die Ästhetik lautet die Frage: Wie viel Redundanz ist optimal? In diesem Sinne entwickelt Gunzenhäuser (1975) die Formel von Birkhoff (1933) weiter. Auch Gunzenhäuser nimmt an, dass die Komplexität (C) eines wahrgenommenen Gegenstandes den Nenner bildet. Im Zähler jedoch ersetzt er die Ordnung (O) durch die Differenz aus der Anfangskomplexität und der Restkomplexität, die verbleibt, nachdem die Ordnung im Gegenstand erkannt wurde. Die Formel lautet dann:

$$M = \frac{C_{Anfang} - C_{Rest}}{C_{Anfang}}$$

Die Differenz im Zähler nennt Gunzenhäuser (1975) *subjektive Redundanz*. Damit soll zum einen zum Ausdruck gebracht werden, dass es, informationstheoretisch gesehen, um eine Informationsentnahme und

damit um eine Informationsminderung geht. Die Redundanz entsteht jedoch erst durch das erkennende Subjekt, das die Anfangskomplexität auf die Restkomplexität reduziert. Die Restkomplexität wird umso niedriger, je geordneter der Gegenstand ist. Damit wird der ästhetische Gesamtwert (M) umso größer, je mehr Ordnung sich vom erkennenden Subjekt aus der Stimulation „herauslesen“ lässt. Die subjektive Redundanz ermöglicht Vorhersagen, warum Hochkomplexes (z. B. ein Ölbild von Velazquez) und deutlich weniger Komplexes (z. B. eine Konturzeichnung von Matisse) ein vergleichbares Ausmaß an ästhetischem Vergnügen erzeugen können.

Ein aktuelles Beispiel für die Anwendung von Ästhetikformeln ist die Programmierung graphischer Interfaces, z. B. das Layout von Websites oder E-Books. Nach Möglichkeit sollen sie so gestaltet sein, dass sie weltweit als ästhetisch wahrgenommen werden. Basierend auf den Gestaltgesetzen und der Formel von Birkhoff (1933) bemühen sich eine Reihe von Ansätzen, Algorithmen zu entwickeln, die einen Mindeststandard für die ästhetische Gestaltung etablieren bzw. die ästhetische Gestaltung automatisieren sollen (vgl. Ngo, Samsudin und Abdullah, 2000; Ngo und Teo, 2000; Ngo, Teo und Byrne, 2002; Ngo, Teo und Byrne, 2003 sowie Hsiao und Chou, 2006). Berücksichtigt werden solche Aspekte, die sich leicht formalisieren lassen, wie Balance, Gleichgewicht, Symmetrie, Abfolge bzw. Rhythmus, Einheitlichkeit, Einfachheit, Dichte, Komplexität oder Ordnung. Hier ein Beispiel: „Balance“ wird definiert über das Verhältnis des „optischen Gewichts“ der verschiedenen Bildelemente. So kann das Ausmaß an Balance (BM) auf einer Webseite anhand folgender Formeln bestimmt werden (Ngo et al., 2000).

$$BM = 1 - \frac{|BM_{vertical}| + |BM_{horizontal}|}{2} \in [0,1]$$

$$BM_{vertical} = \frac{w_L - w_R}{\max(|w_L|, |w_R|)}$$

$$BM_{horizontal} = \frac{w_T - w_B}{\max(|w_T|, |w_B|)}$$

$$w_j = \sum_i^{n_j} a_{ij} d_{ij} \quad j = L, R, T, B$$

Wobei $BM_{vertical}$ und $BM_{horizontal}$ für die vertikalen und horizontalen Balancen stehen, die Indizes L und R für die Bereiche links und rechts der vertikalen Achse und die Indizes T und B für die Bereiche oberhalb und unterhalb der horizontalen Achse; w_j ist das aufsummierte Gewicht des Bereichs j ; a_{ij} ist die Fläche des Objektes i im Bereich j , d_{ij} der Abstand zwischen den Mittellinien des Objektes und dem Rahmen und n_j ist die Anzahl der Objekte im Bereich j . Ist eine Seite perfekt ausbalanciert, so ist $BM = 1$.

2. Gestaltgüte

Bereits in Zusammenhang mit den Gliederungs- und Gruppierungsheuristiken (Kapitel IV) wurde darauf hingewiesen, dass „gute Gestalten“ mit einem positiven Werterleben verbunden sind. Christian von Ehrenfels (1890) definiert das Merkmal der *Gestalthöhe* oder *Gestaltgüte*, das er in unmittelbarem Zusammenhang zu jeder ästhetischen Wahrnehmung sieht, als „Produkt von Einheit und Mannigfaltigkeit“. Aus Sicht der Wahrnehmungsforschung sind Werke der bildenden Kunst zunächst „Gestalten“. Da sie übersummativ und transponierbar sind, können sie Gegenstand der Imagination und der Fantasie werden. Wir können sie aus ihrem ursprünglichen Kontext lösen und sie mit anderen

Gestalten vergleichen. Auf diese Weise werden sie auch Gegenstand der Reflexion.

Wolfgang Metzger hat in einer wenig beachteten Schrift „Der Beitrag der Gestalttheorie zur Frage der Grundlagen des künstlerischen Erlebens“ (1965) Ansätze einer Theorie ästhetischen Erlebens auf der Basis von Gestaltqualitäten vorgestellt. Er geht von Überlegungen Karl Dunckers in dessen posthum veröffentlichter Schrift „Pleasure, Emotion, and Striving“ (1941) aus. Duncker unterscheidet drei Arten lustvollen Erlebens: 1. den sinnlichen Genuss (etwa wenn jemand einen guten Wein trinkt), 2. die Formen der Freude, die aus Gesinnungen oder Wertvorstellungen hervorgehen (etwa wenn eine „gute Sache“ gesiegt hat) und 3. den Kunstgenuss oder die ästhetische Freude.

Demnach können Empfindungen oder Sinnesqualitäten als solche Genuss bereiten (Fall 1). Sie können zweitens, wenn sie in einer geeigneten Anordnung zu Wahrnehmungsgebilden zusammengefasst werden (z. B. einem gesprochenen oder geschriebenen Wort), als Mittel dienen, etwas mitzuteilen. Der Inhalt dieser Mitteilung kann Anlass zur Freude sein (etwa die Nachricht über die Geburt eines Kindes). Drittens können Sinnesqualitäten oder Empfindungen etwas unmittelbar zum Ausdruck bringen. So kann das Streicheln einer Hand auch zärtliche Besorgnis bedeuten und ein Rondo von Mozart schwebende Heiterkeit. Ästhetischer Genuss oder Kunstgenuss ist nach Duncker in jedem Fall Freude an etwas Ausgedrücktem. Charakteristisch ist, dass sie unmittelbar am Vorgang des Ausgedrücktwerdens erlebt wird. Duncker setzt voraus, dass dieser Vorgang vom Gegenstand seinen Ausgangspunkt nimmt und nicht vom Willen des Subjektes. Gelingt diese Herausgrenzung aus der „übrigen Welt der Sorgen, Hoffnungen und Befürchtungen“, dann ist ästhetische Freude auch einem Gegenstand gegenüber möglich, der sonst nur Furcht und Schrecken erwecken würde.

Damit Sinnesqualitäten oder Empfindungen ausdrucksstark sein können, müssen sie zu Wahrnehmungsgebilden organisiert sein, die Gliederungen und Gruppierungen aufweisen. Während aber für die Funktion der Mitteilung (Fall 2) nur eine eindeutige Zuordnung zwischen Wahrnehmungsgebilde und Bedeutungsgehalt notwendig ist, die im Übrigen völlig willkürlich und rein konventionell sein kann, verlangt die Ausdrucksfunktion (Fall 3) eine Nicht-Beliebigkeit der Zuordnung, eine Identität in wesentlichen Zügen der Struktur, die von den Gestalttheoretikern als *Isomorphie* bezeichnet wird (Köhler, 1920). Diese ist verantwortlich, wenn von unterschiedlichen Stimuli die gleichen Gefühle hervorgerufen werden. Nach gestalttheoretischer Auffassung ist ein Gefühl nur ein Spezialfall einer sehr viel allgemeineren Kategorie von Qualitäten, die als *Gestaltqualitäten* bezeichnet werden. Sie sind unlösbar an die Struktur des Gebildes gebunden (zumeist wird das Bild eines Gefäßes, in dem sich eine flüssige Substanz befindet, die nicht ausgeschüttet werden kann, verwendet). Entscheidend ist das besondere Verhältnis von Qualitäten und Strukturen. Es ist nicht umkehrbar. Es gibt weniger Qualitäten als Strukturen. Jede Gestaltqualität hat eine ganze Familie von Strukturen, in denen sie mehr oder weniger ausgeprägt erscheint. „Das bedeutet: Um eine Struktur, aus der sie in völliger Reinheit und Transparenz hervorstrahlt, und die wir daher die prägnante nennen, scharen sich strukturelle Varianten von abnehmender Prägnanz, in denen die fragliche Gestaltqualität immer weniger rein, immer stärker gestört, schließlich nur noch angedeutet und endlich überhaupt nicht mehr erscheint.“ (Metzger 1965, S. 503).

Die Gestaltqualitäten sind weitgehend unabhängig von dem Material, aus dem die sie tragenden Strukturen gebildet sind. Daher kann es auch zu einer engen Verwandtschaft zwischen künstlerischen Äußerungen verschiedener Epochen oder in verschiedenen Gattungen und Medien kommen. Diese Universalitätsannahme rechtfertigt letztlich eine allgemeine Theorie ästhetischen Erlebens (Köhler, 1938). Wenn

menschliche Gefühle Sonderfälle von Gestaltqualitäten sind und jede Gestaltqualität ein strukturelles Korrelat hat, so müssen auch die Gefühle strukturelle Korrelate haben. Nach Auffassung von Arnheim (1980) sind es die *gerichteten Spannungen* oder *dynamischen Strukturen*, welche die Grundlage der ästhetischen Gefühle bilden.

Rudolf Arnheim hat sich intensiv mit der Bedeutung dynamischer Gestaltqualitäten beschäftigt (Arnheim, 1978, 1980). Jede Figur und jedes Objekt weist anschaulich eine gerichtete Bewegung oder eine Spannung auf. Er beschreibt diese Wahrnehmungskräfte, die in Abbildung 95 gezeigt werden, so: „In einem Quadrat oder Rechteck strebt die Bewegung auf die Seiten zu (b), aber es gibt auch eine Bewegung zu den Ecken hin (c). Ein Stern taucht auf, in-dem seine Zacken nach außen streben (d). Wenn ein gleichseitiges Dreieck auf einer seiner Seiten steht, bleibt die Grundlinie ruhig, während die zwei anderen Seiten kräftig nach außen und oben drücken, wobei die Spitze als Scharnier dient (e). Wird das Bild nur sehr kurz gezeigt, entsteht die gleiche Figur dadurch, dass die Spitze von der Grundlinie heftig nach oben drängt (f). Wenn das Quadrat oder Dreieck auf einer der Ecken steht, drücken alle Ecken mehr oder weniger systematisch nach außen (g, h). Die Bewegung zeigt jedoch eine Tendenz, in den horizontalen Richtungen am stärksten zu sein, und in der Vertikalen stößt sie eher nach oben als nach unten. Das wird durch das Quadrat veranschaulicht (b). Die seitliche Bewegung ist am stärksten ausgeprägt, die nach oben ist schwächer, und die nach unten fehlt fast ganz.“ (Arnheim, 1978, S. 441).

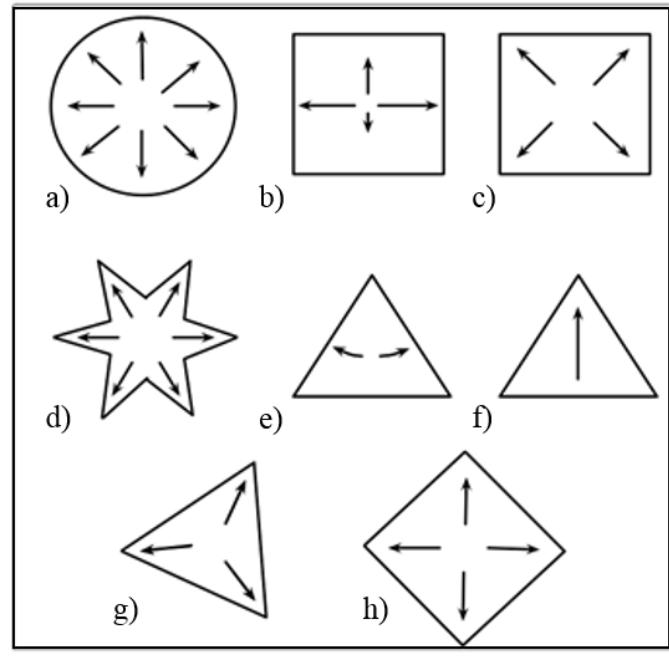


Abbildung 95: Dynamik einfacher geometrischer Formen (Arnheim, 1978). Erläuterung im Text.

Es handelt sich bei diesen Kräften um eine phänomenale, nicht um eine physikalische Eigenschaft. Die Dynamik lässt sich als Vektor darstellen. Wie Arnheim sich die Analyse solcher Spannungsverhältnisse vorstellt, zeigt er nicht nur an einfachen geometrischen Figuren, sondern auch an Kunstwerken (u. a. an der „Beweinung Christi“ von Giotto und der „Auferstehung“ von Piero della Francesca). Jedes Kunstwerk weist ihm eigene Spannungsverhältnisse auf, die auf die Betrachterin oder den Betrachter wirken. Metzger (1968) schlägt folgende Systematik der Gestaltqualitäten vor, die auf Werke der bildenden Kunst anwendbar ist (vgl. auch Kobbert, 1986): (a) Wesenseigenschaften (Anmutungs- und Ausdruckseigenschaften), (b) Struktur- und Gefügequalitäten (Tektonik und Dynamik), (c) Ganzqualitäten (Farb- und Materialqualitäten).

(a) Anmutungs- und Ausdrucksqualitäten

Bei *Anmutungsqualitäten* steht der Eindruck eines Kunstwerks noch vor seiner intellektuellen Analyse im Mittelpunkt. So wirkt ein Kunstwerk auf

die Betrachterin oder den Betrachter anziehend oder abstoßend, erregend oder beruhigend. Es bereitet Wohlgefallen oder Missfallen. Diese Anmutungsweise hat zugleich auch einen mehr oder weniger starken Aufforderungscharakter, mit dem die erste Betroffenheit in eine Explorationstätigkeit oder eine geistige Auseinandersetzung übergeht. Die Anmutungsqualitäten werden ihrerseits durch *Ausdrucksqualitäten* hervorgerufen und diese wiederum basieren auf tektonischen und dynamischen Qualitäten. So kann ein Gewitter „gewaltig“ und eine Landschaft „heiter“ sein. Zum Bild der Künstlerin oder des Künstlers gehört, dass sie oder er eine besondere Sensibilität für die Entdeckung von Ausdrucksqualitäten hat und in der Lage ist, diese über dynamische und tektonische Qualitäten im Werk zur Geltung zu bringen. Die „Lebendigkeit“ in einem Kunstwerk bezieht sich demnach nicht auf das „was“ der Darstellung, sondern auf das „wie“. Ein Bild von einem menschlichen Gesicht ist nicht deshalb lebendig, weil es etwas Lebendiges zeigt, sondern weil es selbst lebendig wirkt. So können auch ein abstraktes Gemälde, ein Gedicht oder ein Musikstück lebendig erscheinen und Ausdrucksqualitäten haben wie „ernst“ und „heiter“. Je nach den Voraussetzungen und der Bereitschaft bei der Betrachterin oder beim Betrachter können dieselben Ausdrucksqualitäten unterschiedliche Anmutungsqualitäten hervorrufen.

(b) Tektonik und Dynamik

Tektonik bezieht sich auf die räumlichen Beziehungsgefüge der Teile zum Ganzen und steht häufig im Mittelpunkt des Interesses um Gestaltqualitäten. Tektonische Qualitäten werden im Bereich der bildenden Kunst gerne an abstrakten Arbeiten veranschaulicht, weil sich hier leicht systematische und plausible Variationen des Originals erzeugen lassen. Sie korrespondieren mit Merkmalen der Komposition. Wie bereits erwähnt hat Arnheim (1978, 1980) sich intensiv mit der Bedeutung dynamischer Gestaltqualitäten beschäftigt. Jede Figur und jedes Objekt

weist anschaulich eine gerichtete Bewegung oder eine Spannung auf, die wir als *Dynamik der Form* wahrnehmen.

(c) Farb- und Materialqualitäten

Farbe ist bei Objekten oft jenes Merkmal, das die Anmutungsqualitäten dominiert. Ein Betrachterin oder ein Betrachter kann etwa 200 Farbtöne unterscheiden. Wird die Helligkeit einer Farbe durch langsame Zunahme der Intensität variiert, so können zusätzlich etwa 500 Helligkeitsabstufungen pro Farbton diskriminiert werden. Schließlich ist es möglich, bei einem Farbton und einer Helligkeitsstufe ca. 20 Sättigungsgrade zu unterscheiden. Insgesamt werden somit etwa 2 Millionen Farbabstufungen erkannt (Gouras, 1991). Von dieser enormen Diskriminationsfähigkeit ist die sprachliche Differenzierungsmöglichkeit abzugrenzen. Das „Munsell Book of Colors“ enthält 1225 Farbmuster, die gleichabständig sind und das „Pantone Farbsystem“ kennt 1200 Farbbezeichnungen. Interessanter als die Frage nach der maximalen Anzahl ist, wissenschaftlich gesehen, die Frage nach dem Minimum: Wie viele Farbnamen werden benötigt, um alle Farbkombinationen zu beschreiben? Dies gelingt bereits mit vier chromatischen Farben: Rot, Gelb, Grün und Blau. Andere Farben wie etwa Braun oder Orange sind nicht notwendig. Allerdings darf auch keine der vier Grundfarben fehlen (vgl. Hurvich, 1981; Abramov und Gordon, 1994). Ethnologische Studien zeigen, dass das chronologische Muster, mit dem Farbnamen neu in eine Sprache aufgenommen werden, fast immer identisch ist. So kennen alle Kulturen Bezeichnungen für Schwarz und Weiß. Kommt ein Name für eine chromatische Farbe hinzu, dann ist es in der Regel zunächst Rot, gefolgt von Grün, Gelb und Blau. Farben wie Orange, aber auch Grau, treten erst später auf. Insgesamt finden sich übereinstimmend 7-11 Farbkategorien (Hardin und Maffi, 1997).

Mit Materialqualitäten sind phänomenale Eigenschaften wie hart vs. weich, rau vs. glatt, kalt vs. warm oder leicht vs. schwer gemeint. Materialqualitäten sind im Gegensatz zu den Farbqualitäten die unscheinbarsten Gestaltqualitäten. Auch wenn sie oft nicht optischer Natur sind, können sie doch über die visuelle Textur des Materials interpretiert werden. So kann die Beobachterin oder der Beobachter den meisten Oberflächen ansehen, ob sie „rau“ oder „glatt“ sind. Für ein Verständnis der Wirkung von Kunstwerken ist gerade das Wechselspiel der Materialqualitäten mit den Anmutungs- und Ausdrucksqualitäten bedeutsam. Es ist zu vermuten, dass ein Teil der Unterschiede in den ästhetischen Urteilen zwischen Lai*innen und Expert*innen aus einer unterschiedlichen Gewichtung einzelner Gestaltqualitäten resultiert.

3. Schönheit als Produkt der Informationsverarbeitung

Aus heutiger Sicht ist der Objektästhetik eine *Subjektästhetik* an die Seite zu stellen. Die Wahrnehmung von Schönheit ist nicht nur an Merkmale der Stimulation gebunden, sondern auch ein Ergebnis der menschlichen Informationsverarbeitung. Kognitions- und neurowissenschaftliche Modelle bemühen sich um kausale Erklärungen, wie es im Laufe des Verarbeitungsprozesses zu ästhetischen Empfindungen und zu ästhetischen Urteilen kommt. Bei aller Heterogenität ist ihnen eines gemeinsam: Sie sehen die Grundlagen des ästhetischen Erlebens und Verhaltens in neuronalen Informationsverarbeitungsprozessen des Gehirns und werden manchmal auch als „Neuroästhetik“ bezeichnet (vgl. u. a. Zeki, 1999, 2001; Skov und Vartanian, 2009).

Drei Forschungsperspektiven lassen sich bei der Untersuchung neuronaler Korrelate ästhetischer Wahrnehmungen unterscheiden: Areal, Verarbeitungsprinzip und Neurotransmitter.

(a) Suche nach dem Areal

In der Folge methodischer Fortschritte bei der Nutzung bildgebender Verfahren (vor allem fMRT) hat die Suche nach spezifischen Hirnregionen als neuronale Substrate des ästhetischen Erlebens zunehmende Bedeutung erlangt (vgl. z. B. Zeki, 1999). Wichtige Fragen, die durch diese Perspektive beantwortet werden könnten, sind beispielsweise: 1. Existieren kortikale Regionen, die ausschließlich oder vorwiegend für die ästhetische Informationsverarbeitung zuständig sind? 2. Gibt es eine Hirnregion, die bei jeder Art von ästhetischem Erleben unabhängig von der Sinnesmodalität aktiv ist und die somit als „Kern“ ästhetischer Empfindungen betrachtet werden kann? 3. Werden je nach Art der ästhetischen Informationsverarbeitung verschiedene Regionen aktiviert – beispielsweise in Abhängigkeit vom Ausmaß der emotionalen Beteiligung? 4. Sind positive und negative Bewertungen durch unterschiedlich starke Aktivierung derselben Region oder durch Aktivierung unterschiedlicher Regionen repräsentiert?

(b) Suche nach dem Verarbeitungsprinzip

Der zweite Ansatz geht nicht von einem definierten räumlichen Aktivitätsmuster, sondern von einer speziellen zeitlichen Struktur bei der Kodierung neuronaler Aktivität als Korrelat ästhetischer Wahrnehmung aus. Ein zeitliches Kodierungsprinzip könnte in einer Vielzahl weitgehend unabhängig und modalitätsspezifisch arbeitender Hirnregionen zum Einsatz kommen, die auch grundlegend für nicht-ästhetische Analysen der Stimuli sind. Konkret würde dies beispielsweise bedeuten, dass ein als schön empfundenes Musikstück genau dieselben zerebralen Areale aktiviert wie eine beliebige Tonfolge gleicher Komplexität – nämlich jene, die an Analysen akustischer Sinnesreize beteiligt sind. Der Unterschied zwischen ästhetischer und nicht-ästhetischer Verarbeitung ergäbe sich durch den Umstand, dass die Neurone in unterschiedlichen „Mustern“

feuern – es würde sich lediglich der zeitliche Aspekt der Aktivität unterscheiden. Eine analoge Veränderung des Verarbeitungsmusters könnte dann auch mit der Betrachtung eines als schön empfundenen Gemäldes oder einer Skulptur einhergehen – in diesem Fall innerhalb der mit visueller Verarbeitung assoziierten Areale. Zur Untersuchung dieser zeitlichen Charakteristiken der neuronalen Aktivität sind zeitlich hochauflösende Verfahren wie EEG und MEG geeignet. Theoretisch könnte es eine weitere neuronale Instanz geben, die speziell auf diese zeitlichen Muster reagiert. In diesem Fall wäre es wieder ein Areal, welches das Korrelat zur ästhetischen Empfindung bildet.

(c) Suche nach dem Neurotransmitter

Schließlich kann auch eine modulierende Beteiligung neuronaler Botenstoffe bei ästhetischen Verarbeitungsprozessen angenommen werden, wobei diese Annahme sowohl mit der Hypothese eines spezifischen Verarbeitungsprinzips als auch mit der Annahme eines spezifischen Areals durchaus in Einklang zu bringen ist. Insbesondere die Monoamine (Dopamin, Serotonin, Noradrenalin) scheinen eine modulierende Wirkung auf die Aktivität verschiedener Hirnregionen auszuüben. Die Ausschüttung dieser Neurotransmitter könnte sowohl ein bestimmtes Verarbeitungsprinzip in bestimmten Hirnregionen „anstoßen“ als auch die (vermehrte) Aktivierung bestimmter Hirnregionen initiieren. Das Ausmaß der Ausschüttung kann mit Methoden wie der Positronen-Emissions-Tomographie (PET) oder der Single-Photon-Emissionscomputertomographie (SPECT) untersucht werden. Sie bedienen sich der Injektion radioaktiv markierter Stoffwechselprodukte in den Blutkreislauf, um Hirnregionen zu identifizieren, in denen ein bestimmtes Stoffwechselprodukt besonders stark „nachgefragt“ wird.

Wie bereits angedeutet, schließen sich die den drei Perspektiven zugrunde liegenden Untersuchungshypothesen keinesfalls gegenseitig aus. In der Praxis ist eine Trennung vor allem der ersten beiden Ansätze sogar kaum möglich. Eine Veränderung des zeitlichen Musters der neuronalen Aktivität bringt eine Veränderung der neuronalen Gesamtaktivität einer Region mit sich, sofern sie als Mittelung über ein zuvor definiertes Zeitintervall erfasst wird (wenngleich diese Veränderung unterschiedlich stark ausgeprägt sein kann). Da im Humanbereich zurzeit noch keine Methoden mit gleichzeitig sehr hoher räumlicher und zeitlicher Auflösung zum Einsatz kommen, hat ein empirisches Vorgehen beide bzw. alle drei Perspektiven zu berücksichtigen. Dies wird zunehmend dadurch gewährleistet, dass parallel mehrere Methoden zur Abdeckung der unterschiedlichen Perspektiven eingesetzt werden. Das von Fechner angestrebte *intensive Mass der Wohlgefälligkeit* – als Maß der *Einzel-Lust* – steht derzeit noch nicht zur Verfügung.

Entsprechend der für dieses Buch charakteristischen Rolle der Theorien wird die Ästhetikforschung nicht nach Themenfeldern oder Methoden geordnet. Vielmehr folgt die Darstellung und Diskussion ausgewählter Theorien und Modelle. Im Einzelnen werden genauer betrachtet: die „Theorie kognitiver Hedonik“ (Martindale, 1984, 2007b), die „Peak-Shift-Hypothese“ (Ramachandran und Hirstein, 1999), die „Theorie der Verarbeitungsflüssigkeit“ (Reber, Schwarz und Winkielman, 2004), das „Universelle Modell ästhetischer Wahrnehmung“ (Redies, 2007) sowie das „Modell der ästhetischen Erfahrung“ (Leder, Belke, Oeberst und Augustin, 2004). Die Auswahl erfolgt aus mehreren Gründen: 1. Alle diese Theorien und Modelle legen den Fokus auf den Prozess der geistigen Verarbeitung. 2. Sie haben in besonderer Weise die wissenschaftliche Diskussion und die empirische Forschungstätigkeit angeregt. 3. Sie sind interessant im Sinne einer allgemeinen Ästhetik, d. h. ihre Annahmen sind nicht a priori auf eine Sinnesmodalität oder eine Kunstgattung beschränkt, selbst wenn sie hieran expliziert werden. Die Reihenfolge der Darstellung ist keine

Gewichtung, sondern entspricht – bis auf das Modell von Leder et al. (2004) – der Chronologie ihres Entstehens. Ihr „Modell der ästhetischen Erfahrung“ ist ein Rahmenmodell zur Strukturierung der Forschung und wird abschließend dargestellt.

Theorie der kognitiven Hedonik

Die Theorie der kognitiven Hedonik (Martindale, 1984, 2007b) erklärt das ästhetische Erleben auf der Basis von Aktivierungsmustern in neuronalen Netzwerken. Wesentlich zum Verständnis der Theorie ist das Konzept *neuronaler Knoten* (cognitive units). Diese stehen durch Verbindungen unterschiedlicher Stärke miteinander in Beziehung (vgl. Abbildung 96).

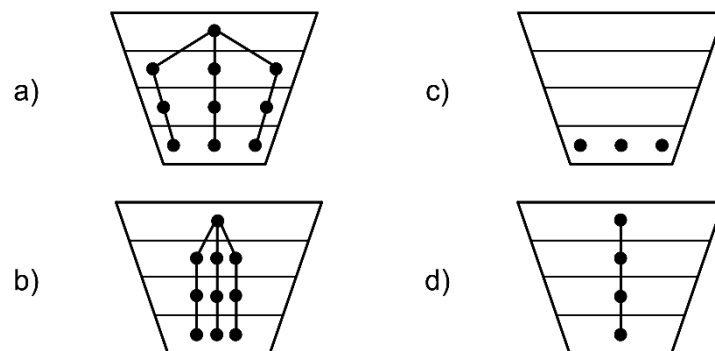


Abbildung 96: Hierarchisches Netzwerkmodell. Hypothetisches Aktivierungsmuster in einem Modul mit vier Schichten, wenn das Material folgende Merkmale aufweist: (a) Einheit und Vielfalt, (b) Einheit ohne Vielfalt, (c) maximale Vielfalt, (d) maximale Einheit. Vertikale Verbindungen wirken exzitatorisch, laterale Verbindungen inhibitorisch. Nach dem Modell kommt es bei (a) durch maximale Exzitation und minimale Inhibition zu der stärksten ästhetischen Empfindung (Martindale, 1984).

Jeder Knoten übernimmt eine spezifische Kodierung. Zum Beispiel speichert er das Konzept „Bild“ oder das Konzept „Skulptur“. Ein Knoten besitzt ein mit ihm assoziiertes *Aktivitätsniveau*, das über die Zeit variieren kann. Wenn ein Knoten ausreichend aktiviert ist, wird sein „Inhalt“

bewusst und rückt, zusammen mit dem Inhalt aller anderen ausreichend aktivierten Knoten, in den Bereich der Aufmerksamkeit. Nicht alle Knoten jedoch können in gleichem Maße aktiviert werden. Vielmehr weisen verschiedene Knoten, beispielsweise in Abhängigkeit von der Häufigkeit, mit der sie bereits früher aktiviert wurden, ein unterschiedliches Aktivitätsniveau auf.

Die Verbindungen zwischen einzelnen Knoten können entweder erregend (exzitatorisch) oder hemmend (inhibitorisch) wirken. Wird ein Knoten ausreichend aktiviert, sendet er Signale zu allen anderen Knoten, mit denen er verbunden ist. Je nachdem, ob im Einzelfall eine erregende oder eine hemmende Verbindung besteht, wird das Aktivitätsniveau dieser Knoten entsprechend erhöht oder verringert. Die Verbindungen zwischen den Knoten werden „gelernt“, d. h. sie können an Stärke zu- oder abnehmen. Knoten sind in *Modulen* organisiert, wobei jedem Modul spezifische Aufgaben zugeschrieben werden. So werden z. B. separate Module für die Analyse von Form, Farbe und Bewegung angenommen. Innerhalb eines Moduls liegen Knoten räumlich nahe beieinander, die ähnliche Konzepte kodieren. Solche, die weniger ähnliche Konzepte kodieren, liegen dem Ausmaß der Unähnlichkeit der zugehörigen Konzepte entsprechend weiter voneinander entfernt.

Jedes Modul ist aus mehreren *Schichten* aufgebaut (hierarchischer Aufbau), wobei vertikale Verbindungen (von einer Schicht zur nächsten) exzitatorisch wirken, laterale Verbindungen (innerhalb einer Schicht) jedoch inhibitorisch. Es gilt: Je näher zwei Knoten einer Ebene beieinander liegen, desto stärker ist der inhibitorische Einfluss, den sie aufeinander ausüben können.

Das Erleben von Schönheit und ästhetischem Vergnügen erklärt Martindale (1984) nun durch folgendes einfaches Prinzip: Ästhetisches Vergnügen entsteht durch Maximierung der aktivierenden Einflüsse auf

einzelne Knoten bei gleichzeitiger Minimierung der hemmenden Einflüsse. Einschränkend fügt er hinzu, dass das Ausmaß an Aktivierung auf jeder Stufe eines Moduls durch die Wirkung der lateralen Hemmung als relativ konstant anzusehen ist und das Ausmaß des ästhetischen Vergnügens somit oft durch die Verteilung aktivierter Knoten innerhalb dieser Schicht bestimmt wird.

Seine Überlegungen präzisiert Martindale in mathematischer Form durch die Gleichung: $P_i = p_i - d_i$, wobei sich P auf den „hedonischen Klang“ eines gegebenen Stimulus bezieht, p auf das Ausmaß des von ihm hervorgerufenen Wohlgefallens, d auf das Ausmaß des von ihm hervorgerufenen Missbehagens und der Index i auf den jeweiligen Knoten. Diesen einfachen Zusammenhang konkretisiert Martindale, indem er p_i und d_i durch separate Gleichungen jeweils wieder mathematisch fasst und dabei eine Vielzahl zusätzlicher Variablen einbezieht, z. B. Schwellenwerte für die Induktion von Behagen und Missbehagen (vgl. ausführlich Martindale, 1984).

Die ästhetische Wirkung des Prinzips „einheitliche Verknüpfung des Mannigfaltigen“ oder „Einheit in der Vielfalt“ ließe sich demnach so erklären: Ein ästhetischer Stimulus aktiviert innerhalb einer Ebene eines Moduls genau solche Knoten, die maximal weit voneinander entfernt sind, und gewährleistet, dass die gegenseitige Hemmung zwischen den aktivierten Knoten auf ein Minimum begrenzt ist. Gleichzeitig jedoch sollte sich dieses Erregungsmuster auf jeder einzelnen Ebene der Hierarchie eines Moduls zeigen, damit insgesamt ein maximales Ausmaß an Erregung zwischen den Ebenen vorliegt.

Betrachten wir das Modell anhand eines fiktiven Beispiels: Eine Studentin der Kunstgeschichte befindet sich im Palazzo Altemps in Rom, der eine bedeutende Sammlung römischer Marmorskulpturen enthält. Beim Betreten eines Ausstellungsraumes fällt ihr Blick auf eine antike Statue, die

sie bei genauerer Betrachtung als „Herkules vernichtet die neunköpfige Hydra“ erkennt. Nach dem Modell von Martindale (1984, 2007b) werden im Gehirn der Betrachterin die neuronalen Knoten für die Konzepte „Skulptur“, „Herkules“ und „Hydra“ aktiviert. Diese wiederum aktivieren weitere Konzepte, so zum Beispiel die Knoten für „Halbgott“ und „Schlange“. Es bildet sich ein Netzwerk von aktivierten und miteinander in Verbindung stehenden neuronalen Knoten. Die weitere Analyse der jungen Frau widmet sich dem Material, aus dem die Statue geschaffen ist. Sie betrachtet den heute vergilbten Marmor und dessen schwarze Einschlüsse. Dies aktiviert im Modul „Farbe“ die neuronalen Knoten für „vergilbt“ und „schwarz“. Aufgrund ihres Hintergrundwissens vermutet sie, dass es sich um die römische Kopie eines griechischen Originals handelt. Ein Blick auf die seitlich angebrachte Beschreibung bestätigt dies. Die neuronalen Knoten für „Original“ und „Kopie“ werden aktiviert (für eine nicht vorgebildete Betrachterin würde hier möglicherweise keine Aktivierung stattfinden).

Stark vereinfacht lässt sich vermuten, dass sowohl die Konzepte „Schlange“ und „Halbgott“ als auch die Konzepte „vergilbt“ und „schwarz“ jeweils nur relativ wenig miteinander gemeinsam haben. Die zugehörigen Knoten befinden sich jeweils in einiger Distanz zueinander und üben nur ein minimales Ausmaß gegenseitiger Hemmung aufeinander aus. Dies führt zu einer *Netto-Aktivierungsmaximierung* in den jeweiligen Modulen und in der Folge zu stärkeren ästhetischen Empfindungen, als wenn die Knoten in den Modulen benachbart wären.

Die Theorie kommt mit wenigen Grundannahmen aus und kann eine Vielzahl empirischer Befunde erklären (vgl. Martindale, 2007b). Sie ist in sich konsistent und ihre Grundannahmen lassen sich, obwohl es eigentlich eine abstrakte Theorie ist, leicht visualisieren. Die mathematische Formulierung der Theorie suggeriert zunächst ein hohes Maß an Konkretheit. Dies kann jedoch nicht darüber hinwegtäuschen, dass eine

empirische Überprüfung schwierig ist. Durch eine geeignete Justierung der Parameter ist eine Anpassung an eine Vielzahl von Daten möglich. Martindale (2007b) macht eine Reihe von Vorhersagen zur Überprüfung seiner Annahmen. Zwei dieser Hypothesen sollen hier beispielhaft erläutert werden: Zum einen geht Martindale davon aus, dass im Falle der sukzessiven Präsentation ähnlicher Stimuli die Präferenz für diese Stimuli mit der Anzahl der aufeinander folgenden Präsentationen kontinuierlich abnimmt. Dies geschieht durch die zunehmende laterale Hemmung. Wird dann jedoch ein andersartiger Stimulus präsentiert, wird der zugehörige Knoten im Vergleich zu den Knoten der vorherigen Stimuli deutlich weniger lateral gehemmt. Er erfährt somit in der Summe mehr Aktivierung und der zugehörige Stimulus erzeugt ein stärkeres ästhetisches Vergnügen. Zum anderen lässt sich aus der Theorie ableiten, dass sowohl die Anzahl als auch die Verschiedenartigkeit von Assoziationen einen Einfluss auf die Präferenz für diesen Stimulus hat: Viele und unterschiedliche Assoziationen sollten demnach zu stärkerem ästhetischen Vergnügen führen, da auf diese Weise die Wahrscheinlichkeit steigt, dass weit auseinanderliegende Knoten auf jeder Ebene aktiviert werden, die sich nur minimal lateral hemmen.

Martindale (2007b) führt mehr als 20 verschiedene Effekte an, die auf der Basis seiner Theorie erklärt werden können. Er unterscheidet verschiedene Gruppen: 1. Effekte, die sich ausschließlich durch die Maximierung der Aktivierung einzelner Knoten ergeben. Ein Beispiel hierfür ist der positive monotone Zusammenhang zwischen Farbsättigung und Präferenz. 2. Effekte der Aktivierungsmaximierung bei gleichzeitiger Hemmungsminimierung. Ein Beispiel ist die unangenehme Wirkung von „weißem Rauschen“. Hier kommt es zwar zu viel Aktivierung, aber auch zu einem Höchstmaß an lateraler Hemmung. 3. Effekte der Maximierung der Anzahl aktivierter Module. Unter Bezugnahme auf eine eigene Untersuchung (Martindale, 1990) wird angenommen, dass Sinnhaftigkeit (meaningfulness) der bedeutendste Faktor mit Wirkung auf ästhetisches

Erleben ist. In Zusammenhang mit dem Erleben von Bedeutung kommt es zur Aktivierung eines speziellen „semantischen Moduls“. 4. Effekte der wiederholten Reizdarbietung, der Aufmerksamkeit und iterative Effekte. Hier bezieht sich Martindale u. a. auf den Mere-Exposure-Effekt (Zajonc, 1968, 1980): Bloße Wiederholung der Stimuluspräsentation führt zu gesteigerten Präferenzen.

Martindale (2007b) gelingt es so, eine Vielfalt bekannter empirischer Befunde einzubinden, die von Farbpräferenzen über Typikalitätseffekte (einschließlich einer Präferenz für Prototypikalität) bis hin zu so komplexen Wirkungen reichen, wie sie durch sprachliche Metaphern ausgelöst werden. Eine offensichtliche Erklärungslücke bezieht sich auf lerntheoretische Effekte wie zum Beispiel das Peak-Shift-Prinzip (s. u.), das bei Ramachandran und Hirstein (1999) eine zentrale Rolle einnimmt. Eine wesentliche Einschränkung der Theorie besteht darin, dass sie sich – zumindest in der Formulierung von 1984 – ausschließlich auf *desinteressiertes Vergnügen* (disinterested pleasure) bezieht, womit eine distanzierte, überwiegend kognitive Haltung des Subjekts gemeint ist. Die sonst in der Ästhetikforschung im Vordergrund stehende enge Anbindung an die Emotionsforschung und die Betonung des Aktivierungsniveaus fehlt (vgl. im Gegensatz dazu Silvia, 2005). Hier grenzt sich die Theorie deutlich ab. Das ästhetische Erleben wird, dem Titel der Theorie („kognitive Hedonik“) folgend, als ein überwiegend kognitiver Vorgang betrachtet. Eine Bewertung der Vereinbarkeit der Theorie mit Befunden neurowissenschaftlicher Studien ist trotz ihrer explizit neuro-kognitiven Orientierung nur begrenzt möglich. Die zentrale Annahme, dass für jedes mentale Konzept ein einzelner Knoten aktiviert wird, ist derzeit nicht überprüfbar. Die Formulierung der Theorie legt nahe, dass diese Knoten als räumlich begrenzt und im Wesentlichen distinkt anzusehen sind, d. h. sich nicht überlappen. In fMRT-Studien werden – aufgrund der Notwendigkeit der statistischen Fehlerkorrektur – aktivierte Hirnareale aber nur dann berichtet, wenn sie eine gewisse Größe überschreiten.

Schließlich macht das Modell keine Aussagen über wahrnehmbare Eigenschaften, die einen Gegenstand als „ästhetisch“ kennzeichnen. Es integriert keine Variablen der Außenwelt. Daher fällt es schwer, Konsequenzen für eine Objektästhetik abzuleiten.

Peak-Shift-Prinzip

Ramachandran und Hirstein (1999) haben in einem viel beachteten Artikel mit dem Titel „The Science of Art, A Neurological Theory of Aesthetic Experience“ acht Prinzipien vorgeschlagen, die für die ästhetische Wahrnehmung grundlegend sein sollen. Bei diesen Prinzipien geht es um die Frage: Wie sieht eine optimale (visuelle) Stimulation des Gehirns aus? Im Einzelnen machen sie folgende Annahmen:

1. Der Peak-Shift-Effekt ist grundlegend für die ästhetische Wahrnehmung.
2. Perzeptuelle Gruppierung und Bindung wirken unmittelbar verstärkend.
3. Es ist effektiver, eine visuelle (Sub-)Modalität einzeln anzusprechen (eine Strichzeichnung sei z. B. effektiver als eine Photographie).
4. Die Extraktion von Kontrastinformation wirkt verstärkend.
5. Perzeptuelles Problemlösen ist verstärkend.
6. Symmetrie hat einen ästhetischen Wert.
7. Es werden Objektinterpretationen bevorzugt, die von einer Vielzahl möglicher Blickpunkte gültig sind.
8. Visuelle Metaphern und Wortspiele wirken ästhetisch.

Im Mittelpunkt ihrer eigenen Überlegungen, aber auch der sich an diese Arbeit anschließenden kontroversen Diskussion, steht der *Peak-Shift-Effekt*. Ramachandran und Hirstein (1999) postulieren, dass ein wesentliches Merkmal von Kunst die Verstärkung, das überdeutliche Zur-

Schau-Stellen der Charakteristika eines Stimulus ist. Exemplarisch verdeutlichen sie dies anhand von Gesichtsdarstellungen. Ein Karikaturist beschreibt eine Person, indem er den Unterschied zwischen dem individuellen Gesicht und einem durchschnittlichen Gesicht ermittelt und diesen Unterschied dann überbetont. Zentral ist die Maximierung der Differenz zwischen dem dargestellten Stimulus und einem durchschnittlichen Stimulus. Im Mittelpunkt steht dabei in der Regel eine bestimmte Eigenschaft des Stimulus (z. B. die Form des Gesichtes; vgl. Abbildung 88).



Abbildung 97: Le passé, le présent, l'Avenir (Louis-Philippe d'Orléans), Lithographie (Honoré Daumier, 1834). Ein Beispiel für die Überbetonung der Stimuluseigenschaft „Form“.

Durch die damit verbundene Fokussierung auf die *Rasa*, die Essenz des Stimulus, wird er als besonders ästhetisch wahrgenommen. Als

Schlüsselmechanismus wird der Peak-Shift vermutet, ein Effekt, der ursprünglich im Rahmen der Lernforschung an Tauben entdeckt wurde (Hanson, 1959; Jenkins und Harrison, 1962) und auf der instrumentellen Konditionierung beruht: Nicht der mit Belohnung assoziierte Ton selbst erzeugt die maximale Reaktionsrate, sondern ein Ton, der das Charakteristische an der Differenz zwischen diesem und dem nicht mit Belohnung assoziierten Ton verstärkt.

Diesen im Tierexperiment gut belegten Befund übertragen Ramachandran und Hirstein (1999) auf die ästhetische Wahrnehmung beim Menschen: Mittels des Peak-Shift findet eine Fokussierung auf die Essenz eines Stimulus statt, die vom Gehirn belohnt und als besonders ästhetisch empfunden wird. Plakativ fassen sie ihre These in dem Satz zusammen: „All art is caricature“. Anhand des eingangs konstruierten Beispiels lässt sich die Hypothese von Ramachandran und Hirstein folgendermaßen verdeutlichen: Die Kunststudentin betritt das Museum nicht als „unbeschriebenes Blatt“. In ihrem bisherigen Leben hat sie bereits eine Vielzahl menschlicher Körper gesehen. Auf diese Weise hat sie eine Vorstellung davon, wie ein durchschnittlicher männlicher und wie ein durchschnittlicher weiblicher Körper aussieht. Dabei hat sie eine Präferenz für solche männlichen Körper entwickelt, die den Unterschied zwischen den charakteristischen männlichen Merkmalen (breite Schultern, muskulöse Oberarme, markantes Gesicht etc.) und den typisch weiblichen Merkmalen überakzentuieren. Diesem Stereotyp, dieser „Rasa“ des männlichen Körpers entspricht die Herkulesstatue. Deshalb wird sie als besonders attraktiv wahrgenommen.

Problematisch ist, dass von den Autoren der Zusammenhang zwischen phylogenetischen und ontogenetischen Einflüssen auf den Peak-Shift-Effekt nicht thematisiert wird. Es bleibt unklar, inwiefern die Präferenz für bestimmte *Super-Stimuli* biologisch bereits angelegt ist oder während der individuellen Entwicklung erworben wird. Zumindest bei dem von

Ramachandran und Hirstein (1999) gewählten Beispiel des weiblichen Körpers scheint aus evolutionärer Sicht eine biologische Grundlegung naheliegend. So fand z. B. Singh (1993) bei Männern eine Tendenz zur Bevorzugung einer deutlich akzentuierten weiblichen Brust und Taille. Eine solche reizspezifische Präferenz ist jedoch in der lerntheoretischen Perspektive des Peak-Shift nicht vorgesehen. Insgesamt beruht die Theorie von Ramachandran und Hirstein zwar auf lerntheoretischen Annahmen, die jedoch nicht konsequent auf den Bereich der Ästhetik angewendet werden.

Auch in sich ist die Theorie wenig konsistent. Vor allem die Verknüpfung der Idee einer „Rasa“ mit dem Slogan „All art is caricature“ erscheint problematisch. Fraglich ist, wie die „Rasa“, die Essenz eines Stimulus, mit einer durch den Peak-Shift-Effekt erworbenen Präferenz übereinstimmen kann. Letzterer kann die Präferenz für einen Stimulus stets nur in Abgrenzung von einem anderen, zuvor gelernten Reiz erklären. Somit kann nicht davon ausgegangen werden, dass durch den Peak-Shift die universelle Essenz eines Stimulus erlernt wird. Streng genommen sind Rasa und Peak-Shift nur vereinbar, wenn man für jeden Menschen entweder eine individuelle Rasa annimmt, die sich erst im Laufe des Lebens herausbildet, oder wenn die gesamte Wirkung des Peak-Shift-Effekts der phylogenetischen Entwicklung des Menschen zugeschrieben wird.

Die Peak-Shift-Hypothese von Ramachandran und Hirstein (1999) wurde vielfach kritisiert. Im Mittelpunkt stand vor allem die Frage der Übertragbarkeit des Peak-Shift-Effekts auf die Ästhetik und die von Ramachandran und Hirstein postulierte Allgemeingültigkeit für den Bereich der Kunst. Nach Hyman (2006) handelt es sich, mit Blick auf die besondere Bedeutung ausgeprägter, geschlechtsprototypischer körperlicher Attribute bei der Partnerwahl, eher um eine Theorie der sexuellen Anziehung als eine Theorie der Ästhetik. Mangan (1999) weist

darauf hin, dass die Theorie auf der Generalisierung von Annahmen beruht, die auf der Basis einer nicht repräsentativen Auswahl von Kunstwerken formuliert wurde. Ebenso leicht hätten sich unzählige Beispiele finden lassen, die die zentrale These widerlegen. Kindy (1999) macht darauf aufmerksam, dass kubistische Werke, die durch die bewusste Konzentration auf das Wesen des Dargestellten nach Ramachandran und Hirstein besonders geschätzt werden sollten, von einer breiten Öffentlichkeit keineswegs favorisiert werden und Wallen (1999) betont, dass viele bedeutende fotografische Werke, die gerade wegen ihres Realismus geschätzt werden, nicht mit dem Ansatz von Ramachandran und Hirstein vereinbar sind. Martindale (1999) führt als prominentes Gegenbeispiel die italienische Malerei zur Zeit der Renaissance an: Hier habe das künstlerische Bestreben gerade darin bestanden, die Dinge möglichst „natürlich“ darzustellen. Sicherlich hat Mangan Recht, wenn er ausführt, dass Verzerrung alleine noch kein Garant für die Entstehung von Kunst ist, unabhängig davon, ob es eine ästhetische Präferenz für Übersteigerungen auf einer Dimension gibt.

Empirische Unterstützung findet die Peak-Shift-Hypothese durch Befunde aus der experimentellen Forschung zur Wahrnehmung menschlicher Gesichter (vgl. Roye, Höfel und Jacobsen, 2008). Ein mittlerweile mehrfach bestätigter Befund ist, dass ein „Durchschnittsgesicht“, das am Computer aus einer Vielzahl realer Gesichter berechnet wurde, als attraktiver wahrgenommen wird als die ursprünglichen Gesichter, aus denen es entstanden ist (z. B. Langlois und Roggman, 1990). In Übereinstimmung mit Ramachandran und Hirstein konnten Perrett, May und Yoshikawa (1994) jedoch zeigen, dass Gesichter, in denen bestimmte Merkmale übertrieben dargestellt sind – zum Beispiel überdurchschnittliche Größe der Augen – als noch attraktiver wahrgenommen werden als die Durchschnittsgesichter. Sie verbuchen einen „Attraktivitätsgewinn“.

Die Peak-Shift-Hypothese von Ramachandran und Hirstein (1999) lässt sich am besten auf die schon von den Autoren selbst gewählten Beispiele anwenden. Zum einen besitzt sie einen augenscheinlichen Erklärungswert für die Wirkung von Karikaturen. Zum anderen bietet sie eine Möglichkeit, die physische Attraktivität im Zusammenhang mit sexueller Selektion zu erklären. Unklar bleibt jedoch, wie stark die Überakzentuierung ausfallen muss. Gilt hier ein „je mehr, desto besser“ oder finden sich eher nichtlineare Zusammenhänge? Das fundamentale Problem aber ist, dass die insgesamt acht Annahmen keine Theorie im strengen Sinne darstellen. Es sind einzelne Hypothesen, die sich empirisch prüfen lassen und die eine, wenn auch unterschiedliche, empirische Evidenz aufweisen. So sind zum Beispiel die Annahmen 2 und 6 (Gruppierung und Symmetrie) gut bestätigt. Es wird jedoch keine Struktur für diese Annahmen expliziert. Ihr Verhältnis untereinander bleibt unklar.

Theorie der Verarbeitungsflüssigkeit

Die „Theorie der Verarbeitungsflüssigkeit“ von Reber, Schwarz und Winkielman (2004) postuliert die direkte Abhängigkeit der ästhetischen Bewertung eines Stimulus und des ästhetischen Vergnügens (pleasure) vom Verarbeitungsaufwand. Die grundlegende Hypothese lautet, dass ein positiver Zusammenhang zwischen der Leichtigkeit, mit der ein Stimulus verarbeitet wird, und der ästhetischen Wirkung dieses Stimulus besteht. Reber et al. (2004) unterscheiden zwischen einer perzeptuellen und einer konzeptuellen Flüssigkeit. *Perzeptuelle Flüssigkeit* bezeichnet die Leichtigkeit, mit der die physikalischen Eigenschaften des Stimulus identifiziert werden können. Beispielsweise enthält eine Bleistiftskizze, die ein menschliches Gesicht nur anhand seiner Kontur darstellt, weniger Information als eine andere, die zusätzlich durch Schattierung auch eine Räumlichkeit erzeugt. Erstere kann somit perzeptuell flüssiger verarbeitet werden. *Konzeptuelle Flüssigkeit* umfasst dagegen die Leichtigkeit, mit der die geistigen Operationen ausgeführt werden können, die notwendig sind,

um dem Stimulus eine Bedeutung zu geben bzw. ihn in bestehende semantische Wissensstrukturen einzuordnen. Hier geht es – am Beispiel der Skizze – um das Erkennen des Gesichts. Ob eine detailreiche oder detailarme Skizze besser erkannt wird, diese Frage kann nicht generell beantwortet werden, sondern hängt nicht zuletzt davon ab, wie bedeutsam oder redundant die Informationen sind, die ausgelassen werden.

Je flüssiger perzeptuelle und konzeptuelle Verarbeitung sind, desto ästhetischer erscheint der Stimulus. Die Verarbeitungsflüssigkeit ergibt sich aus einer Interaktion der Eigenschaften des Stimulusmaterials mit den kognitiven und affektiven Prozessen beim Wahrnehmungssubjekt. Auf Seite der Stimuluseigenschaften erleichtern beispielsweise Symmetrie und Kontrast die Verarbeitung, indem sie die Menge an Information reduzieren (Symmetrie) bzw. es dem visuellen System erleichtern, diese zu extrahieren (Kontrast). Die Verarbeitungsflüssigkeit soll nicht nur einen Einfluss auf das unmittelbare ästhetische Vergnügen und das spontane Urteil, sondern darüber hinaus auch Auswirkungen auf reflektierte ästhetische Urteile haben. Die Idee ist, dass spontane und subjektive Eindrücke als (heuristische) Informationen in solche reflektierten Urteile eingehen können. Schließlich nehmen Reber und Kollegen einen moderierenden Einfluss von Erwartungen und Attributionen auf die Leichtigkeit der Verarbeitung an. Ist die Verarbeitung flüssiger als vorab erwartet (im Sinne einer positiven Differenz zu einem Erwartungswert), wirkt sich dies positiv auf das ästhetische Vergnügen aus und geht mit einem besonders starken Affekt einher. Bei einer negativen Differenz zum Erwartungswert ist der Affekt hingegen abgeschwächt. Attributionen bezeichnen die rückgeschlossene Ursache eines Ereignisses. Wird die Flüssigkeitserfahrung auf eine irrelevante Quelle zurückgeführt, schwächt dies deren Einfluss.

Das Verständnis der Theorie soll wieder durch unser Beispiel erleichtert werden. Die Studentin betrachtet die Statue. Die Beleuchtung ist

ausreichend, um einen angemessenen Kontrast zu erzielen (perzeptuelle Flüssigkeit). Durch ihre kunstgeschichtliche Expertise erkennt sie ohne Schwierigkeiten, dass es sich bei der Darstellung um Herkules mit Hydra handelt (konzeptuelle Flüssigkeit). Die leichte Erkennbarkeit der Darstellung führt bei ihr direkt zu einem positiven Gefühl. Dieses bleibt jedoch relativ gering, da sie sich ihrer Expertise durchaus bewusst ist (Einfluss der Erwartung).

Mit einer Zusammenstellung konsistenter empirischer Befunde stützen Reber et al. (2004) ihre Theorie. Zwei besonders bedeutsame und gut replizierte Ergebnisse sind: 1. Die wiederholte Darbietung desselben Stimulus führt zu gesteigerten Präferenzen für diesen Stimulus (Mere-Exposure-Effekt, vgl. Zajonc, 1968, 1980). 2. Prototypische Stimuli werden gegenüber nicht-prototypischen präferiert (vgl. Hekkert und van Wieringen, 1990; Martindale und Moore, 1988; Whitfield und Slatter, 1979). Der Theorie folgend sind beide Effekte auf die Flüssigkeit bei der Informationsverarbeitung zurückzuführen: Mit steigender Anzahl an Präsentationen gelingt eine flüssige Verarbeitung zunehmend besser und ein prototypischer Stimulus kann per se besonders leicht verarbeitet werden.

Eine wichtige Herausforderung für die Theorie der Verarbeitungsflüssigkeit ist die Beobachtung, dass Menschen in der Regel nicht möglichst einfache Stimuli, sondern Stimuli mittlerer Komplexität als ästhetisch beurteilen (Berlyne, 1971). Im Sinne des Ansatzes von Reber et al. (2004) liegt es zunächst nahe anzunehmen, dass zunehmende Stimuluskomplexität mit verringerter Leichtigkeit bei der Verarbeitung und so auch mit weniger positiven Affekten einhergeht. Eine solch eindimensionale Betrachtungsweise missachtet jedoch die „internen“ Bedingungen bei den Wahrnehmenden. Reber et al. (2004) bieten ergänzende Erklärungen an, wie die Befunde zur Komplexität mit ihrer Theorie in Übereinstimmung gebracht werden können. Eine bedeutsame

Rolle spielt der weiter oben beschriebene Einfluss der Erwartungen. Auch bei komplexem Stimulusmaterial kann die Verarbeitung leichter sein als ursprünglich angenommen (z. B. indem Regelmäßigkeiten in einem zunächst unstrukturiert erscheinenden Stimulus erkannt werden). Ein besonders starkes Erleben ästhetischen Vergnügens ist die Folge. Außerdem ist die Unterscheidung zwischen perzeptueller und konzeptueller Flüssigkeit relevant. Eine Zunahme perzeptueller Komplexität mag durch eine gleichzeitige Abnahme konzeptueller Komplexität mehr als kompensiert werden (z. B. indem eine zusätzliche Schraffur in einer Skizze die Eindeutigkeit der Darstellung erhöht). Überdies ist bekannt, dass insbesondere Expert*innen komplexe gegenüber einfachen Stimuli bevorzugen (McWhinnie, 1968). Nach Reber et al. (2004) ist dies darauf zurückzuführen, dass Expert*innen durch ihren häufigen Umgang mit entsprechendem Stimulusmaterial auch komplexe Stimuli „flüssig“ verarbeiten können. Expertentum ist durch eine konzeptuelle Flüssigkeit gekennzeichnet.

Reber und Kollegen integrieren eine ganze Reihe bestehender Befunde der empirischen Ästhetik in ihre Theorie und liefern einen neuen Erklärungsrahmen. Sie erbringen jedoch keinen direkten Nachweis, dass Verarbeitungsflüssigkeit eine kausale Wirkvariable auf das ästhetische Vergnügen darstellt. Die zur Stützung der Theorie dargestellten Befunde umfassen Zusammenhänge der Form: Manipulation X, die plausiblerweise eine Veränderung der Verarbeitungsflüssigkeit bewirkt, führt zu Änderungen in den berichteten Präferenzen. Der Nachweis, dass eine solche Veränderung in der Verarbeitungsflüssigkeit tatsächlich erfolgt, würde aber dessen quantitative Bestimmung erfordern. Eine Definition des Konstrukts „Verarbeitungsflüssigkeit“, die konkret genug ist, um eine solche Messung zu ermöglichen, fehlt. Messen ließen sich hingegen Eigenschaften von Stimuli. Es ist allerdings kaum spezifiziert, wie genau diese mit den inneren Verarbeitungsprozessen der Rezipient*innen zusammenwirken, um eine flüssige Verarbeitung zu bewirken. Unklar ist,

unter welchen (inneren) Bedingungen welches Ausmaß an Komplexität präferiert wird. Da also weder die Verarbeitungsflüssigkeit bei den Wahrnehmenden direkt gemessen werden kann, noch eine Beziehung zwischen Stimuluseigenschaften und ästhetischem Vergnügen aus der Theorie abgeleitet werden kann, ist es schwer, überprüfbare Vorhersagen zu formulieren.

Universelles Modell der ästhetischen Wahrnehmung

Als vierter Ansatz sei das „Universelle Modell der ästhetischen Wahrnehmung“ von Redies (2007) vorgestellt, das funktionelle resonatorische Wechselwirkungen zwischen zentralnervösen Neuronen als neurophysiologisches Korrelat der ästhetischen Wahrnehmung annimmt. Je größer das Ausmaß dieser Wechselwirkungen ist, das ein Stimulus auslöst, desto stärker ist die zugehörige ästhetische Empfindung.

Das Modell basiert auf der Idee der *effizienten Kodierung* (sparse coding): Eine sensorische Information wird dann effizient vom Gehirn kodiert, wenn in Reaktion auf einen sensorischen Stimulus möglichst wenige Neurone aktiv werden bzw. die Aktivität von Neuronen auf diesen Stimulus über die Zeit hin reduziert wird (Olshausen und Field, 2004). Grundlage für diese Annahme sind Forschungsergebnisse, die darauf hindeuten, dass einzelne Neurone aus den visuellen Arealen des Cortex auf komplexe visuelle Szenen – im Vergleich zu lokalen einfachen Stimuli – mit einer Reduktion ihrer Aktivität reagieren (Vinje und Gallant, 2002). Insbesondere gibt es nach Redies (2007) Befunde, die andeuten, dass solche Stimuli maximal effizient kodiert werden können, mit denen das Gehirn im Laufe der Phylogenese und der Ontogenese häufiger „Erfahrung“ erworben hat. Dies ist insbesondere bei „natürlichen Szenen“ der Fall.

Der entscheidende Punkt in der Theorie von Redies (2007) ist nun die Annahme, dass Kunst sich eben jener Mechanismen bedient, die eigentlich für die effiziente Kodierung komplexer natürlicher Szenen entwickelt wurden. Die These lautet, dass Künstler*innen ihre Werke so gestalten, dass eine Betrachterin oder ein Betrachter sie „effizient kodieren“ kann: „[T]he artist adapts his art to the visual system that, in turn, is adapted to the natural environment“ (Redies, 2007, S. 105). Für diese effiziente Kodierung ist nicht der Inhalt ausschlaggebend, sondern die Art der Darstellung. Entscheidend ist das Vorkommen von Invarianten. Hiervon hängt der ästhetische Wert eines Objektes ab, den die Betrachterin oder der Betrachter dem Gegenstand zuschreibt.

Methodisch gesehen geht Redies (2007) so vor, dass natürliche Szenen und Kunstwerke (graphische Werke aus unterschiedlichen Epochen) einer statistischen Analyse unterzogen werden. Als Ausgangsmaterial werden Graustufenbilder verwendet. Diese werden mittels der Fourieranalyse verglichen. Unabhängig vom jeweiligen Bildthema zeigen sich statistische Gemeinsamkeiten, die mithilfe des Konzeptes der *Skaleninvarianz* interpretiert werden. Ein Beispiel: Beim Hineinzoomen in eine Szene bleibt die Menge an Struktur sowohl bei den natürlichen Szenen als auch bei den Kunstwerken über einen weiten Zoombereich hinweg konstant. Solche Invarianzen finden sich hingegen nicht bei der Analyse von Graustufenbildern beliebiger Objekte mit geringer ästhetischer Wirkung (z. B. Fotografien von Alltagsobjekten und Innenräumen; vgl. Redies, Hasenstein und Denzler, 2007). Beispiele der von Redies, Hasenstein und Denzler untersuchten Vorlagen sowie Ergebnisse der Fourieranalysen zeigt Abbildung 98.

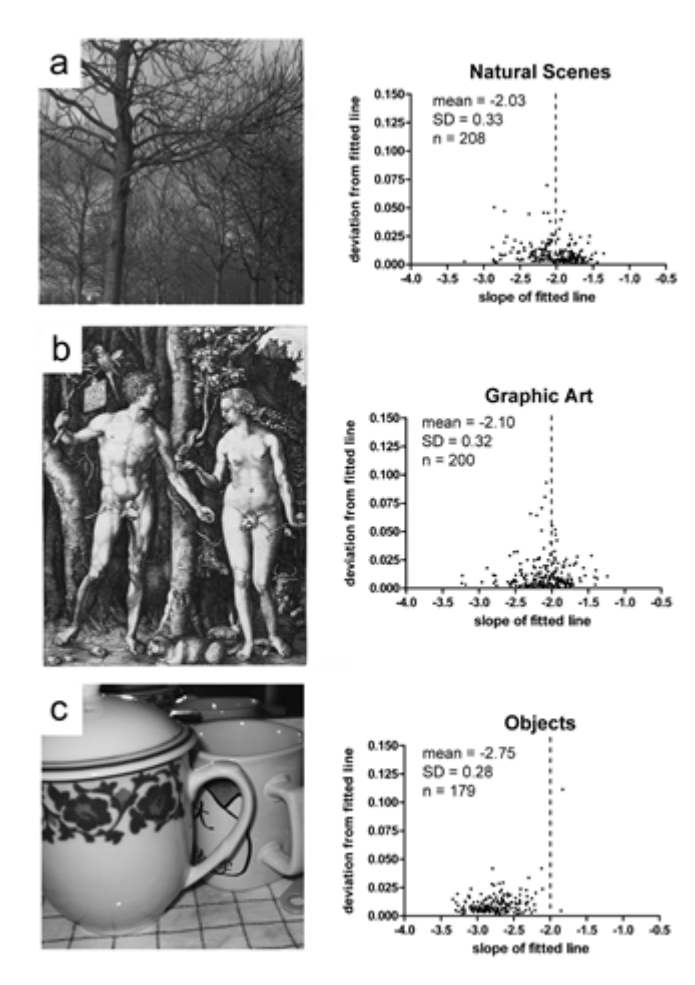


Abbildung 98: Beispiele der von Redies, Hasenstein und Denzler untersuchten Fotografien sowie Ergebnisse der Fourieranalysen für drei unterschiedliche Stimulusklassen: Beispielpfhaft sind eine natürliche Szene (a), ein Kunstwerk (b) und ein Haushaltsobjekt (c) dargestellt. Die Streudiagramme veranschaulichen für jede dieser Stimuluskategorien die Ergebnisse der Fourieranalysen aller verwendeten Stimuli: Auf der Abszisse ist die Steigung des linearen Zusammenhangs zwischen Signalstärke und Frequenz angegeben, auf der Ordinate die tatsächliche Abweichung von diesem Zusammenhang. Erkennbar sind Steigungswerte nahe -2 (Skaleninvarianz) für Kunstwerke und natürliche Szenen, nicht jedoch für Haushaltsobjekte (Redies, Hasenstein und Denzler, 2007).

Nach diesem Modell existieren keine spezifischen neuronalen Mechanismen, die für die ästhetische Wahrnehmung oder die Wahrnehmung von Kunstwerken verantwortlich sind. Die statistischen Eigenschaften eines Kunstwerkes entlehnt die Künstlerin oder der Künstler nicht den Eigenschaften des Dargestellten, sondern vielmehr

jenen, die sich gesetzmäßig in komplexen natürlichen Szenen finden: „Artists create art with statistical properties that are not necessarily the same as those of the subject depicted. Rather, artists adjust the image statistics of their subjects during the process of creation so that, in their works of art, statistics are more similar to those encountered in complex natural scenes“ (Redies, 2007, S. 103).

Das Modell von Redies (2007) beruht auf einer intensiv verfolgten Forschungslinie (z. B. Hoyer und Hyvärinen, 2002; Lennie, 2003; Vinje und Gallant, 2000, 2002; für einen Überblick vgl. Simoncelli und Olshausen, 2001), die auf Überlegungen von Barlow (1961) zurückgeht. Ihre grundlegende Idee: Um Ressourcen zu sparen, kodiert das Gehirn eingehende Informationen effizient, indem es Redundanzen aus dem vorliegenden Stimulusmaterial entfernt. Redies et al. (2007) verknüpfen diesen Gedanken mit der Ästhetikforschung. Das Modell besticht durch seine Einfachheit. Es wird ein einziger neuronaler Mechanismus angenommen, der das Ausmaß der durch ein Kunstwerk hervorgerufenen ästhetischen Wirkung zu erklären sucht. Individuelle, von Person zu Person variierende Merkmale werden ausgespart. Weniger überzeugt die Theorie in Bezug auf die Konkretheit der Vorhersagen. Das hängt unmittelbar mit dem Forschungsstand zur „adaptiven Resonanz“ zusammen. Zwar umschreiben die Konzepte „effiziente Kodierung“ (als Prozess) und „adaptive Resonanz“ (als Zustand) sprachlich angemessen die zugehörigen Abläufe im Gehirn. Offen bleibt aber zurzeit die Frage, wie, neuronal gesehen, die effiziente Kodierung im Gehirn erreicht wird bzw. wie über die effiziente Kodierung die adaptive Resonanz entsteht. Hierzu lassen sich lediglich eine Reihe von gleichberechtigten Hypothesen anführen, die eng mit der empirischen Bestimmung des Phänomens zusammenhängen (vgl. z. B. den Unterschied zwischen *lifetime sparseness* – über die Zeit hinweg wird jedes Neuron nur selten aktiv – und *population sparseness* – nur ein geringer Teil einer Zellpopulation wird durch einen gegebenen Stimulus aktiviert – nach Willmore und Tolhurst, 2001). Die

Weiterentwicklung des Modells hängt stark von der neuronalen Grundlagenforschung zur „effizienten Kodierung“ ab.

Die Ergebnisse zweier fMRT-Studien (Jacobsen, Schubotz, Höfel und v. Cramon, 2006; Kawabata und Zeki, 2004) sprechen dafür, dass sich neuronale Aktivierungsunterschiede für „schöne“ und „hässliche“ Stimuli nicht etwa in den visuellen, sondern vor allem in den frontalen Arealen des Cortex finden lassen. Nach Jacobsen et al. (2006) befinden sich die aktivierten Bereiche in teilweiser Überlappung mit Arealen, die an der Bewertung sozialer und moralischer Reize beteiligt sind. Unterschiede in der Aktivierung visueller Hirnareale finden sich nur dann, wenn die Stimuli auch inhaltlich unterschiedliche Darstellungen aufweisen. Die bewusste Bewertung eines Stimulus als schön oder hässlich findet demnach eher in frontalen Hirnregionen statt.

Redies universelles Modell liefert Erklärungen für solche ästhetischen Effekte, die sich aus der Art ihrer Darstellung ergeben. Befunde, die ästhetische Einflüsse in Abhängigkeit von anderen Stimulusmerkmalen beschreiben, wie Effekte der Prototypikalität (z. B. Martindale, Moore und West, 1988), der Komplexität (z. B. Berlyne, 1970, 1974; Eisenman, 1968; Eisenman und Rappaport, 1967) oder auch der Farbsättigung (z. B. Martindale und Moore, 1988) müssen unberücksichtigt bleiben. Das Gleiche gilt für Einflüsse, die sich aus der allgemeinen Erregung oder der Stimmungslage ergeben. Denkbar wäre, dass eine zu hohe oder zu niedrige Erregung verhindert, dass sich ein Zustand der adaptiven Resonanz im Gehirn aufbauen kann.

Die Annahme, ein Bild werde deshalb als „ästhetisch“ beurteilt, weil es die Skaleninvarianz der Darstellung gewährleistet, hat indirekte Evidenz aus fast allen Bereichen der visuellen Wahrnehmung. Die Extraktion von Invarianzen, die das System unabhängig von den Spezifika der aktuellen Wahrnehmungssituation macht (z. B. Besonderheiten der Beleuchtung),

gehört zu den grundlegenden Aufgaben des Sehsystems. Das Ziel besteht wesentlich darin, trotz sich permanent ändernder Stimulation eine stabile „visuelle Welt“ zu erzeugen, die Orientierung, Entscheidung und Bewegung erlaubt. Die vielleicht noch interessantere Frage ist, ob bedeutende Werke der Kunst nicht dadurch gekennzeichnet sind, dass sie noch besser an die Eigenschaften des Nervensystems angepasst sind als jede „natürliche Szene“.

Konzeptuell ist das Universelle Modell der ästhetischen Wahrnehmung der Theorie der Verarbeitungsflüssigkeit ähnlich. Während Redies (2007) davon ausgeht, dass ästhetische Stimuli „effizient“ kodiert werden, nehmen Reber und Kollegen an, dass sie „flüssig“ verarbeitet werden. Während Redies eine neurophysiologisch orientierte Perspektive einnimmt, beziehen sich Reber et al. (2004) auf psychologische Studien. Beide Theorien nehmen unterschiedliche Blickwinkel auf dieselbe grundlegende Idee ein: Distale Reize werden dann als ästhetisch empfunden, wenn der Verarbeitungsaufwand gering bzw. geringer als erwartet ist.

Prozessmodell der ästhetischen Erfahrung

Das „Modell der ästhetischen Erfahrung“ von Leder, Belke, Oeberst und Augustin (2004) wurde vor allem in Hinblick auf die ästhetische Wahrnehmung von visueller Kunst entwickelt (vgl. Abbildung 99). Als „Eingangsinformation“ setzt es ein Kunstwerk voraus, das aufgrund von Kontextfaktoren als solches erkannt und klassifiziert wurde. Ein solcher Kontextfaktor könnte z. B. die Tatsache sein, dass es sich in einem Museum befindet. Das Modell sieht fünf Stufen der Verarbeitung ästhetischer Reize vor.

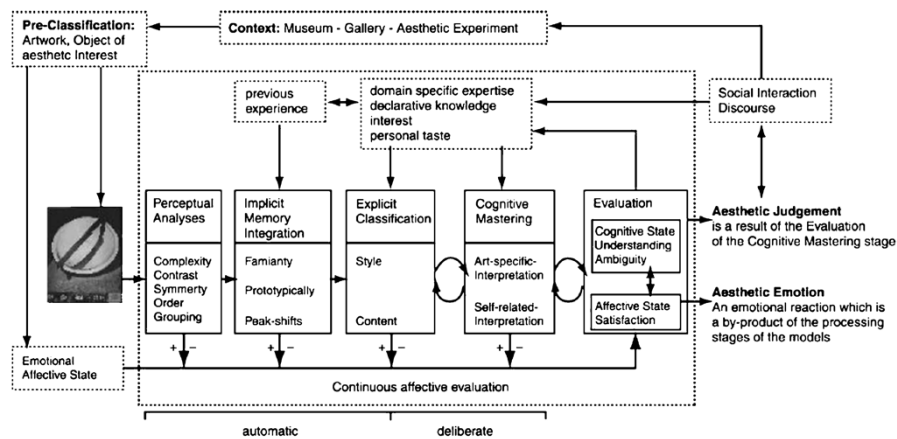


Abbildung 99: Prozessmodell der ästhetischen Erfahrung (Leder, Belke, Oeberst und Augustin, 2004). Erläuterung im Text.

Die ersten beiden Stufen sind automatisiert und einer bewussten Kontrolle nicht zugänglich, während die Verarbeitung ab der dritten Stufe zunehmend vom Subjekt gesteuert wird und auf Expertise beruht. Die Stufen bilden keine strenge Hierarchie. Der Informationsfluss verläuft nicht rein seriell von niedrigen zu höheren Stufen, sondern die Verarbeitung kann in „Schleifen“ auch auf niedrigere Stufen „zurückfallen“. Das Modell unterscheidet zwei separate Ausgaben (Outputs): das *ästhetische Urteil* und die *ästhetische Emotion*. Die Verarbeitung ästhetischer Reize führt sowohl zu kognitiven als auch zu affektiven Bewusstseinszuständen. Zum besseren Verständnis sollen die fünf Stufen des Modells im Einzelnen vorgestellt werden.

Die erste Stufe (perceptual analyses) umfasst einfache sensorische Analysen, d. h. die von Vorerfahrungen noch (weitgehend) unabhängige Verarbeitung des proximalen Stimulus. Als Stimuluseigenschaften, die die Analyse auf dieser Stufe beeinflussen, werden von den Autoren vor allem Komplexität, Symmetrie, Kontrast, Ordnung und Gruppierung genannt – fünf Variablen, denen in der Ästhetikforschung immer schon eine hohe Bedeutung zugesprochen wurde (vgl. z. B. Berlyne, 1970, 1974). Die Verarbeitung auf dieser ersten Stufe erfolgt innerhalb von wenigen

Millisekunden. Da noch kein Abgleich mit Gedächtnisinhalten stattfindet, ist an dieser Stelle ein Erkennen oder Wiedererkennen von Objekten nicht möglich. Vielmehr werden die Strukturmerkmale einer visuellen Szene analysiert und zueinander in Beziehung gesetzt. Beteiligt sind solche Areale des Gehirns, die für die Analyse visueller Informationen verantwortlich sind und sich im Okzipitallappen befinden.

Auf der zweiten Stufe (implicit memory integration) erfolgt eine „implizite“, d. h. nicht notwendigerweise bewusste Verarbeitung auf der Basis von Gedächtnisspuren. Variablen, die auf dieser Stufe eine wichtige Rolle spielen, sind zum Beispiel Familiarität und Prototypikalität. Familiarität meint den Bekanntheitsgrad bzw. die individuelle Vertrautheit mit einem distalen Reiz. Prototypikalität bezeichnet das Ausmaß, in dem ein Reiz repräsentativ für eine Kategorie von Reizen ist. So ist z. B. eine „Madonna“ von Raffael für die meisten Betrachter*innen ein typischeres Beispiel für eine „Madonna“ als eine Darstellung desselben Themas von El Greco. Bereits auf dieser zweiten Verarbeitungsstufe spielt die Kunstexpertise eine nicht unbedeutende Rolle. So werden Expert*innen ästhetische Stimuli aufgrund umfangreicherer Erfahrungen zu einem größeren Anteil als „familiär“ beurteilen und ihnen werden mehr und andere Kategorien zur Verfügung stehen als Lai*innen, was zu differenzierteren Vergleichen und in der Folge auch zu differenzierteren Prototypen führt. Darüber hinaus ist aus Sicht von Leder et al. (2004) auf dieser Verarbeitungsstufe der von Ramachandran und Hirstein (1999) vorgeschlagene Peak-Shift-Effekt relevant.

Auf der dritten Stufe des Modells (explicit classification) wird der ästhetische Stimulus zum ersten Mal als zusammengehöriges Objekt erkannt und als solches mit mentalen Klassifikationssystemen verglichen. Er ist dem Bewusstsein jetzt zugänglich und kann benannt werden. Stil und Inhalt eines Kunstwerks werden jetzt relevant und dienen als Basis für die Klassifikation: Hintergrundwissen, Interesse und persönlicher Geschmack

sind die Grundlage, auf der das Kunstwerk bewertet wird. Können Kunstrezipient*innen eine erfolgreiche Klassifikation vornehmen, wirkt dies nach Leder et al. (2004) belohnend und trägt damit unmittelbar zum weiteren Bestreben der Rezipient*innen bei, an diesem Objekt ästhetische Erfahrungen zu suchen.

Unterschiede zwischen Expert*innen und Lai*innen werden hier noch deutlicher. Berücksichtigen Lai*innen für eine Klassifikation hauptsächlich inhaltliche Gesichtspunkte, also die Frage, was in einem Kunstwerk dargestellt wird, sind für die Exper*innen auch Kategorien relevant, die spezifisch für den Bereich der Kunst sind, etwa das Wissen über die Künstlerin oder den Künstler, die Technik des Kunstwerks und die kunsthistorische Einordnung. Auch ist nach Leder et al. (2004) das Prinzip der Generalisierung auf dieser Verarbeitungsstufe anzusiedeln: Die Betrachterin oder der Betrachter kann die bisherigen Erfahrungen mit Kunst auf neue, unbekannte Objekte generalisieren. Dieser Vorgang ist mit Lust oder Vergnügen verbunden.

Die vierte und fünfte Stufe des Modells (cognitive mastering und evaluation) bilden, wie schon die beiden vorherigen Stufen, eine Feedback-Schleife (vgl. Abbildung 90). Das Ergebnis der vierten Verarbeitungsstufe wird auf der Evaluationsstufe des Modells fortwährend bewertet. Diese permanente Evaluation wiederum beeinflusst den Verarbeitungsprozess auf Stufe vier. Die Stufe der Evaluation hat im Wesentlichen die Aufgabe, eine Erfolgsabschätzung des durch die vorherige vierte Stufe gelieferten Verarbeitungsergebnisses vorzunehmen. Wird sie als nicht erfolgreich wahrgenommen, kann über Feedback-Schleifen eine erneute Verarbeitung auf den Stufen zwei bis vier veranlasst werden. Erfolgskriterium ist ein zufriedenstellendes Verständnis (entspricht einem erfolgreichen „kognitiven Mastering“ als einer Möglichkeit neben einer Veränderung im Ambiguitätsniveau). Angenommen wird, dass die Wahrnehmung von Mehrdeutigkeiten zu einer intensiveren Beschäftigung mit einem

Kunstwerk anregt, mit dem Ziel, die Mehrdeutigkeiten zu reduzieren (vgl. auch Kebeck, 2006).

Parallel zur Bewertung des kognitiven Zustandes erfolgt in Stufe fünf auch eine Evaluation der affektiven Reaktionen in Bezug auf den ästhetischen Stimulus. Der affektive Zustand kann durch das Ergebnis jeder vorherigen Verarbeitungsstufe beeinflusst werden. Eine bewusste Wahrnehmung des affektiven Zustandes ist bereits während der Verarbeitung des Kunstwerks auf einer der „niedrigeren“ Verarbeitungsstufen möglich. Diese explizite Trennung von kognitiver und affektiver Verarbeitung spiegelt sich letztlich auch in den beiden separaten „Ausgängen“ des Modells „ästhetisches Urteil“ und „ästhetische Emotion“ wider. Nach Ansicht von Leder et al. (2004) sind diese als relativ unabhängig voneinander zu betrachten: Das Ergebnis der kognitiven Bewertung beeinflusst das Ergebnis der affektiven Bewertung nicht und vice versa. Ein Kunstwerk kann als angenehm oder lustvoll erlebt werden, einer kognitiven Bewertung jedoch nicht standhalten.

Betrachten wir auch dieses mehrstufige Modell anhand des Beispiels, das bereits zur Veranschaulichung der anderen Modelle gewählt wurde: Die Studentin weiß, dass sie sich in einem Museum befindet. Sie erkennt die Statue als Kunstwerk; das ermöglicht die Verarbeitung der visuellen Eindrücke unter ästhetischen Gesichtspunkten. Zuerst ist ihr Gehirn mit der perzeptuellen Analyse von Kontrast, Symmetrie, Komplexität etc. beschäftigt. Auf der nächsten Stufe werden Familiarität und Prototypikalität der Darstellung auf Basis bisheriger Seherfahrungen beurteilt. Hat die Betrachterin ähnliche Statuen bereits zuvor gesehen? Entspricht die Darstellung einem im Gedächtnis gespeicherten Prototypen? Anschließend erfolgt eine explizite und bewusste Klassifikation des Kunstwerks anhand von Stil und Inhalt der Darstellung. Hier hilft der Studentin ihre Expertise: Sie erkennt Herkules und Hydra. Sie schätzt die Entstehung aufgrund der Proportionen, der Bewegtheit der

Darstellung und anderer Merkmale auf das 3. Jahrhundert vor Christus und ordnet die Statue dem Hellenismus zu. Das Bewusstsein, über ausreichend Wissen zu verfügen, um eine solche Einschätzung vornehmen zu können, löst bei ihr ein positives Gefühl aus. Auf einer vierten Analysestufe folgen kunstspezifische und selbstbezogene Interpretationen, die einer konstanten Evaluation unterliegen. Die Studentin spielt gedanklich verschiedene andere Versionen des Themas durch und betrachtet Aufbau und Dynamik der Szene. Nachdem sie verschiedene Möglichkeiten abgewogen hat, kommt sie zu einem vorläufigen Urteil. Das Gefühl, das Kunstwerk besser zu verstehen, intensiviert ihren positiven Gefühlszustand. Als Ergebnis ihrer kognitiven Bemühungen kommt die Studentin zu dem Schluss, dass es sich bei der Statue um ein ansprechendes, aber keinesfalls außergewöhnliches Beispiel hellenistischer Kunst handelt. Trotzdem gefällt ihr die Skulptur gut. Ästhetisches Urteil und ästhetische Emotion stehen hier nicht in Übereinstimmung.

Das Prozessmodell der ästhetischen Erfahrung von Leder et al. (2004) ist ein Rahmenmodell für die Forschung. Es integriert unterschiedliche Forschungsperspektiven. Der stufenweise und hierarchische Aufbau der fünf Verarbeitungsschritte weist auf die Vorstellung von Modulen und Verarbeitungspfaden in der kognitiven Neurowissenschaft, die Berücksichtigung von Einflussfaktoren wie „Kontext“ oder „soziale Interaktion“ lässt auf sozialwissenschaftliche Einflüsse schließen. Eine Vielzahl von Stimulusfaktoren, die sich in der Ästhetikforschung bereits als relevant erwiesen haben (Komplexität, Prototypikalität, Gruppierung etc.) wird integriert. Insgesamt bietet das Modell einen differenzierten begrifflichen Rahmen für die aktuelle Diskussion innerhalb der empirischen Ästhetikforschung.

Ein Rahmenmodell ist als Ganzes einer empirischen Prüfung nicht zugänglich. Auch in dem Modell von Leder et al. (2004) sind die meisten Abläufe innerhalb der Stufen und die Interaktionen zwischen den Stufen

hierfür nicht detailliert genug spezifiziert. Die Abstraktheit der Stufen erschwert es, eindeutige Gegenbeispiele zu finden bzw. ihre spezifischen Merkmale empirisch zu überprüfen. Auch werden nur wenige Angaben gemacht, welche Hirnareale an der ästhetischen Bewertung beteiligt sein sollen, obwohl die fünf zentralen Stufen des Modells eine modularisierte neuronale Verarbeitung nahelegen. Das Modell von Leder et al. (2004) erlaubt bisher eher beschreibende als erklärende oder prädiktive Aussagen. Hier leistet es einen wichtigen Beitrag zur Systematisierung und Integration der aktuellen Forschung.

4. Wahrnehmung von Kunstwerken

Die Suche nach den Gesetzmäßigkeiten ästhetischen Erlebens ist nicht identisch mit der Frage nach der Wahrnehmung von Kunstwerken. Die von Fechner ausgehende Experimentelle Ästhetik hat bisher vorwiegend die Lust-Unlust-Dimension ästhetischer Empfindungen in den Vordergrund gestellt. Oft wurde auch diese noch reduziert auf ein Gefallen oder Missfallen. Die Frage sinnlicher Erkenntnis hat dagegen kaum eine Rolle gespielt. Unter einem Primat der Methode galt es als zu schwierig, den Erkenntnisvorgang selbst zum Gegenstand der Untersuchung zu machen. Das hat der empirischen Ästhetik von Seiten der philosophischen Ästhetik schon früh den Vorwurf eingebracht, sich nur mit den Akzidenzien der ästhetischen Erfahrung zu beschäftigen (Hartmann, 1886, 1924). Diese Einschränkung ist aus heutiger Sicht nicht mehr notwendig. Sowohl die veränderten Möglichkeiten der Stimulusmanipulation als auch die Verwendung neurophysiologischer Methoden zur Abbildung der zentralnervösen Aktivität liefern neue Ansatzpunkte. Die sinnliche Erkenntnis kann selbst zum Forschungsgegenstand werden und mit rationaler Erkenntnis verglichen werden. Auch wenn ein großer Teil der Theorien und Experimente zur ästhetischen Wahrnehmung sich explizit auf Kunstwerke als Stimuli bezieht, so sind sie aus Sicht der Wahrnehmungsforschung doch ein Spezialfall, der gesonderter

Überlegungen bedarf (für Parallelen zwischen der Tätigkeit von Künstler*innen und Forscher*innen vgl. Spillmann, 2007; zur Frage, inwieweit unser Gehirn „künstlerische Kreativität“ aufweist, vgl. Zeki, 2001).

Auch bei der Betrachtung von Kunstwerken ist die Umgebungswahrnehmung zunächst die Referenz. Ihre Vorgaben sind a priori Wahrscheinlichkeiten, denen die Betrachterin oder der Betrachter sich nicht entziehen kann. Die automatischen Gliederungs- und Gruppierungsprozesse können nicht unterlaufen werden (zu ihrer Anwendung im Bereich der bildenden Kunst vgl. Koenderink, 2015; Kebeck, 2019). Der Versuch einer konsistenten Interpretation der Stimulation mithilfe der für die Umgebungswahrnehmung entstandenen Heuristiken steht am Beginn des Betrachtungsprozesses. Während der Betrachtung aber gehören Faktoren wie Schnelligkeit und Eindeutigkeit nicht zu den zentralen Anforderungen. Es besteht eine *evolutionäre Differenz* zwischen dem Sehsystem und dem Vorgang der Betrachtung. Das visuelle System wurde nicht für die Wahrnehmung von Kunstwerken geschaffen und es hat sich nicht unter diesem Aspekt entwickelt. Es ist auf die Beobachtung von Veränderungen ausgerichtet, nicht auf die Betrachtung eines Zustandes. Angepasst ist es an Situationen, in denen aus den Veränderungen schnell und zuverlässig eine stabile Welt erstellt werden muss. In der Umgebungswahrnehmung sind wir Beobachter*innen und Akteur*innen, keine Betrachter*innen.

Ausgerichtet auf ein Wiedererkennen von bereits Bekanntem, unterläuft das visuelle System die aktuelle Stimulation. Es arbeitet eher mit Vorhersagen, was zu sehen sein sollte. Die Suche nach Stabilität beschleunigt das Sehen, während für den Vorgang der Betrachtung eine Verlangsamung gefordert wäre. Auf diese Weise verhindern die Voreinstellungen des Sehsystems eine intensivere Auseinandersetzung mit dem Kunstwerk. Der Sehprozess wird frühzeitig abgebrochen. Die für

ästhetische Erfahrungen kennzeichnende Veränderung des Zeiterlebens bleibt aus. Die Außenzeit steht nicht still. Die Ambiguitäten des Kunstwerks werden nicht registriert. Es besteht ein Antagonismus zwischen dem Kunstwerk und seiner Wahrnehmung.

Damit aus der Beobachtung eine Betrachtung wird, muss das Sehen eine neue Qualität erlangen. Ein Zentrierungswechsel ist erforderlich. Die Bereitschaft der Betrachterin oder des Betrachters, sich vom Kunstwerk anregen zu lassen. Der Impuls geht vom Werk aus und ist Voraussetzung für eine ästhetische Erfahrung. Ein Prozess, der nicht durch willentliche Anstrengung erzwungen werden kann. Ein Wechsel in der Distanz oder in der Ansicht kann hilfreich sein. Jedes Werk hat seine eigenen Abstände und Perspektiven. Es holt die Betrachterin oder den Betrachter zu sich heran oder distanziert ihn. Es gibt ihm einen Standort vor oder legt eine Bewegung und die Richtung dieser Bewegung nahe. Nicht alle Orte im Umfeld einer künstlerischen Arbeit erzeugen dieselbe Spannung.

Neurophysiologisch könnte diese Form der Spannung mit dem bereits erwähnten Zustand der *adaptiven Resonanz* korrespondieren. Das Konzept der Resonanz ist uns aus der Musik vertraut und meint das verstärkte Mitschwingen eines Körpers in Abhängigkeit von der Schwingung eines anderen Körpers. Voraussetzung für das Entstehen von Resonanzfrequenzen ist eine zeitlich veränderliche Einwirkung auf das System. Übertragen auf die Wahrnehmung bildender Kunst: Das Kunstwerk dient als Anregung. Im Prozess der Betrachtung verändert sich die Stimulation und beim der Betrachterin oder dem Betrachter entwickelt sich eine Resonanz. Zwar bleibt der distale Reiz – das Kunstwerk – konstant, nicht aber die neuronale Aktivität, die die Grundlage des Sehens bildet. Die Beschreibung des Betrachtungsvorgangs mithilfe der Resonanzmetapher gilt zunächst dem Kunstwerk. Es öffnet einen Raum für Wahrnehmungen und Reflexionen. Die Betrachterin oder der Betrachter ist kein passiver Resonanzkörper. Im Rahmen eines längeren

Betrachtungsvorgangs kommt es durch permanente Augenbewegungen und ständige Fixationswechsel immer wieder zu neuen Gliederungen und Gruppierungen, der Basis für die sich verstärkende Wirkung der Resonanz. Konstitutive Voraussetzungen sind die zeitliche Erstreckung des Vorgangs und die Bereitschaft der Betrachterin oder des Betrachters, sich vom Bild anregen zu lassen. Gerade die Ambiguitäten des Bildes schaffen günstige Voraussetzungen für eine Resonanzverstärkung. Die Intensität der ästhetischen Erfahrung korrespondiert mit der Qualität des Resonanzvorgangs.

Das Kunstwerk spielt – notwendigerweise – mit der evolutionären Differenz. Dieses Spiel ist gleichzeitig ein Spiel mit den Analyse- und Konstruktionsprinzipien der visuellen Welt. Kunstwerke lösen die Dinge aus ihrer funktionalen Gebundenheit und machen sie durch neue Gliederungen und Zusammenfassungen der Anschauung zugänglich. Dabei unterlaufen sie die Konstruktionsregeln, bringen sie in einen Konflikt und fordern die Betrachterin oder den Betrachter heraus. Ist die neue Ordnung überzeugend, so formt sie das nachfolgende Sehen. Das Besondere dieser Herausforderung besteht darin, dass die ästhetischen Eigenschaften der Objekte im Mittelpunkt stehen. Während es in der Umgebungswahrnehmung zwingend ist, eine Eindeutigkeit der Interpretation zu erreichen, bilden bei einer ästhetischen Erfahrung gerade die Ambiguitäten den Ausgangspunkt. Bei Kunstwerken sind die Möglichkeiten der Betrachterin oder des Betrachters hinsichtlich der Ambiguitätsreduktion begrenzt. Mehrdeutigkeiten sind intendiert und können nicht aufgelöst werden. Ambiguität erlaubt die Erfahrung eines Zugleichs und korrespondiert eher mit unseren Empfindungen als mit unserem logischen Denken. Etwas kann gleichzeitig schön und traurig sein, erschreckend und anziehend. Ambiguität stellt bestehende Kategorien infrage.

Auf seiner Suche nach Eindeutigkeit kann das evolutionär geprägte Sehen überlistet werden. Bedingt durch die Augenbewegungen sind auch bei einem stationären Objekt und einer stationären Beobachterin oder einem stationären Beobachter die Stimulationen nie identisch. Bei längerer oder wiederholter Betrachtung eines Kunstwerks treten andere Aspekte in den Vordergrund als beim ersten Eindruck. Nicht selten kommt es zu neuen Gliederungen und Gruppierungen, zu einer neuen Interpretation. Da der Informationswert an Bedeutung verliert, gewinnen die formalen und ästhetischen Aspekte an Gewicht. Der Akt des Sehens wird bewusst. Vorfreude, aber auch Enttäuschung, sind Gefühle, die der längeren oder mehrfachen Betrachtung vorbehalten sind. Vergleiche zu anderen Werken werden wahrscheinlich. Die wiederholte Betrachtung hat eine eigene Qualität.

Da das visuelle System keine dauerhafte Mehrdeutigkeit akzeptiert, werden die Ambiguitäten zum Gegenstand von Reflexionsprozessen. Es kommt zu einem produktiven Wechsel von Anschauung und Reflexion. Eine neue Ganzheit entsteht, die uns eine plausible Verknüpfung von zunächst Widersprüchlichem zeigt. Die Reflexion vergleicht die neue Erfahrung mit unserem Wissen und ist bemüht, sie auf eine Begrifflichkeit zu beziehen. Gelingt dies, entsteht eine Einheit von Anschauung und Begriff. Den Prozess von besonderer Qualität hat Rudolf Arnheim (1977) als *anschauliches Denken* bezeichnet. Die Betrachterin oder der Betrachter wechselt beständig zwischen dem Kunstwerk und seiner Imagination. Ein Vorgang, der Vergnügen bereitet und den Betrachtungsprozess offen gestaltet. Die Ambiguitäten des Sichtbaren bilden den idealen Ausgangspunkt für visuelle Andeutungen, für einen unbemerkten Wechsel von der Wahrnehmung zur Fantasie.

Glossar

Absorptionsspektrum. Im Zusammenhang mit Sehpigmenten; relatives Ausmaß der Absorption von Licht in einem Sehpigment in Abhängigkeit von der Wellenlänge des Lichtes; unterschiedlich für die einzelnen Photorezeptoren der Netzhaut.

Adaptation. Anpassung eines Sinnessystems an die Umgebungsbedingungen.

Akkommodation (Verschiebung des Brennpunkts). Die Fähigkeit des Auges, sich durch eine Veränderung der Krümmung der Linse auf ein Objekt, welches sich in einer beliebigen Entfernung befindet, so einzustellen, dass dieses Objekt auf der Netzhaut scharf abgebildet wird.

Aktionspotential. Rasche und vorübergehende Änderung des Membranpotentials einer Nerven- oder Muskelzelle nach dem „Alles-oder-Nichts-Gesetz“. Pflanzte sich entlang der Nervenfasern fort und leitet Information weiter.

Algorithmus. Eine aus endlich vielen Schritten bestehende, eindeutige Handlungsanweisung zur Lösung einer Problemstellung.

Amodale Figuren. (s. amodale Konturen)

Amodale Konturen. Amodale Kanten. Konturen oder Kanten, die als vollständige Konturen oder Kanten gesehen werden, obwohl die Stimulation unvollständig ist.

Augenbewegungen. Kommen im Alltag ständig in Form von Konvergenz- und Divergenz-, Folge- und sakkadischen Bewegungen vor. Konvergenzbewegungen meinen ein Einwärtsdrehen der Augen bei Fixation eines sich nähernden Objektes, Divergenz meint die Gegenbewegung bei sich entfernenden Objekten. Folgebewegungen sind relativ langsame, parallele Bewegungen beider Augen, wenn z. B. ein fahrendes Auto beobachtet wird, ohne den Kopf zu bewegen. Für sakkadische Augenbewegungen siehe Sakkaden.

Augenbewegungspfade. Registrierung von Augenbewegungen. Verbindung der Punkte, die beim Betrachten nacheinander fixiert werden. Gesteuert durch Reizmerkmale und Intentionen der Betrachterin oder des Betrachters.

Ausfüllfunktion (Filling-in). Sorgt dafür, dass fehlende Informationen eines Netzhautabbildes, wie beispielsweise durch den blinden Fleck erzeugt, ausgefüllt werden. Ein ähnlicher Mechanismus erfolgt bei der Herstellung eines stabilen Netzhautabbildes. Nach kurzer Zeit nimmt das fixierte Bild Farbe und Helligkeit der Umgebung an (Filling-in).

Bewegungsparallaxe. Eine Informationsquelle für die Wahrnehmung räumlicher Tiefe. Wenn sich (a) eine Beobachterin oder ein Beobachter

bewegt oder wenn sich (b) Objekte im Gesichtsfeld bewegen, verschieben sich die zugehörigen Abbildungen auf der Netzhaut. Das Ausmaß dieser Verschiebung ist abhängig von der Entfernung: Je näher ein Objekt, umso größer ist die Verschiebung. Diese wichtige Information fehlt bei der Wahrnehmung von Bildern.

Bildgebende Verfahren. Verfahren, die Form und Erregung neuronaler Strukturen sichtbar machen, z. B. funktionale Magnetresonanztomographie (fMRT).

Binokularität (binokulares Sehen). Meint im Gegensatz zum monokularen Sehen das Sehen mit beiden Augen. Ist für die räumliche Wahrnehmung bedeutsam.

Bipolare Zellen (Bipolarzellen). Nervenzellen in der Retina. Der Name bezieht sich auf ihren bipolaren Aufbau. Ihre Aufgabe ist es, die Informationen der lichtempfindlichen Fotorezeptoren zu sammeln, zu gewichten und an die Ganglienzellen weiterzuleiten.

Bottom-up-Verarbeitung. Von den proximalen Reizen ausgehende „aufwärts“ gerichtete Verarbeitung, bei der zuerst kleine Einheiten wie Elementarmerkmale analysiert werden. Das Gegenstück ist die Top-down-Verarbeitung.

Computationales Modell. Ein am Computer erstelltes, mathematisches Modell, das komplexe, oft nichtlineare Zusammenhänge zwischen mehreren Parametern simuliert. Hier: Simulation der visuellen Informationsverarbeitung.

Center-Surround-Zellen. Spezieller Typ von Neuronen, die entweder ein exzitatorisches oder ein inhibitorisches Zentrum beziehungsweise Umfeld aufweisen. Die resultierenden, differenzierten Erregungsmuster dienen der Kontrastbildung.

Corpus geniculatum laterale (CGL). Ein Kern im Thalamus, der Informationen über den Sehnerv von der Retina erhält und von dem aus Fasern zum primären visuellen Cortex (V1) ziehen.

Cornea. Glasklarer, von Tränenflüssigkeit benetzter, nach außen gewölbter Teil der Augenhaut, welcher bei der Lichtbrechung eine wichtige Rolle spielt.

Cortex. Die größte Gehirnstruktur beim Menschen. Der Cortex umschließt die Oberfläche und die Seiten des Gehirns und besteht aus mehreren Schichten. Der Cortex hat sehr großen Anteil an der sensorischen Informationsverarbeitung.

Detektoren. Sind hypothetische neuronale Elemente, die auf Unterschiede der Luminanz reagieren. Detektoren für Ecken, Linien und deren Orientierungen sind im primären visuellen Cortex (V1) zu finden.

Dioptrischer Apparat. Eine zusammenfassende Bezeichnung für alle lichtbrechenden Elemente des Auges.

Disparität (binokulare Disparität). Beruht auf dem Abstand zwischen dem linken und dem rechten Auge, wodurch auf der jeweiligen Netzhaut leicht unterschiedliche Abbilder entstehen. Dieser Unterschied bildet einen wichtigen Tiefenhinweis für die räumliche Wahrnehmung.

Distaler Stimulus (distaler Reiz). Meint in Abgrenzung zum proximalen Stimulus die Gegenstandsmerkmale.

Dynamische Stimulation. Beschreibt den Umstand, dass sich das Abbild der Außenwelt auf der Netzhaut ständig ändert, da sich sowohl Beobachterin oder Beobachter als auch die Umgebung im Regelfall in Bewegung befinden.

Einfache Zellen. Die einfachen Zellen des primären visuellen Cortex (V1) reagieren optimal auf Linien einer bestimmten Richtungsorientierung.

Erstskizze (Marr). David Marr versuchte am Computer die menschliche Wahrnehmung zu simulieren. In einem Bild lassen sich Regionen identifizieren, die signifikante Helligkeitsunterschiede aufweisen. Findet man diese Veränderungen bei einer Reihe von benachbarten Regionen, kann daraus eine Linie konstruiert werden. Durch Kombination über verschieden grobe Skalierungen (Filter) hinweg kommt es zur Erstellung einer Erstskizze.

Farbkonstanz. Relative Unabhängigkeit der wahrgenommenen Farbe einer Oberfläche von der Wellenlängenverteilung der Beleuchtung, sodass sich die Wahrnehmung der Farbe gewöhnlich wenig ändert, wenn sich die Art der Beleuchtung ändert.

Farbmischung. Es wird eine additive und eine subtraktive Farbmischung unterschieden. Additive Farbmischung liegt vor, wenn sich Leuchtfelder verschiedener Wellenlängen überlagern (die Wellenlängen addieren sich). Subtraktive Farbmischung hingegen besteht, wenn Malfarben verschiedener Farbpigmente gemischt werden (die Pigmente wirken wie Filter auf verschiedene Wellenlängen).

Farbton. Die Wahrnehmung einer chromatischen Farbe wie Rot, Grün, Blau oder Gelb bzw. von Kombinationen dieser Farben unabhängig von ihrer Helligkeit. Abhängig von den Wellenlängen, die das Auge erreichen.

Figur-Grund-Unterscheidung. Grundlegende Aufteilung eines visuellen Musters in eine Figur und einen Grund. Die Figur wird als vor dem Grund liegend wahrgenommen.

Fixationen. Zeitintervalle, in denen das Auge relativ stillsteht, während die Beobachterin oder der Beobachter ein Bild oder einen Umweltausschnitt anblickt. In diesen Pausen, die Bruchteile von Sekunden dauern, entnimmt die Beobachterin oder der Beobachter dem Bild oder dem Umweltausschnitt die Informationen. Die Fixationen

müssen nicht immer gleich lang sein. Die Fixationsdauer bezeichnet die Länge eines solchen Zeitintervalls.

Fließmuster (optisches). Die Veränderung des Abbildungsmusters auf der Netzhaut, das entsteht, wenn (a) eine Beobachterin oder ein Beobachter sich bewegt (dann verschiebt sich das gesamte Fließmuster) oder wenn (b) sich Objekte im Gesichtsfeld bewegen (dann verschieben sich lediglich Teile). Eine Vorwärtsbewegung erzeugt ein sich ausdehnendes Muster, eine Rückwärtsbewegung ein sich zusammenziehendes.

Formkonstanz. Relative Unabhängigkeit der wahrgenommenen Form eines Objektes von Veränderungen des Netzhautabbildes, z. B. bei einer Neigung oder Drehung des Objektes.

Fourieranalyse. Ein mathematisches Verfahren, um jede beliebige periodische Schwingung als eine Überlagerung von Sinusschwingungen unterschiedlicher Amplituden und Frequenzen darzustellen; wichtiges methodisches Werkzeug der Wahrnehmungsforschung.

Fovea centralis. Ein kleiner Bereich der Netzhaut, der nur Zapfen enthält und der Ort des schärfsten Sehens ist. Die Fovea liegt auf der Sichtlinie, sodass immer dann, wenn eine Person ein Objekt direkt anblickt, dessen Bild (oder ein Teil des Bildes) auf die Fovea fällt. Der Bereich des fovealen Sehens ist sehr klein. Er umfasst nur ca. einen Sehwinkelgrad, was bei ausgestrecktem Arm der Größe des Daumennagels entspricht.

Magnetresonanztomographie, funktionale (fMRT). Eine Aufnahmetechnik, die Hirnaktivitäten misst und lokalisiert, ohne invasiv in das Gehirn einzugreifen. Das Gehirn wird in ein starkes magnetisches Feld gebracht und mit Radiowellen stimuliert, sodass die Moleküle Energie ausstoßen, die zur dreidimensionalen Rekonstruktion des Gehirns genutzt werden kann.

Ganzfeld, homogenes. Das gesamte Gesichtsfeld ist von einer einheitlichen Stimulation ausgefüllt. Es gibt keine Inhomogenitäten, sodass die Netzhaut trotz der Sakkadensprünge immer gleichmäßig gereizt wird.

Gegenfarbencellen. Ein Neuron, das auf Wellenlängen eines Teils des Spektrums mit Erregung antwortet und auf Wellenlängen eines anderen Teils des Spektrums mit Hemmung. Eine doppelte Gegenfarbencelle antwortet im On-Zentrum auf einen Wellenlängenbereich mit Erregung und auf die Wellenlängen der Gegenfarbe mit Hemmung. Für das Umfeld ergibt sich das entgegengesetzte Muster.

Gesichtsfeld. Das Gesichtsfeld umfasst alles, was ohne Bewegung der Augen, des Kopfes oder des Rumpfes gesehen werden kann.

Größenkonstanz. Auch wenn bei größerer Entfernung eines Objektes oder einer Person das zugehörige Netzhautbild kleiner wird, wird doch z.

B. eine sich entfernende Person nicht als kleiner wahrgenommen, sondern nur als „weiter entfernt“.

Gruppierungs- und Gestaltgesetze. Eine Reihe grundlegender, von den Gestaltpsychologen in der ersten Hälfte des 20. Jahrhunderts beschriebener Gesetzmäßigkeiten, die angeben, unter welchen Aspekten einzelne Elemente zu größeren und bedeutungsvollen Einheiten zusammengefasst werden (z. B. Nähe, Gleichheit, glatter Verlauf usw.).

Helligkeit. Begriff, der sich auf die Wahrnehmung der Intensität des reflektierten Lichtes unabhängig von der Wellenlänge bezieht. Der Eindruck umfasst die gesamten Graustufen (Weiß bis Schwarz).

Helligkeitskonstanz. Relative Unabhängigkeit der wahrgenommenen Helligkeit einer Oberfläche von der Beleuchtungsintensität; wichtige Voraussetzung für das Bildlicht.

Heuristik. Vorgehensweise, die das Auffinden von Lösungen zu Problemen ermöglichen soll, zu denen kein mit Sicherheit zum Erfolg führender Algorithmus bekannt ist.

Homomorphie. Zwei Abbildungen sind dann homomorph, wenn jedem Element von A ein Element von B eindeutig zugeordnet werden kann (in Abgrenzung zur Isomorphie, bei der die Abbildung auch in umgekehrter Richtung eindeutig sein muss).

Hyperkomplexe Zellen (endinhibierte Zellen). Reagieren optimal auf Streifen, Ecken oder Winkel mit einer bestimmten Ausrichtung und einer festgelegten Bewegungsrichtung.

Hypersäulen. Eine Zusammenfassung von Zellen des visuellen Cortex mit etwa einem Millimeter Kantenlänge, die eine Positionssäule für einen bestimmten Bereich der Netzhaut, je eine Säule für Augendominanz des rechten und des linken Auges und einen vollständigen Satz von Orientierungssäulen umfasst. Eine Hypersäule kann man sich als Verarbeitungseinheit für einen kleinen Bereich der Netzhaut vorstellen.

Induktion (induktives Schließen). Nicht deterministisches, schlussfolgerndes Denken, dessen Ergebnis mit Unsicherheit behaftet ist.

Inferenzprozesse. Schlussfolgernde Denkprozesse.

Inferotemporaler Cortex (IT). Das inferotemporale Areal erhält Informationen aus dem visuellen Cortex (V4) und ist Teil der ventralen „Was-Bahn“. Im inferotemporalen Areal werden u. a. Informationen über die Form von wahrgenommenen Objekten verarbeitet.

Invarianz. Reizverhältnisse höherer Ordnung, die auch bei einer Bewegung der Beobachterin und des Beobachters oder einer Veränderung der Beleuchtungsbedingungen konstant bleiben (s. Helligkeits-, Farb-, Größen- und Formkonstanz).

Isoluminanz. Von einer Oberfläche wird an allen Punkten dieselbe Menge oder Intensität an Licht ausgestrahlt bzw. reflektiert, sodass keine Helligkeitsunterschiede wahrgenommen werden können.

Kanonische Ansichten. Von jedem Gegenstand gibt es einige Ansichten, die gewohnter oder charakteristischer sind als andere. Die kanonischen Ansichten werden möglicherweise im Gedächtnis bevorzugt gespeichert. Aus diesem Grund ist die Zeit verkürzt, die zum Erkennen und Benennen eines Objekts benötigt wird, das in einer kanonischen Ansicht gezeigt wird.

Kategorienselektivität. Annahme, dass einige Neuronenverbände im visuellen Cortex spezifisch auf eine Kategorie von Objekten reagieren.

Komplexe Zellen. Die komplexen Zellen des primären visuellen Cortex reagieren optimal auf Linien mit einer bestimmten Ausrichtung und einer bestimmten Bewegungsrichtung.

Konstanzphänomene. (s. Helligkeits-, Farb-, Größen und Formkonstanz). Bestehen für eine Vielzahl von visuellen Wahrnehmungsqualitäten und sorgen dafür, dass ein Objekt unter wechselnden Bedingungen stabil wahrgenommen werden kann.

Konvergenz, neuronale. Mehrere Rezeptorzellen werden in der weiteren Verarbeitung auf ein einzelnes Neuron verschaltet. Die Signale summieren sich und es kommt leichter zu einem Überschreiten der „Alles-oder-Nichts“-Schwelle.

Kontrastbildung. Prozess des visuellen Systems, bei dem bestehende Luminanzunterschiede im visuellen Feld verstärkt werden, um Oberflächen einfacher voneinander abgrenzen zu können.

Laterale Hemmung. Hemmung, die sich in einem neuronalen Schaltkreis seitlich ausbreitet. In der Netzhaut sind die Horizontal- und Amakrinzellen am Vorgang der lateralen Hemmung beteiligt. Sie bewirkt eine Verstärkung der bestehenden Reizunterschiede und ist eine wesentliche Grundlage für die weitere Informationsverarbeitung.

Linse. Konvex gekrümmter, elastischer, kristallklarer Körper, der das durch die Pupille eintretende Licht auf der Netzhaut bündelt, sodass ein scharfes Bild entsteht.

Luminanz (Leuchtdichte). Maß der Menge oder Intensität, mit dem sichtbares Licht von einer Oberfläche ausgestrahlt oder reflektiert wird.

Magnifikation. Prozess, bei dem sich mit zunehmender Annäherung eines Objektes der Sehwinkel und damit die Abbildung des Objektes auf der Retina vergrößert.

Magno-System („Wo-Bahn“). In einem vereinfachten Funktionsmodell des visuellen Systems u. a. für die Objektlokalisierung zuständige Verbindung zwischen der Area striata und dem Parietallappen. Auch dorsale Bahn genannt.

Mediotemporales Areal (MT). Das mediotemporale Areal erhält Informationen aus dem visuellen Cortex und ist Teil der dorsalen „Wo-Bahn“. Die Informationen werden an den parietalen Cortex weitergeleitet. Im mediotemporalen Areal werden u. a. Bewegungsinformationen verarbeitet.

Minifikation. Prozess, bei dem sich mit zunehmender Entfernung der Sehwinkel und damit die Abbildung des Objektes auf der Retina verkleinert.

Monokulares Sehen. Sehen mit einem Auge.

Monokulare Tiefenhinweise. Um Hinweise über die dritte Dimension zu erhalten, nutzt das visuelle System verschiedene Informationen. Die monokularen sind besonders bei der Bildbetrachtung wichtig. Dazu gehören: Abstand zur Horizontlinie, atmosphärische Perspektive, Licht und Schatten, relative Größe, Texturgradient, Verdeckung.

Multiperspektivität. Fähigkeit des visuellen Systems, verschiedene Ansichten eines Objekts zu erstellen, sodass es aus unterschiedlichen Perspektiven identifiziert werden kann.

Multistabilität. Auf der Basis eines distalen Stimulus (z. B. eines Bildes) gibt es zwei oder mehr Perzepte, die in sich konsistent sind, sich aber in der Priorität abwechseln.

Neuron. Nervenzelle: grundlegende Einheit des Nervensystems, deren Funktion darin besteht, Informationen zu übertragen und zu speichern. Umfasst den Zellkörper mit den Erbinformationen und die Fortsätze (Dendriten), die die Signale zum Zellkörper leiten, sowie ein Axon, in dem das Signal von der Zelle aus weitergeleitet wird.

Ökologische Optik. Theorie von J. J. Gibson zum Verständnis der Leistungen des visuellen Systems.

Okulomotorik. Fähigkeit zur Drehung des Auges, um die horizontale, vertikale und sagittale Achse.

Ontogenese. Individuelle Entwicklungsgeschichte eines einzelnen Lebewesens von der befruchteten Eizelle bis zum Tod.

Optische Anordnung. Ergibt sich aus der Struktur des von den Oberflächen reflektierten Lichts und ist für einen Standort spezifisch. Alle Sehwinkel oder Raumwinkel zusammen bilden die optische Anordnung.

Optisches Fließmuster. (s. Fließmuster).

Parietaler Cortex (PT). Der parietale Cortex schließt nach frontal an den okzipitalen Cortex an. Er ist Teil der dorsalen „Wo-Bahn“ und erhält seine Informationen u. a. aus dem mediotemporalen Areal.

Parvo-System („Was-Bahn“). In einem vereinfachten Funktionsmodell des visuellen Systems für die Objekterkennung zuständige Verbindung zwischen Area striata und Temporallappen. Auch ventrale Bahn genannt.

Peripheres Sehen. Gegenstück zum fovealen Sehen. Als Peripherie wird der Bereich der Netzhaut bezeichnet, der sich um die Fovea centralis und die angrenzende Parafovea herum befindet. Einen Großteil der Dinge, die wir wahrnehmen, sehen wir peripher. Hier ist die Abbildung unscharf und Farben können nur bei größeren Flächen erkannt werden.

Perzept. Ergebnis des Wahrnehmungsvorganges.

Photopisches Sehen. Sehen bei Tageslicht oder hellem Kunstlicht. In diesem Falle sind die Zapfen aktiv, sie ermöglichen das Farbsehen.

Phylognese. Stammesgeschichtliche Entwicklung der Gesamtheit aller Lebewesen einer Art.

Positronen-Emissions-Tomographie (PET). Ein bildgebendes Verfahren, das Areale erhöhten Zuckerstoffwechsels mithilfe von radioaktiv markierten Zuckermolekülen dreidimensional darstellen kann. Es eignet sich zur Untersuchung von aktiven Bereichen im Gehirn.

Prägnanz. Annahme, dass die Gesamtorganisation eines visuellen Feldes nach einem Maximum an Einfachheit und Ordnung strebt (genauer: Prägnanztendenz).

Primärer visueller Cortex (Primäres visuelles Areal, V1). Für die visuelle Informationsverarbeitung wichtiger Bereich im Hinterhauptslappen, in dem die von der Netzhaut über das CGL kommende Information weiterverarbeitet wird. Der primäre visuelle Cortex ist retinotop organisiert.

Proximaler Stimulus, Proximaler Reiz. Meint in Gegenüberstellung zum distalen Stimulus die Reizverhältnisse am Sinnesorgan (auf der Netzhaut).

Pupille. Kreisförmige Öffnung der Iris im Auge, durch welche Licht in die hintere Augenkammer dringen kann. Durch Verkleinerung und Vergrößerung der Pupille kann das Ausmaß des Lichteinfalls reguliert werden.

Querdisparation. Dreidimensionale Objekte mit ihrer räumlichen Erstreckung erzeugen auf den Netzhäuten der beiden Augen leicht unterschiedliche Abbildungen. Diese Winkeldifferenz wird als Querdisparation bezeichnet. Sie ist abhängig vom aktuellen Fixationspunkt.

Refraktion. Brechkraft der Linse, mit deren Hilfe ein scharfes Bild auf der Netzhaut erzeugt werden kann.

Repräsentation. Neben dem „Erleben“ (Perzept, Vorstellungsbild) ist es sinnvoll, die mentale Repräsentation als ein theoretisches Konstrukt zu unterscheiden. Gemeint sind in der Kognitionsforschung zumeist Annahmen über die zugrundeliegende Art der Speicherung von Informationen im Gedächtnis.

Retina, Netzhaut. Das von den Oberflächen reflektierte Licht fällt durch die Linse in unser Auge und trifft auf die Retina. Sie bedeckt fast die gesamte Rückwand des Auges und hat einen Schichtenaufbau. In der untersten Schicht befinden sich die Photorezeptoren (Zapfen und Stäbchen), die für eine Informationsumwandlung der Lichtreize in elektrische Potentiale sorgen.

Retinotopie, retinotope Organisation. Meint eine neuronale Kartierung. Sie gibt die Entsprechung von Orten an, die auf der Netzhaut nebeneinander liegen und sich auch auf weiteren Verarbeitungsstufen (z. B. CGL und primärer visueller Cortex) in unmittelbarer Nachbarschaft befinden.

Rezeptives Feld. Das rezeptive Feld eines Neurons ist derjenige Bereich auf der Rezeptorfläche (beim Sehen ist es die Retina), der bei einer Reizung die Entladungsrate dieses Neurons beeinflusst.

Rezeptoren, Photorezeptoren. Ein sensorischer Rezeptor ist ein Neuron, das für Signale aus der Umwelt empfindlich ist und diese physikalischen Signale in bioelektrische Signale umwandelt. Im visuellen System nehmen die Zapfen und Stäbchen die Rolle der Rezeptoren ein. Da sie auf Lichteinfall reagieren, werden sie auch Photorezeptoren genannt.

Rezeptorpotential. Wird eine Rezeptorzelle gereizt, entsteht eine länger anhaltende Spannungsänderung, deren Größe in einem direkten Zusammenhang zur Stärke des Reizes steht, aber keinen Einfluss auf die Größe der Amplitude des nachfolgenden Aktionspotentials hat, möglicherweise aber auf die Anzahl der Aktionspotentiale.

Sakkaden. Sakkadische Augenbewegungen sind schnelle und sprunghafte Bewegungen der Augen zwischen zwei Fixationen. Sie finden ständig statt. Sie können willkürlich ausgelöst werden, haben aber einen ballistischen Charakter (ihr Zielpunkt ist vorab nicht genau zu bestimmen).

Sakkadische Suppression. Bezeichnet die Unterdrückung der Informationsaufnahme während, kurz vor und nach einer Sakkade (s. Sakkaden).

Sättigung. Meint mit Bezug auf die Farbwahrnehmung den relativen Anteil von Weiß in einer chromatischen Farbe. Je weniger Weiß eine Farbe enthält, desto gesättigter ist sie.

Schatten. Eine Region relativer Dunkelheit innerhalb einer beleuchteten Region. Er entsteht, wenn ein Objekt ganz oder teilweise das Licht verdeckt. In Abhängigkeit von der Lichtquelle werden verschiedenen Arten von Schatten unterschieden, z. B. Kern- und Halbschatten.

Sehbahn. Weg der neuronalen Verbindungen, über die die bioelektrischen Signale aus der Retina zum CGL und dann zum primären visuellen Cortex gelangen.

Simultankontrast. Veränderung der wahrgenommenen Helligkeit und/oder Farbe einer Oberfläche durch umgebende Helligkeiten oder Farben. Der Simultankontrast ist am größten, wenn die Fläche von einer andersfarbigen Fläche umschlossen wird.

Sklera. Ledrige Außenhaut des Auges, die das Auge gegen äußere Einflüsse abschirmt.

Skotopisches Sehen. Sehen bei Nacht oder geringer Beleuchtung. In diesem Fall sind besonders die Stäbchen aktiv. Sie sind für das Schwarz-Weiß-Sehen zuständig. Wellenlängenunterschiede und damit Farben können von ihnen nicht wahrgenommen werden.

Spektrum. Elektromagnetisches Spektrum, sichtbares Licht. Das visuelle System des Menschen kann nur Licht bestimmter Wellenlängen wahrnehmen. Dieser Bereich reicht von ca. 400–700 Nanometer.

Stäbchen. Eine der beiden Photorezeptor-Arten der Retina. Wird aufgrund ihres Aussehens als Stäbchen bezeichnet. Das Stäbchensystem ist im Dunkeln besonders empfindlich, kann aber keine feinen Einzelheiten oder Details auflösen (skotopisches Sehen).

Strahlungslicht. Licht, welches von einer Energiequelle ausgesendet wird (nach Gibson).

Subjektives Augengrau. Perzept, das einige Sekunden nach der künstlichen Herstellung eines stabilen Netzhautbildes entsteht. Es weist keinerlei Strukturierung auf.

Textur. Jede Oberfläche weist eine gewisse Struktur auf, die für sie charakteristisch ist. Sie beeinflusst so, wie auftreffendes Licht reflektiert wird. Texturen spielen besonders bei der Abgrenzung von Oberflächen und bei der Tiefenwahrnehmung eine große Rolle.

Texturdichte. Die Textur einer Oberfläche kann von sehr fein bis sehr grob sein. Die Texturdichte gibt die Anzahl der Elemente pro Flächeneinheit an.

Texturgradient. Das Muster, das von einer regelmäßig texturierten Fläche bei seiner Abbildung auf der Netzhaut erzeugt wird. Die Elemente eines Texturgradienten werden mit größerer Entfernung immer kleiner und ihre Dichte nimmt zu. Die Information kann für die Wahrnehmung räumlicher Tiefe und Entfernung genutzt werden. Der Verlauf des Gradienten gibt darüber hinaus Auskunft über die Krümmung einer Fläche bzw. ihren Neigungswinkel gegenüber der Beobachterin oder des Beobachters.

Top-down-Verarbeitung. Verarbeitung, die mit der Analyse von Informationen auf den höchsten Stufen beginnt, etwa semantischer Information und die die Verarbeitung neu eingehender Informationen durch bereits vorhandene Information (Gedächtnis) erleichtert (s. im Gegensatz dazu Bottom-up-Verarbeitung).

Trajektorie. Verlauf bzw. Kurve, die sich hier als Resultat einer Differentialgleichung ergibt.

Transduktion. Wandelt den äußeren Reiz (Licht) in ein (analoges) bioelektrisches Signal um.

Transformation. Merkmal der Reizfortleitung, bei der das graduelle Signal der Rezeptoren auf der Retina in ein Aktionspotential („Alles-oder-Nichts-Gesetz“) umgewandelt wird.

Umgebungslicht. Das von Oberflächen reflektierte Licht (nach Gibson).

V4. Das Areal V4 ist Teil des visuellen Cortex. Hier findet u. a. ein wesentlicher Teil der Farbverarbeitung statt.

Visuelle Suche. Es geht um die Frage: Welche Merkmale im Gesichtsfeld werden in welcher Kombination und Reihenfolge im Laufe des Wahrnehmungsprozesses analysiert und wie werden diese Merkmale integriert?

Visuelle Welt. Alles, was mit der visuellen Wahrnehmung (Auge plus Gehirn) erfasst werden kann. Umfasst sowohl stationäre wie sich bewegenden Beobachter*innen. Wesentlich sind hierbei Oberflächen und das von ihnen reflektierte Licht.

Visueller Cortex. Der visuelle Cortex ist die Region des Gehirns, in der die Signale aus der Retina weiterverarbeitet werden. Der Sehnerv läuft, nachdem er das CGL durchquert hat, zum primären visuellen Cortex, V1. Von dort aus werden die Signale immer spezialisierter in höheren Zentren (V2–V4) weiterverarbeitet.

Visuelles Feld. Das visuelle Feld eines Auges ist begrenzt durch die Augenhöhle und die Nase. Unter den Bedingungen der Umgebungswahrnehmung weist es eine typische Gliederung in drei Bereiche auf: (Erd-)Boden, Horizont und Himmel.

Visuelles Gedächtnis. In Mehrspeichermodellen geht man davon aus, dass sich das visuelle Gedächtnis aus funktional unterschiedlichen Teilen zusammensetzt: dem ikonischen Gedächtnis, dem visuellen Kurzzeitgedächtnis (Arbeitsspeicher) und dem Langzeitgedächtnis, welches wiederum aus einem episodischen, prozeduralen und semantischen Gedächtnis besteht.

Zapfen. Photorezeptor in der Netzhaut, der in seiner Form Ähnlichkeit mit einem Zapfen hat und für das Sehen bei Tage und bei heller Beleuchtung zuständig ist (photopisches Sehen). Zapfen befinden sich vor allem in der Nähe der Fovea centralis und sind für das Farbsehen und für das Sehen von Details verantwortlich.

Zentralperspektive. Aus Sicht der Wahrnehmungsforschung ist die Zentralperspektive ein Projektionsverfahren, das den Abbildungsverhältnissen einer monokularen und stationären Beobachterin oder eines Beobachters nahekommt. Es setzt die Festlegung

eines Augenpunktes voraus. Von diesem werden nach dem Modell der Sehpyramide Linien zu den Eckpunkten der abzubildenden Gegenstände gezeichnet. Sie entsprechen den „Sehstrahlen“ und schneiden die Bildfläche an den Punkten, an denen die Eckpunkte der zugehörigen Figuren einzuzeichnen sind.

Zentrum-Umkreis-Filter. Vorschlag von David Marr, wonach in einem kreisförmigen Feld anhand eines Algorithmus unterschiedliche Gewichtungen der einzelnen Bereiche vorgenommen werden, sodass eine Akzentuierung des Zentrums resultiert. Das Zentrum erscheint dann heller als die Umgebung.

Literatur

- Abramov, I. & Gordon, J. (1994). Color appearance: On seeing red, or yellow, or green, or blue. *Annual Review of Psychology*, 45, 451-485.
- Ahlström, U. (1995). Perceptual unit formation in simple motion patterns. *Scandinavian Journal of Psychology*, 36(4), 343-354.
- Alais, D., Blake, R. & Lee, S. H. (1998). Visual features that vary together over time group together over space. *Nature Neuroscience*, 1(2), 160-164.
- Allesch, C. G. (1987). *Geschichte der psychologischen Ästhetik*. Göttingen: Hogrefe.
- Allesch, C. G. (2006). *Einführung in die psychologische Ästhetik*. Wien: WUV.
- Arnheim, R. (1977). *Anschauliches Denken*. Köln: DuMont.
- Arnheim, R. (1978). *Kunst und Sehen. Eine Psychologie des schöpferischen Auges*. Berlin: de Gruyter.
- Arnheim, R. (1980). *Die Dynamik der architektonischen Form*. Köln: DuMont.
- Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, 61, 183-193.
- Attneave, F. (1955). Symmetry, information, and memory for patterns. *The American Journal of Psychology*, 68(2), 209-222.
- Attneave, F. & Arnoult, M. D. (1956). The quantitative study of pattern and shape perception. *Psychological Bulletin*, 53(6), 452-472.
- Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. In W. A. Rosenblith (Hrsg.), *Sensory communication* (S. 217-234). Cambridge: MIT Press.
- Barlow, H. B. & Reeves, B. C. (1979). The versatility and absolute efficiency of detecting mirror symmetry in random dot displays. *Vision research*, 19(7), 783-793.
- Bayes, T. (1763). *An Essay towards solving a problem in the doctrine of chance*. London.
- Beck, J. & Gibson, J. (1955). The relation of apparent shape to apparent slant in the perception of objects. *Journal of Experimental Psychology*, 50, 125-133.
- Beck, D. M. & Palmer, S. E. (2002). Top-down influences on perceptual grouping. *Journal of Experimental Psychology. Human Perception and Performance*, 28(5), 1071-1084.
- Berkley, G. (1987). *Versuch über eine neue Theorie des Sehens*. (Philosophische Bibliothek) Hamburg: Felix Meiner. (Original 1709).
- Berlyne, D. E. (1970). Novelty, complexity and hedonic value. *Perception and Psychophysics*, 8, 279-286.

- Berlyne, D. E. (1971). *Aesthetics and psychobiology*. New York: Appleton-Century-Crofts.
- Berlyne, D. E. (1974). *Studies in the new experimental aesthetics. Steps toward an objective psychology of aesthetic appreciation*. New York: Wiley.
- Biederman, I. (1985). Human image understanding: Recent research and a theory. *Computer Vision, Graphics, and Image Understanding*, 32, 29-73.
- Biederman, I. (1987). Recognition-by-components: A theory of human image understanding. *Psychological Review*, 94(2), 115-117.
- Birkhoff, G. D. (1928). Quelques éléments mathématiques de l'art. In G. D. Birkhoff (Hrsg.), *Collected mathematical papers* (S. 288-306); Vol. 3. New York: Dover Publishers.
- Birkhoff, G. (1933). *Aesthetic measure*. Cambridge: Harvard University Press.
- Blake, R. & Shiffrar, M. (2007). Perception of Human Motion. *Annual Review of Psychology*, 5, 847-873.
- Bohil, C. J., Bradly, A. & Biocca, A. (2011). Virtual reality in neuroscience research and therapy. *Nature Reviews. Neuroscience*, Dec., 752-762.
- Boring, E. & Holway, A. (1940). The moon illusion and the angle of regard. *American Journal of Psychology*, 53, 109-116.
- Börjesson, E. & Ahlström, U. (1993). Motion structure in five-dot patterns as a determinant of perceptual grouping. *Perception & Psychophysics*, 53(1), 2-12.
- Brooks, J. L. (2015). Traditional and new principles of perceptual grouping. In J. Wagemans (Hrsg.), *The Oxford Handbook of Perceptual Organisation* (S. 57-87). Oxford: Oxford University Press.
- Bülhoff, H., Edelman, S. & Tarr, M. (1994). *How are three-dimensional objects represented in the brain?* CogSci Memo No. 5, Max-Planck-Institut für biologische Kybernetik. Tübingen.
- Caputo, G. (1998). Texture brightness filling-in. *Vision Research*, 38(6), 841-851.
- Casati, R. (2001). *Die Entdeckung des Schattens*. Berlin: Berlin Verlag.
- Casco, C., Campana, G., Han, S. & Guzzon, D. (2009). Psychophysical and electrophysiological evidence of independent facilitation by collinearity and similarity in texture grouping and segmentation. *Vision Research*, 49(6), 583-593.
- Castelo-Branco, M., Goebel, R., Neuenschwander, S. & Singer, W. (2000). Neural synchrony correlates with surface segregation rules. *Nature*, 405(6787), 685-689.
- Chalupa, L. M. & Werner, J.S. (2004). (Hrsg.). *The Visual Neuroscience*. Vol. 1. & Vol. 2. Cambridge: MIT Press.

- Coren, S., Porac, C. & Ward, L. M. (1984). *Sensation and perception*. Orlando: Academic Press.
- Dempster, A. P. (1967). Upper and lower probabilities induced by a multivalued mapping. *The Annals of Mathematical Statistics*, 38(2), 325-339.
- De Weerd, P., Desimone, R. & Ungerleider, L. (1998). Perceptual filling-in: A parametric study. *Vision Research*, 38, 2721-2734.
- Duchowski, A. (2007). *Eye Tracking Methodology. Theory and Practice*. 2nd Edition. London: Springer.
- Duncker, K. (1941). On pleasure, emotion and striving. *Philosophical and Phenomenological Research*, 1, 391-430.
- Ecco, U. (2004). *Die Geschichte der Schönheit*. München: Hanser Verlag.
- Ecco, U. (2007). *Die Geschichte der Häßlichkeit*. München: Hanser Verlag.
- Ehrenfels, C. (1890). Über Gestaltqualitäten. *Vierteljahresschrift für wissenschaftliche Philosophie*, XIV(3).
- Ehrenstein, W. (1941). Über Abwandlungen der Hermannschen Helligkeitserscheinung. *Zeitschrift für Psychologie*, 150, 83-91.
- Eisenman, R. (1968). Semantic differential ratings of polygons varying in complexity-simplicity and symmetry-asymmetry. *Perceptual and Motor Skills*, 26, 1243-1248.
- Eisenman, R. & Rappaport, J. (1967). Complexity preference and semantic differential ratings of complexity-simplicity and symmetry-asymmetry. *Psychonomic Science*, 7(4), 147-148.
- Elder, J. H. & Goldberg, R. M. (2002). Ecological statistics of Gestalt laws for the perceptual organization of contours. *Journal of Vision*, 2(4), 324-353.
- Eysenck, H. J. (1942). The experimental study of the 'good gestalt': A new approach. *Psychological Review*, 49, 345-364.
- Fantoni, C. & Gerbino, W. (2003). Contour interpolation by vector-field combination. *Journal of Vision*, 3(4), 281-303.
- Farid, H. (2002). Temporal synchrony in perceptual grouping: a critique. *Trends in Cognitive Sciences*, 6(7), 284-288.
- Farid, H. & Adelson, E. H. (2001). Synchrony does not promote grouping in temporally structured displays. *Nature Neuroscience*, 4(9), 875-876.
- Fechner, G.T. (1865). Über die Frage des goldenen Schnittes. *Archiv für die zeichnenden Künste*, 11, 100-112.
- Fechner, G. T. (1871). Zur experimentalen Ästhetik. *Abhandlungen der Königlich Sächsischen Gesellschaft der Wissenschaften*, 9, 555-635. (Auch: beigegeben zu G. Th. Fechner, Vorschule der Aesthetik. Nachdruck Hildesheim: Olms, 1978).

- Fechner, G. T. (1876). *Vorschule der Aesthetik*. Nachdruck der 3. Aufl. 1925, Hildesheim: Olms, 1978.
- Field, D. J., Hayes, A. & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local "association field". *Vision Research*, 33(2), 173-93.
- Frith, C. (2010). *Wie unser Gehirn die Welt erschafft*. Heidelberg: Spektrum Akademischer Verlag.
- Garner, W. (1974). *The processing of information and structure*. Potomac: Erlbaum.
- Garner, W. R. & Clement, D. E. (1963). Goodness of pattern and pattern uncertainty. *Journal of Verbal Learning and Verbal Behavior*, 2, 446-452.
- Gegenfurtner, K. R. (2001). *Color vision from genes to perception*. Cambridge: Cambridge University Press.
- Gegenfurtner, K. R. & Kiper, D. (2003). Color Vision. *Annual Review of Neuroscience*, 26, 181-206.
- Gegenfurtner, K. R. & Rieger, J. (2000). Sensory and cognitive contributions of color to the perception of natural scenes. *Current Biology*, 10, 805-808.
- Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology*, 59, 167-192.
- Gelb, A. (1929). Die Farbkonstanz der Sehdinge. In A. Bethe (Hrsg.), *Handbuch der normalen und pathologischen Physiologie* (S. 549-678). Berlin: Springer.
- Gibson, J. J. (1950). *The perception of the visual world*. Boston: Houghton Mifflin.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. New York: Houghton Mifflin.
- Gibson, J. J. (1979). *The ecological approach to visual perception*. Hillsdale: Erlbaum.
- Giese, M. A. & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nature Reviews Neuroscience*, 4(3), 179-192.
- Goldstein, E. B. (2002). *Wahrnehmungspsychologie*. Berlin: Spektrum.
- Gouras, P. (1991). Color Vision. In E. Kandel, J. Schwartz & T. Jessell (Hrsg.), *Principals of neural science* (S. 467-480). New York: Elsevier.
- Graumann, C. F. (1960). *Grundlagen einer Phänomenologie der Perspektivität*. Berlin: Walther De Gruyter.
- Grossberg, S. (1994). 3-D vision and figure-ground separation by visual Cortex. *Perception & Psychophysics*, 55(1), 48-120.
- Grossberg, S. (1997). Cortical dynamics of three-dimensional figure-ground perception of two-dimensional pictures. *Psychological Review*, 104(3), 618-658.

- Gunzenhäuser, R. (1975). *Maß und Information als ästhetische Kategorien*. Baden-Baden: Agis-Verlag (2.Aufl.).
- Guttman, S. E., Gilroy, L. A. & Blake, R. (2007). Spatial grouping in human vision: Temporal structure trumps temporal synchrony. *Vision Research*, 47(2), 219-230.
- Han, S., Song, Y., Ding, Y., Yund, E. W. & Woods, D. L. (2001). Neural substrates for visual perceptual grouping in humans. *Psychophysiology*, 38(6), 926-935.
- Hanson, H. M. (1959). Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology*, 58, 321-334.
- Hardin, C. L. & Maffi, L. (1997). *Color categories in thought and language*. Cambridge: Cambridge University Press.
- Hartmann v., E. (1886). *Die deutsche Ästhetik seit Kant*. Berlin: Ducker.
- Hartmann v., E. (1924). *Philosophie des Schönen* (2. Aufl.). Berlin: Wegweiser-Verlag.
- Hekkert, P. (2006). Design aesthetics: Principles of pleasure in design. *Psychology Science*, 48(2), 157-172.
- Hekkert, P. & van Wieringen, P. C. W. (1990). Complexity and prototypicality as determinants of the appraisal of cubist paintings. *British Journal of Psychobiology*, 81, 483-495.
- Helmholtz, H. v. (1876). Optisches über Malerei. In Herrmann v. Helmholtz. *Vorträge und Reden*. Bd. 2. Braunschweig: Vieweg & Sohn.
- Henderson, J. M. & Hollingworth, A. (1999). High-level scene perception. *Annual Review of Psychology*, 50, 243-271.
- Hess, R. F., May, K. A. & Dumoulin, S. O. (2015). Contour integration: Psychophysical, neurophysiological and computational perspectives. In J. Wagemans (Hrsg.), *Oxford Handbook of Perceptual Organization* (S. 189-206). Oxford: Oxford University Press.
- Hochberg, J. (1971). Perception I: Color and Shape. In J. W. Kling & L. A. Riggs (Hrsg.), *Woodworth & Schlossberg's Experimental Psychology*. New York: Holt, Rinehart & Winston.
- Hochberg, J. & Brooks, V. (1960). The psychophysics of form: Reversible perspective drawings of spatial objects. *American Journal of Psychology*, 73, 337-354.
- Hochberg, J. & McAlister, E. (1953). A quantitative approach to figure 'goodness'. *Journal of Experimental Psychology*, 46, 361-364.
- Hoffman, D. D. (2001). Mereology of visual form. *Proceedings of the Fourth International Workshop on Visual Form*, Capri, Italy, 40-50.
- Hoyer, P. O. & Hyvärinen, A. (2002). A multi-layer sparse coding network learns contour coding from natural images. *Vision Research*, 42, 1593-1605.

- Hsiao, S. W. & Chou, J. R. (2006). A gestalt-like perceptual measure for home page design using a fuzzy entropy approach. *International Journal of Human-Computer Studies*, 64(2), 137-156.
- Hume, D. (1777). *Enquiries Concerning the Human Understanding*. Oxford: Clarendon Press.
- Hummel, J. & Biedermann, I. (1992). Dynamic binding in a neural network for shape recognition. *Psychological Review*, 99(3), 480-517.
- Humphreys, G. W. & Riddoch, M. J. (1993). Interactions between object and space systems revealed through neuropsychology. In D. E. Meyer & S. Kornblum (Hrsg.), *Attention and Performance (Volume 24)* (S. 183-218). Cambridge, MA: MIT Press.
- Hüppe, A. (1984). *Prägnanz – ein gestalttheoretischer Grundbegriff*. München: Profil.
- Hurvich, L. M. (1981). *Color vision*. Sunderland: Sinauer.
- Hyman, J. (2006, Januar). *Art and neuroscience*. Paper presented at the Symposium on Art and Cognition. Heruntergeladen von www.interdisciplines.org/artcognition/papers/15
- Jacobsen, T. (2004). Individual and group modelling of aesthetic judgment strategies. *British Journal of Psychology*, 95, 41-56.
- Jacobsen, T., Schubotz, R. I., Höfel, L. & von Cramon, D. Y. (2006). Brain correlates of aesthetic judgment of beauty. *NeuroImage*, 29, 276-285.
- Jenkins, H. M. & Harrison, R. H. (Hrsg.) (1962). Generalization gradients of inhibition following auditory discrimination learning. *Journal of the Experimental Analysis of Behavior*, 5, 435-441.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14, 201-211.
- Johansson, G. (1986). Visuelle Bewegungswahrnehmung. In *Wahrnehmung und visuelles System*. Heidelberg: Spektrum der Wissenschaft.
- Kaminski, G. & Immelmann, K. (1987). Die ökologische Perspektive. In Deutsches Institut für Fernstudien an der Universität Tübingen (Hrsg.), *Funkkolleg Psychobiologie. Verhalten bei Mensch und Tier. Studienbegleitbrief* (S. 57-119). Weinheim: Beltz.
- Kanizsa, G. (1955). Marzini quasi-perscettivi in campi con stimolazione omogenea. *Rivistia di Psicologica*, 49(1), 7-30.
- Kanizsa, G. (1979). *Organization in vision. Essays on gestalt perception*. New York: Praeger.
- Kawabata, H. & Zeki, S. (2004). Neural correlates of beauty. *Journal of Neurophysiology*, 91, 1699-1705.
- Kebeck, G. (1991). *Fehleranalyse als Methode der Gedächtnisforschung*. Wiesbaden: Deutscher Universitäts-Verlag.

- Kebeck, G. (1994). *Wahrnehmung. Theorien, Methoden und Forschungsergebnisse der Wahrnehmungspsychologie*. Weinheim: Juventa.
- Kebeck, G. (2005). *Bild und Betrachter. Auf der Suche nach Eindeutigkeit*. Regensburg: Schnell & Steiner.
- Kebeck, G. (2018). Die Ordnung der visuellen Welt. In G. Kebeck & C. Gliese (Hrsg.), *Sehdinge und Schaustücke*. Bielefeld: Kerber.
- Kebeck G. (2019). *Gliederung und Gruppierung*. In G. Kebeck & A. K. Schulze (Hrsg.), *Übersummativität*. Köln: Snoek.
- Kebeck, G. & Schroll, H. (2011). *Experimentelle Ästhetik*. Wien: Facultas.
- Kellman, P. J., Garrigan, P. B., Kalar, D. & Shipley, T. F. (2010). Good continuation and relatability: Related but distinct principles. *Journal of Vision*, 3(9), 120-120.
- Kellman, P. J. & Shipley, T. F. (1991). A theory of visual interpolation in object perception. *Cognitive Psychology*, 23(2), 141-221.
- Kelly, F. & Grossberg, S. (2000). Neural dynamics of 3-D surface perception: Figure-ground separation and lightness perception. *Perception & Psychophysics*, 62(8), 1596-1618.
- Kindy, J. (1999). Of time and beauty. *Journal of Consciousness Studies: Controversies in Science & the Humanities*, 6(6-7), 61-63.
- Klütsch, C. (2007). *Computer Grafik. Ästhetische Experimente zwischen zwei Kulturen. Die Anfänge der Computerkunst in den 1960er Jahren*. Wien: Springer Verlag.
- Kobbert, M. J. (1986). *Kunstpsychologie. Kunstwerk, Künstler und Betrachter*. Darmstadt: Wissenschaftliche Buchgesellschaft.
- Koffka, (1935). *Principles of Gestalt psychology*. London: Kegan Paul.
- Köhler, W. (1920). *Die physischen Gestalten in Ruhe und im stationären Zustand*. Braunschweig: Vieweg & Sohn.
- Köhler, W. (1938). *The place of value in the world of facts*. New York: Liveright. (deutsch: *Werte und Tatsachen*, Berlin: Springer, 1968).
- Konderink, J. J. (2015). Perceptual organization in art. In J. Wagemans (Hrsg.), *The Oxford Handbook of Perceptual Organisation* (S. 886-916). Oxford: Oxford University Press.
- Kraft, J. & Brainard, D. (1999). Mechanisms of color constancy under natural viewing. *Proc. Nat. Acad. Sci. USA*, 96(1), 307-312.
- Kubovy, M. (1979). How Probable and Useful are Subjective Probability and Utility? *Psyc critiques*, 24(3), 174-175.
- Kubovy, M. (1986). *The psychology of perspective and Renaissance art*. Cambridge: Cambridge University Press.
- Landwehr, K. (1986). Die ökologische Auffüllung der Welt I – Ein Vergleich der Prinzipien der Analyse optischer Stimulus-Information in der Gestalttheorie und in der ökologischen Optik. *Gestalt Theory*, 8, 186-203.

- Landwehr, K. (1988). Die ökologische Auffüllung der Welt II – Homogenitäts-Inhomogenitäts-Übergänge im Ganzfeld. *Gestalt Theory*, 10, 21-34.
- Langlois, J. H. & Roggman, L. A. (1990). Attractive faces are only average. *Psychological Science*, 1, 115-121.
- Leder, H., Belke, B., Oeberst, A. & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgements. *British Journal of Psychology*, 95, 489-508.
- Lee, S. H. & Blake, R. (1999). Visual form created solely from temporal structure. *Science*, 284(5417), 1165-1168.
- Leeuwenberg, E. L. J. (1971). A perceptual coding language for visual and auditory patterns. *American Journal of Psychology*, 84, 307-349.
- Leeuwenberg, E. L. J. (1978). Quantification of certain visual pattern properties: Saliency, transparency, and similarity. In E. L. J. Leeuwenberg & H. F. J. M. Buffart (Hrsg.), *Formal theories of visual perception* (S. 277-298). New York: Wiley.
- Leibowitz, H. & Bourne, L. (1956). Time and intensity as determiners of perceived shape. *Journal of Experimental Psychology*, 51, 277-281.
- Lennie, P. (2003). The cost of cortical computation. *Current Biology*, 13, 493-497.
- Levinthal, B. R. & Franconeri, S. L. (2011). Common-fate grouping as feature selection. *Psychological Science*, 22(9), 1132-1137.
- Lewin, K. (1936/1969). *Grundzüge der topologischen Psychologie*. Bern: Huber.
- Lindsay, P. & Norman, D. (1977). *Human Information Processing*. New York: Academic Press.
- Livingstone, M. (2002). *Vision and art. The biology of seeing*. New York: Abrams.
- Logethis, N. (2000). Das Sehen - Ein Fenster zum Bewußtsein. *Spektrum der Wissenschaft*, 37-43.
- Lopes, D. (1996). *Understanding Pictures*. Oxford: Clarendon Press.
- Luce, R. D. (2002). A psychophysical theory of intensity proportions, joint presentations, and matches. *Psychological Review*, 109(3), 520-532.
- Mach, E. (1911). *Die Analyse der Empfindungen und das Verhältnis des Physischen zum Psychischen*. Jena: Fischer.
- Mangan, B. (1999). It don't mean a thing if it ain't got that swing. *Journal of Consciousness Studies: Controversies in Science & the Humanities*, 6(6-7), 56-58.
- Marr, D. (1982). *Vision*. San Francisco: Freeman.
- Marr, D. & Poggio, T. (1976). Cooperative computation of stereo disparity. *Science*, 194, 283-287.

- Martindale, C. (1984). The pleasures of thought: A theory of cognitive hedonics. *The Journal of Mind and Behavior*, 5(1), 49-80.
- Martindale, C. (1990). *The clockwork muse: The predictability of artistic change*. New York: Basic Books.
- Martindale, C. (1999). Peak shift, prototypicality and aesthetic preference. *Journal of Consciousness Studies: Controversies in Science & the Humanities*, 6(6-7), 52-54.
- Martindale, C. (2007a). Recent trends in the psychological study of aesthetics, creativity, and the arts. *Empirical Studies of the Art*, 25(2), 121-141.
- Martindale, C. (2007b). A neural-network theory of beauty. In C. Martindale, P. Locher & V. M. Petrov (Hrsg.), *Evolutionary and neurocognitive approaches to aesthetics, creativity and the arts* (S. 181-194). Amityville: Baywood Publishing Company.
- Martindale, C. & Moore, K. (1988). Priming, prototypicality, and preference. *Journal of Experimental Psychology: Human Perception and Performance*, 14(4), 661-670.
- Martindale, C., Moore, K. & Borkum, J. (1990). Aesthetic preference: Anomalous findings of berlyne's psychobiological theory. *American Journal of Psychobiology*, 103, 53-80.
- Martindale, C., Moore, K. & West, A. (1988). Relationship of preference judgments to typicality, novelty, and mere exposure. *Empirical Studies of the Arts*, 6, 79-96.
- McWhinnie, H. J. (1968). A review of research on aesthetic measure. *Acta Psychologica*, 28, 363-375.
- Metelli, F. (1974). The perception of transparency. *Scientific American*, 230, 90-98.
- Metzger, W. (1930). Optische Untersuchungen am Ganzfeld. II. Mitteilung: Zur Phänomenologie des homogenen Ganzfeldes. *Psychologische Forschung*, 13, 6-29.
- Metzger, W. (1965). Der Beitrag der Gestalttheorie zur Frage der Grundlagen des künstlerischen Erlebens. In M. Stadler & H. Crabus (Hrsg.), *Gestaltpsychologie* (S. 497-508). Frankfurt am Main: Kramer.
- Metzger, W. (1967). Figural-Wahrnehmung. In W. Metzger (Hrsg.), *Handbuch der Psychologie (Bd.1). Allgemeine Psychologie*. Göttingen: Hogrefe.
- Metzger, W. (1968). *Psychologie* (4. Aufl.). Darmstadt: Dietrich Steinkopff.
- Metzger, W. (1975). *Gesetze des Sehens*. Frankfurt a. Main: Waldemar Kramer.
- Milman, M. (1984). *Das Trompe-l'oeil*. Genf: Skira.

- Ngo, D. C. L., Samsudin, A. & Abdullah, R. (2000). Aesthetic measures for assessing graphic screens. *Journal of Information Science and Engineering*, 16, 97-116.
- Ngo, D. C. L. & Teo, L. S. (2000). A mathematical theory of interface aesthetics. *Visual mathematics (Online-Journal)*, 2(4).
- Ngo, D. C. L., Teo, L. S. & Byrne, J. G. (2002); Evaluating interface aesthetics. *Knowledge and Information Systems*, 4, 46-79.
- Ngo, D. C. L., Teo, L. S. & Byrne, J. G. (2003). Modelling interface aesthetic. *Information Science*, 152, 25-46.
- Olshausen, B. A. & Field, D. J. (2004). Sparse coding of sensory inputs. *Current Opinion in Neurobiology*, 14, 481-487.
- Paden, A., Cap, M., Yong, S. Z., Yershov, D. & Frazzoli, E. (2016). Survey of motion planning and control techniques for self-driving urban vehicles. *IEEE Transactions on Intelligent Vehicles*, 1(1), 33-55.
- Palmer, S. E. (1982). Symmetry, transformations, and the structure of perceptual systems. In F. Beck (Hrsg.), *Organization and representation in perception*. Hillsdale: Erlbaum.
- Palmer, S. E. (1992). Common region: A new principle of perceptual grouping. *Cognitive Psychology*, 24(3), 436-447.
- Palmer, S. E. (1999). *Vision science. Photons to phenomenology*. Cambridge: MIT Press.
- Palmer, S. E. & Rock, I. (1994). Rethinking perceptual organization: The role of uniform connectedness. *Percept Psychophys*, 67(4), 727-740.
- Paradiso, M. A. & Nakayama, K. (1991). Brightness perception and filling-in. *Vision Research*, 31(7/8), 1221-1236.
- Pelli, D. G. (1999). Close encounters - An artist shows that size affects shape. *Science*, 285, 844-846.
- Perrett, D. I., May, K. A. & Yoshikawa, S. (1994). Facial shape and judgments of female attractiveness. *Nature*, 368, 239-242.
- Pinel, J. (2001). *Biopsychologie*. Berlin: Spektrum.
- Pirenne, M. H. (1970). *Optics, painting, photography*. Cambridge: Cambridge University Press.
- Poggio, T. (1986). Wie Computer und Menschen sehen. In *Wahrnehmung und visuelles System*. Heidelberg: Spektrum.
- Pomerantz, J. R. & Cragan, A. I. (2015). Emergent features and feature combination. In J. Wagemans (Hrsg.), *Oxford Handbook of Perceptual Organization* (S. 88-107). Oxford: Oxford University Press.
- Ramachandran, V. S. (1988). Perception of shape from shading. *Nature*, 331, 163-166.
- Ramachandran, V. S. (1995). 2-D or not 2-D - that is the question. In R. Gregory, J. H. P. Harris & D. Rose (Hrsg.), *The Artful Eye* (S. 249-267). Oxford: Oxford University Press.

- Ramachandran, V. S. (1999). Author's response. *Journal of Consciousness Studies: Controversies in Science & the Humanities*, 6(6-7), 72-75.
- Ramachandran, V. S. & Hirstein, W. (1999). The science of art: A neurological theory of aesthetic experience. *Journal of Consciousness Studies: Controversies in Science & the Humanities*, 6(6-7), 15-51.
- Rausch, E. (1974). Das Eigenschaftsproblem in der Gestalttheorie der Wahrnehmung. In W. Metzger (Hrsg.), *Handbuch der Psychologie* (2. Aufl., Bd. 1, S. 866-953). Göttingen: Hogrefe.
- Rausch, E. (1982). *Bild und Wahrnehmung*. Frankfurt: Waldemar Kramer.
- Reber, R. & Schwarz, N. (1999). Effects of perceptual fluency on judgments of truth. *Consciousness and Cognition*, 8, 338-342.
- Reber, R., Schwarz, N. & Winkielman, P. (2004). Processing fluency and aesthetic pleasure: Is beauty in the perceiver's processing experience. *Personality and Social Psychology Review*, 8, 364-382.
- Redies, C. (2007). A universal model of esthetic perception based on the sensory coding of natural stimuli. *Spatial Vision*, 21, 97-117.
- Redies, C., Hasenstein, J. & Denzler, J. (2007). Fractal-like image statistics in visual art: Similarity to natural scenes. *Spatial Vision*, 21, 137-148.
- Rock, I. (1977). In defense of unconscious inference. In W. Epstein (Hrsg.), *Stability and constancy in visual perception: Mechanisms and processes* (S. 321-373). New York: Wiley.
- Rock, I. (1985). *Wahrnehmung. Vom visuellen Reiz zum Sehen und Erkennen*. Heidelberg: Spektrum.
- Roelfsema, P. R. (2006). Cortical algorithms for perceptual grouping. *Annual Review of Neuroscience*, 29, 203-227.
- Roelfsema, P. R., Lamme, V. a F. & Spekreijse, H. (2004). Synchrony and covariation of firing rates in the primary visual cortex during contour grouping. *Nature Neuroscience*, 7(9), 982-991.
- Rosenholtz, R. (2015). Texture perception. In J. Wagemans (Hrsg.), *Oxford Handbook of Perceptual Organization* (S. 167-186). Oxford: Oxford University Press.
- Roye, A., Höfel, L. & Jacobsen, T. (2008). Aesthetics of faces: Behavioral and electrophysiological indices of evaluative and descriptive judgment processes. *Journal of Psychophysiology*, 22(1), 41-57.
- Rubin, E. (1921). *Visuell wahrgenommene Figuren. Studien in psychologischer Analyse*. København: Gyldendalske Boghandel. (Original 1915).
- Runeson, S. (1977). On the possibility of smart perceptual mechanisms. *Scandinavian Journal of Psychology*, 18(3), 172-179.

- Sekuler, A. B. & Bennett, P. J. (2001). Generalized common fate: Grouping by common luminance changes. *Psychological Science*, 12(6), 437-444.
- Seymour, K., Karnath, H.-O. & Himmelbach, M. (2008). Perceptual grouping in the human brain: Common processing of different cues. *Neuroreport*, 19(18), 1769-1772.
- Shafer, G. (1967). *A Mathematical Theory of Evidence*. Princeton: Princeton University Press.
- Shapiro, A. & Todorovic, D. (2017). (Hrsg.). *The Oxford Compendium of Visual Illusions*. Oxford: Oxford University Press.
- Shepard, R. S. (1990). *Mind Sights*. New York: Freeman.
- Shevell, S. K. & Kingdom, F. A. (2008). Color in complex scenes. *Annual Review of Psychology*, 59, 143-166.
- Shubnikov, A. V. & Koptsik, V. A. (1974). *Symmetry in science and art*. New York: Plenum.
- Silvia, P. J. (2005). Emotional responses to art: From collation and arousal to cognition and emotion. *Review of General Psychology*, 9(4), 342-357.
- Simoncelli, E. P. & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, 24, 1193-1216.
- Singer, W. & Gray, C. M. (1995). Visual feature integration and the temporal correlation hypothesis. *Annual Review of Neuroscience*, 18, 555-586.
- Singh, D. (1993). Adaptive significance of female physical attractiveness: Role of waist-to-hip ratio. *Journal of Personality and Social Psychology*, 65, 293-307.
- Skov, M. & Vartanian, O. (2009). *Neuroaesthetics*. Amityville: Baywood Publishing.
- Spillmann, L. (2007). Artists and vision scientists can learn a lot from each other, but do they? *Gestalt Theory*, 29(1), 13-39.
- Stevens, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural and social prospects*. New York: Wiley.
- Stiller, C., Kammel, S., Lulcheva, I. & Ziegler, J. (2008). Probabilistische Methoden in der Umfeldwahrnehmung kognitiver Automobile. *Automatisierungstechnik*, 56, 563-574.
- Sun, Z., Bebis, G. & Miller, R. (2006). On-road vehicle detection: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(5), 694-711.
- Tarr, M. J. & Warren W. H. (2002). Visual reality in behavioral neuroscience and beyond. *Nature Neuroscience*, 5 Suppl. (Nov.), 1089-1092.

- Thinès, G., Costall, A. & Butterworth, G. (Hrsg.). (1991). *Michotte's experimental phenomenology of perception*. Hillsdale, NJ England: Lawrence Erlbaum.
- Ullman, S. (1979). *The interpretation of visual motion*. Oxford: MIT Press.
- Van der Helm, P. A. (2015a). Simplicity in perceptual organization. In J. Wagemans (Hrsg.), *Oxford Handbook of Perceptual Organization* (S. 1027-1045). Oxford: Oxford University Press.
- Van der Helm, P. A. (2015b). Symmetry Perception. In J. Wagemans (Hrsg.), *Oxford Handbook of Perceptual Organization* (S. 108-128). Oxford: Oxford University Press.
- Vinje, W. E. & Gallant, J. L. (2000). Sparse coding and decorrelation in primary visual cortex during natural vision. *Science*, 287, 1273-1276.
- Vinje, W. E. & Gallant, J. L. (2002). Natural stimulation of the nonclassical receptive field increases information transmission efficiency. *Journal of Neuroscience*, 22, 2904-2915.
- Wagemans, J. (1993). Skewed symmetry: A nonaccidental property used to perceive visual forms. *Journal of Experimental Psychology. Human Perception and Performance*, 19(2), 364-380.
- Wagemans, J. (1995). Detection of visual symmetries. *Spatial Vision*, 9(1), 9-32.
- Wagemans, J. (2015). (Hrsg.). *The Oxford Handbook of Perceptual Organisation*. Oxford: Oxford University Press, 886-916.
- Wallen, R. (1999). Response to Ramachandran and Hirstein. *Journal of Consciousness Studies: Controversies in Science & the Humanities*, 6(6-7), 68-72.
- Werner, J. S. & Chalupa, L. M. (2013). (Hrsg.). *The New Visual Neuroscience*. Cambridge, MA: MIT Press.
- Wertheimer, M. (1912). Experimentelle Studien über das Sehen von Bewegung. *Zeitschrift für Psychologie*, 61, 161-265.
- Wertheimer, M. (1923). Untersuchungen zur Lehre von der Gestalt II. *Psychologische Forschung*, 2(4), 301-350.
- Weyl, H. (1952). *Symmetry*. Princeton: Princeton University Press.
- Whitfield, T. W. & Slatter, P. E. (1979). The effects of categorization and prototypicality on aesthetic choice in a furniture selection task. *British Journal of Psychology*, 70, 65-75.
- Willmore, B. & Tolhurst, D. J. (2001). Characterizing the sparseness of neural codes. *Network: Computational Neural Systems*, 12, 255-270.
- Wouterlood, D. & Boselie, F. (1992). A good-continuation model of some occlusion phenomena. *Psychological Research*, 54(4), 267-277.
- Yarbus, A. L. (1967). *Eye movements and vision*. New York: Plenum Press.

- Zajonc, R. B. (1968). Attitudinal effects of mere exposure. *Journal of Personality and Social Psychology*, 9(2, Pt2), 1-27.
- Zajonc, R. B. (1980). Feeling and thinking: Preferences need no inferences. *American Psychologist*, 35, 151-175.
- Zeki, S. (1983a). Color coding in the cerebral Cortex: The reaction of cells in monkey visual Cortex to wavelengths and colors. *Neuroscience*, 9, 741-765.
- Zeki, S. (1983b). Color coding in the cerebral Cortex: The responses of wavelength-selective and color-coded cells in monkey visual Cortex to changes in wavelength composition. *Neuroscience*, 9, 767-781.
- Zeki, S. (1999). *Inner vision. An exploration of art and the brain*. Oxford: Oxford University Press.
- Zeki, S. (2001). Artistic creativity in the brain. *Science*, 293, 51-52.
- Zimmer, A. C. (1984). Foundations for the measurement of phenomenal symmetry. *Gestalt Theory*, 6, 118-157.
- Zimmer, A. C. (1995). Multistability - More than just a freak phenomenon. In P. Kruse & M. Stadler (Hrsg.), *Ambiguity in Mind and Nature* (S. 99-138). Berlin: Springer.

Abbildungen

Abbildung 1: Vergleich von retinalem Bild und wahrgenommenem Bild. Entnommen am 20.7.2012 von <http://de.wikipedia.org/wiki/Gesichtsfeld> ©Wikimedia Deutschland e.V.

Abbildung 2: Versuchsanordnung zum homogenen Ganzfeld.

Abbildung 3: Visuelles Feld. Gibson, J. J. (1950). *The perception of the visual world*. Abb. 19. ©Houghton Mifflin Company, Boston.

Abbildung 4: Piero della Francesca (ca. 1465). *Geißelung Christi*. ©Palazzo Ducale, Urbino.

Abbildung 5: Gijysbrechts, C. N. (um 1670). Rückseite eines Gemäldes, Öl auf Leinwand, 66,6 x 86,5 cm, © Statens Museum for Kunst, Copenhagen.

Abbildung 6: Natürliche Szenen. Entnommen am 24.7.2013 von <http://www.wikimedia.de/wiki> ©Wikimedia Deutschland e.V.

Abbildung 7: Geisler, W. S. (2008). Natural Scene Statistics. Visual Perception and the Statistical Properties of Natural Scenes. *Annual Review of Psychology*, 59, S. C-1, Abb.1.

Abbildung 8: Elektromagnetisches Spektrum.

Abbildung 9: Texturgradienten (longitudinale und frontale Oberflächen). Gibson, J. J. (1950). *The perception of the visual world*. Abb. 32 und 33. ©Houghton Mifflin Company, Boston.

Abbildung 10: Optische Anordnung. Gibson, J. J. (1979). *The ecological approach to visual perception*. ©Houghton Mifflin Company, Boston.

Abbildung 11: Optisches Fließmuster. Gibson, J. J. (1950). *The perception of the visual world*. Abb. 53 und Abb. 54. ©Houghton Mifflin Company, Boston.

Abbildung 12: Lichteinfall und retinale Reizverarbeitung. Goldstein, E.B. (2008). *Wahrnehmungspsychologie*. ©Spektrum-Verlag, Berlin.

Abbildung 13: Funktionsweise von rezeptiven Feldern.

Abbildung 14: Absorptionsspektren. Neitz, J. & Neitz, M. (2011). Spectral Sensitivities of S, L and M Cones. The genetics of normal and defective color vision. *Vision Research*, 51, S. 637, Abb. 3.

Abbildung 15: Verlauf der Sehbahn. Pinel, J. P. (2001). *Biopsychologie*. Farbtafel XIV. ©Spektrum-Verlag, Berlin.

Abbildung 16: Hypersäulen. Wolfe, J. M., Kluender, K. R. & Levi, D. M. (2009). Model of a Hypercolumn. *Sensation and Perception* (2te Ed.). S. 68, Abb. 3.23. © Sinauer Associates, Sunderland.

Abbildung 17: Schematische Darstellung des Parvo-Systems und des Magno-Systems.

Abbildung 18: Diskrete und kontinuierliche Übergänge.

Abbildung 19: Charakteristische Texturen von Oberflächen. Gibson, J. J. (1979). *The ecological approach to visual perception*. ©Houghton Mifflin Company, Boston.

Abbildung 20: Zwei Texturfelder mit sichtbarer Grenze.

Abbildung 21: Geone als Grundbausteine der Objektwahrnehmung.

Biedermann, I. (1987). Recognition-by components: A theory of human image understanding. *Psychological Review*, 94, 115-147. ©APA Journals, Washington, D.C.

Abbildung 22: Versteckte und unauffindbare Taschenlampe.

Biedermann, I. (1987). Recognition-by components: A theory of human image understanding. *Psychological Review*, 94, 115-147. ©APA Journals, Washington, D.C.

Abbildung 23: Kanonische und nicht-kanonische Ansicht eines Dreirads.

Bülthoff, H. H., Edelman, S. & Tarr, M.J. (1994). How are three-dimensional objects represented in the brain? *CogSci Memo*, No. 5.

©Max-Planck-Institut für biologische Kybernetik, Tübingen.

Abbildung 24: Augenbewegungspfade bei verschiedenen Aufgabenstellungen. Yarbus, A. L. (1967). *Eye movements and vision*.

©Plenum Press, New York.

Abbildung 25: Auswahl von Gruppierungs- und Gestaltgesetzen. Coren, S., Porac, C. & Ward, L. M. (1984). *Sensation and perception*. ©Academic Press, Orlando.

Abbildung 26: Invarianz gegenüber Spiegelungen und Rotationen;

Takahashi, J., Kawachi, Y. & Gyoba, J. (2012). Internal criteria underlying affective responses to visual patterns. *Gestalt Theory*, 34, 67-80.

Abbildung 27: Perspektive.

Abbildung 28: Horizont.

Abbildung 29: Beleuchtung.

Abbildung 30: Planität.

Abbildung 31: Toplogie.

Abbildung 32: Textur.

Abbildung 33: Einfachheit und Prägnanz.

Abbildung 34: Nähe.

Abbildung 35: Gleichheit und Ähnlichkeit.

Abbildung 36: Kollinearität.

Abbildung 37: Geschlossenheit.

Abbildung 38: Aufgehen ohne Rest.

Abbildung 39: Symmetrie.

Abbildung 40: Konnektivität.

Abbildung 41: Gemeinsames Schicksal.

Abbildung 42: Simultanität.

Abbildung 43: Hysterese.

Abbildung 44: Fixationswechsel.

Abbildung 45: Eigenbewegung.

Abbildung 46: Intention.

Abbildung 47: Simultaner Helligkeitskontrast.

Abbildung 48: White'sche-Illusion.

Abbildung 49: Amodales Dreieck. Kanizsa, G. (1955). Marzini quasi-percettivi in campi con stimolazione omogenea. *Rivista di Psicologia*, 49(1), 7-30.

Abbildung 50: Ehrenstein-Figur. Ehrenstein, W. (1941). Über Abwandlungen der L. Hermannschen Helligkeitserscheinung. *Zeitschrift für Psychologie*, 150, 83-91.

Abbildung 51: Phänomenale Transparenz. Kanizsa, G. (1979). *Organization in vision. Essays on gestalt perception*. S. 155, Abb. 9.3a. ©Praeger Publishers, New York.

Abbildung 52: Topologische Voraussetzungen für phänomenale Transparenz. Kanizsa, G. (1979). *Organization in vision. Essays on gestalt perception*. S. 156-157, Abb. 9.3; ©Praeger Publishers, New York.

Abbildung 53: Necker-Würfel.

Abbildung 54: Projektionen des Necker-Würfels. Zimmer, A. C. (1995). Multistability - More than just a freak phenomenon. In Kruse, P. & Stadler, M. (Hrsg.), *Ambiguity in Nature and Mind. Multistability in Cognition*. S. 111, Abb. 11. ©Springer-Verlag, Berlin.

Abbildung 55: Amesche Stühle. Perception Demonstration. Center, Princeton University, ©Princeton University Press.

Abbildung 56: Konstruktion einer Figur nach dem „bump-mapping“ Verfahren.

Abbildung 57: Figurenstudie. Marco Pitteri (a) und Francesco Bartolozzi (b) nach Giovanni Battista Piazzetta. ©Getty Center for the History of Art and Humanities, Santa Monica, California.

Abbildung 58: Mehrdeutige Figuren: Kugeln oder Mulden. Ramachandran, V. (1995). 2-D or not 2-D - that is the question. In R. Gregory, J. H. P. Harris & D. Rose (Hrsg.), *The Artful Eye* (S. 249-267). S.250, Abb. 11.1. ©Oxford University Press, Oxford.

Abbildung 59: Ramachandran, V. S. (1995). (a) Reihe Mulden, Reihe Kugeln. S. 250, Abb. 11.2; (b) Ramachandran, V. S. (1995). Einbettung in eine eindeutige Gesamtsituation. S. 251, Abb. 11.3. In R. Gregory, J. H. P. Harris & D. Rose (Hrsg.), *The Artful Eye* (S. 249-267). ©Oxford University Press, Oxford.

Abbildung 60: (a) Ramachandran, V. S. (1995). Kreise als Kugeln. S. 252, Abb. 11.4 (a); (b) Ramachandran, V. S. (1995). Kreise als Mulden. S. 252, Abb. 11.4 (b). In R. Gregory, J. H. P. Harris & D. Rose (Hrsg.), *The Artful Eye* (S. 249-267). ©Oxford University Press, Oxford.

Abbildung 61: Ramachandran, V. S. (1995). Demonstrationen des Zentrum-Umfeld-Effektes. S. 255, Abb. 11.7, 11.8, 11.9. In R. Gregory, J. H. P. Harris & D. Rose (Hrsg.), *The Artful Eye* (S. 249-267). ©Oxford University Press, Oxford.

Abbildung 62: Cézanne, P. (1890-1894). Die Kalkbrennerei, Aquarell. ©Musée d'Orsay, Paris.

Abbildung 63: Anordnung der Markierungspunkte. Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14, 202.

Abbildung 64: Punktmuster eines tanzenden Paares. Johansson, G. (1986). Visuelle Bewegungswahrnehmung. In: Wahrnehmung und visuelles System. Heidelberg: Spektrum der Wissenschaft.

Abbildung 65: Wahrgenommene Kausalität. Michotte, A., Thinès, G. (Hrsg.), Costall, A. (Hrsg.) & Butterworth, G. (Hrsg.). (1991). *Michotte's experimental phenomenology of perception*. ©Lawrence Erlbaum Associates, Inc., Hillsdale, New England.

Abbildung 66: Pendelproblem.

Abbildung 67: Gedrehte Tische. Shepard, R. S. (1990). *Mind Sights*. S. 48, Abb. A2. ©Freeman, New York.

Abbildung 68: Helligkeitskonstanz in natürlicher Umgebung (Meyerowitz, 1985).

Abbildung 69: (a) Hopper, E. (1963). Sonne in einem leeren Zimmer, Öl auf Leinwand, 63x90 cm, Privatbesitz. (b) und (c) eigene Bearbeitung.

Abbildung 70: Fotos einer Wiesenlandschaft zu drei Tageszeiten. Gegenfurtner, K. & Kiper, D. (2003). Color Vision. *Annual Review of Neuroscience*, 26, 181-206. ©Fischer-Verlag, Frankfurt am Main.

Abbildung 71: Ein Spiel mit der Größenkonstanz. ©Bettman Archive.

Abbildung 72: Formkonstanz. Lindsay, P. & Norman, D. (1977). *Human Information Processing*. ©New York: Academic Press.

Abbildung 73: Nixon vor seinem Wahlkampfplakat. *Time Magazine* (1968, 19. März) [Fotografie von einer Fotografie] ©Time Magazine, New York.

Abbildung 74: Homomorphismus. Palmer, S. E. (1999). *Vision science. Photons to Phenomenology*. © MIT Press, Cambridge, Massachusetts.

Abbildung 64: Stufen der visuellen Informationsverarbeitung. Palmer, S. E. (1999). *Vision science. Photons to phenomenology*. S. 85, Abb. 2.4.1 ©MIT Press, Cambridge, Massachusetts.

Abbildung 75: Umwandlung eines Graufstufenbildes in eine numerische Verteilung. Palmer, S. E. (1999). *Vision science. Photons to phenomenology*. Abb.2.4.2, 2.4.3. ©MIT Press, Cambridge, Massachusetts.

Abbildung 76: Schematische Darstellung der Verarbeitungsschritte bis zur Erstskeizze. Marr, D. (1982). *Vision*. S. 55, Abb. 2-5. ©Freeman, San Francisco.

Abbildung 77: Wirkungsweise des mathematischen Filters. Marr, D. (1982). *Vision*. S. 55, Abb. 2-9. ©Freeman, San Francisco.

Abbildung 78: Bildzerlegung durch mathematische Filter. Marr, D. (1982). *Vision*. S. 69, Abb. 2-20. ©Freeman, San Francisco.

Abbildung 79: Konstruktion von Bildgrenzen durch Texturmerkmale. Marr, D. (1982). *Vision*. S. 94, Abb. 2-34. ©Freeman, San Francisco.

Abbildung 80: Illustration der 2.5 D Skizze. Marr, D. (1982). *Vision*. S. 129, Abb. 3-12. ©Freeman, San Francisco.

Abbildung 81: Schematische Darstellung zu Ullmanns Versuchsaufbau mit rotierendem Zylinder. Marr, D. (1982). *Vision*. S. 206, Abb. 3-52. ©Freeman, San Francisco.

Abbildung 82: Hierarchische Repräsentation einer Form in der dreidimensionalen Beschreibung. Marr, D. (1982). *Vision*. S. 306, Abb. 5-3. ©Freeman, San Francisco.

Abbildung 83: Pfeifenreinigerfiguren. Marr, D. (1982). *Vision*. S. 299, Abb. 5-1. ©Freeman, San Francisco.

Abbildung 84: Extremkletterin in einer Steilwand. Kosicki, G. (1987). In G. Kaminski & K. Immelmann, *Funkkolleg Psychobiologie: Lebensräume: die ökologische Perspektive* © Weinheim: Beltz.

Abbildung 85: Suchbild. Dalmatinerhund. ©R. C. James.

Abbildung 86: Vexierbild. Napoleon auf Elba.

Abbildung 87 Suchbild. Jesusdarstellung [Eigene Bearbeitung]. (a) und (b): Rock, I. (1985). *Wahrnehmung. Vom visuellen Reiz zum Sehen und Erkennen*. © Spektrum Verlag, Heidelberg

Abbildung 88: Durchsichtiges Messer. Kanizsa, G. (1979). *Organization in vision. Essays on Gestalt Perception*. ©Praeger Publishers, New York.

Abbildung 89: Ehrenstein-Figur. Ehrenstein, W. (1941). Über Abwandlungen der L. Hermannschen Helligkeitserscheinung. *Zeitschrift für Psychologie*, 150, 83-91.

Abbildung 90: Beispiel eines CC-Loops. Grossberg, S. (1994). 3-D vision and figure-ground separation by visual Cortex. *Attention, Perception, & Psychophysics*, 55, 76, Abb. 22a.

Abbildung 91: Vergleich von BCS und FCS, 30. Bruchmann, M. (2007). A Feature-Based Approach to Perceptual Texture Filling-In. Dissertation der Johannes-Gutenberg-Universität Mainz.

Abbildung 92: Funktionale Architektur des FACADE-Modells. Kelly, F. & Grossberg, S. (2000). Neural dynamics of 3-D surface perception: Figure-ground separation and lightness perception. *Attention, Perception, & Psychophysics*, 62, 1600, Abb.5.

Abbildung 93: Birkhoff Polygone. Kobbert, M. J. (1986). *Kunstpsychologie. Kunstwerk, Künstler und Betrachter*. S. 15, Abb. 2. © Wissenschaftliche Buchgesellschaft, Darmstadt.

Abbildung 94: Chinesische Vasen und Umrisszeichnungen mit Gitter. Fotografien von Hobson, 1927, Zeichnungen von Birkhoff, 1928; Klütsch, C. (2007). *Computer Grafik. Ästhetische Experimente zwischen zwei Kulturen. Die Anfänge der Computerkunst in den 1960er Jahren*. S. 52, Abb. 6-9. © Springer Verlag, Wien.

Abbildung 95: Dynamik einfacher geometrischer Formen. Arnheim, R. (1978). *Kunst und Sehen. Eine Psychologie des schöpferischen Auges*. S. 441, Abb. 274. © de Gruyter, Berlin.

Abbildung 96: Hierarchisches Netzwerkmodell. Martindale, C. (1984). The pleasures of thought: A theory of cognitive hedonics. *The Journal of Mind and Behavior*, 5, 68, Abb. 6.

Abbildung 97 : Le passé, le présent, l' Avenir (Louis-Philippe d'Orléans), Lithographie von Honoré Daumier, 1834.

Abbildung 98: Beispiele der von Redies, Hasenstein und Denzler untersuchten Photographien sowie Ergebnisse der Fourieranalysen für drei unterschiedliche Stimulusklassen. Redies, C., Hasenstein, J. & Denzler, J. (2007). Fractal-like image statistics in visual art: Similarity to natural scenes. *Spatial Vision*, 21, S. 140, Abb. 1A & I, S. 141, Abb. 2C, S. 143, Abb. 3C, F & E; Stimulus a) von diesen übernommen aus van Hateren & van der Schaaf, 1998.

Abbildung 99: Prozessmodell der ästhetischen Erfahrung. Leder, H., Belke, B., Oeberst, A. & Augustin, D. (2004). A model of aesthetic appreciation and aesthetic judgements. *British Journal of Psychology*, 95, 492, Abb. 1.