

2 Wahrnehmung

Unser Körper ist sprichwörtlich übersät mit Sensoren, um Sichtbares, Geräusche, Gerüche und körperlichen Kontakt zu erfassen. Milliarden von Neuronen verarbeiten sensorische Information und liefern das Entdeckte an die höheren Zentren des Gehirns. Die enorme Größe dieses informationsverarbeitenden Systems führt für die Kognition auf höherer Ebene zu einem Problem, das darin besteht zu entscheiden, worauf bei all der verarbeiteten sensorischen Information die Aufmerksamkeit gerichtet werden soll. Auf dieses Problem wird in ► Kap. 3 näher eingegangen. Im vorliegenden Kapitel liegt der Schwerpunkt jedoch darauf, wie diese sensorischen Systeme erkennen, was in der Außenwelt ist. Wir konzentrieren uns dabei auf die visuelle Wahrnehmung und in geringerem Umfang auf die Wahrnehmung von Sprache, da dies die beiden wichtigsten Wahrnehmungssysteme der menschlichen Spezies sind.

Dieses Kapitel behandelt die folgenden Fragen:

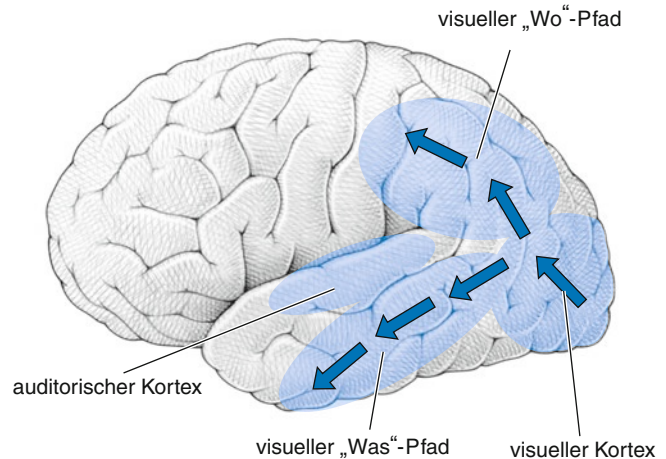
- Wie gewinnt das Gehirn Informationen aus dem optischen Signal?
- Wie wird visuelle Information in Gegenstände umgesetzt?
- Wie werden optische und sprachliche Muster erkannt?
- Inwiefern beeinflusst die Situation das Erkennen von Mustern?

2.1 Die Verarbeitung visueller Information

2.1.1 Visuelle Wahrnehmung im Gehirn

■ Abbildung 2.1 zeigt die kortikalen Regionen, die ausschließlich mit der Verarbeitung von Informationen aus dem Gesichtssinn und dem Gehör befasst sind, und veranschaulicht, dass beim Menschen genau wie bei anderen Primaten erheblicher neuronaler Aufwand betrieben wird, um visuelle Informationen zu verarbeiten. Auch andere kortikale Regionen sind, ebenso wie zahlreiche subkortikale Strukturen, an der visuellen Informationsverarbeitung beteiligt. Diese Informationen gelangen von den Augen über die subkortikalen Strukturen zur primären Sehrinde auf der „Rückseite“ des Gehirns.

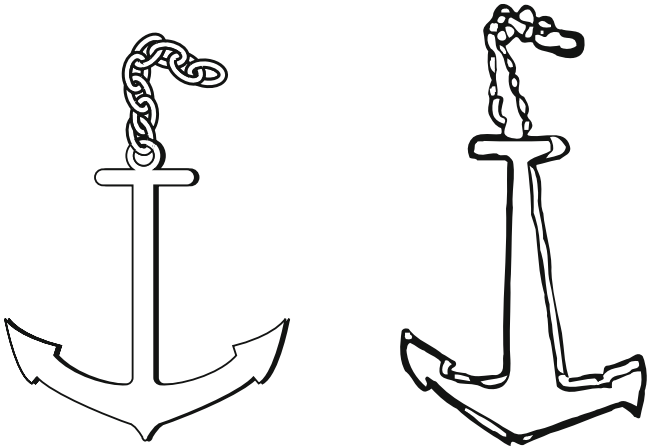
Visuelle Agnosie ist eine Unfähigkeit, Objekte zu erkennen, die weder auf einen allgemeinen Verlust der intellektuellen Fähigkeiten noch auf einen Verlust grundlegender sensorischer Wahrnehmung zurückzuführen ist. Ein Fall visueller Agnosie lag bei einem Soldaten vor, der aufgrund eines Unfalls eine Kohlenmonoxidvergiftung erlitten hatte, die zu einer Gehirnschädigung führte. Er konnte Objekte anhand ihrer ertasteten Form, ihres Geruchs oder ihres Klanges erkennen, war jedoch nicht in der Lage, das Bild eines Kreises von dem eines Quadrats zu unterscheiden oder Gesichter und Buchstaben zu erkennen



■ **Abb. 2.1** Einige der kortikalen Strukturen, die beim Sehen und Hören beteiligt sind: der visuelle Kortex, der auditorische Kortex, der visuelle Wo-Pfad und der visuelle Was-Pfad. (Ungerleider & Mishkin, 1982)

(Benson & Greenberg, 1969). Andererseits konnte er Helligkeitsabstufungen und Farben unterscheiden und angeben, in welche Richtung sich ein Objekt bewegte. Demzufolge war sein sensorisches System in der Lage, visuelle Informationen aufzunehmen, aber seine Gehirnschädigung hatte auf irgendeine Weise zu einem Verlust der Fähigkeit geführt, visuelle Informationen in eine Wahrnehmungserfahrung umzusetzen. Dieses Fallbeispiel zeigt, dass Wahrnehmung viel mehr ist als die einfache Aufnahme sensorischer Informationen.

Allgemein wird visuelle Agnosie in **apperzeptive Agnosie** und **assoziative Agnosie** eingeteilt (ein Überblick findet sich in Farah, 1990). Dem erwähnten Patienten von Benson und Greenberg wird eine apperzeptive Agnosie zugeschrieben. Solche Patienten sind nicht in der Lage, einfache Figuren wie Kreise oder Dreiecke zu erkennen oder Figuren nach Vorlagen zu zeichnen. Im Gegensatz dazu können Patienten mit assoziativer Agnosie einfache Formen erkennen und sogar komplexe Objekte erfolgreich nachzeichnen. Sie sind jedoch nicht in der Lage, die komplexen Objekte zu erkennen. In ■ Abb. 2.2 sind das Original und die Zeichnung eines Ankers dargestellt, die ein Patient mit assoziativer Agnosie anfertigte (Ratcliff & Newcombe, 1982). Trotz seiner Fähigkeit, eine relativ akkurate Zeichnung anzufertigen, war es dem Patienten nicht möglich, das Dargestellte als Anker zu erkennen (er bezeichnete es als Regenschirm). Es wird allgemein angenommen, dass bei Patienten mit apperzeptiver Agnosie Probleme bei relativ frühen visuellen Prozessen bestehen, wohingegen bei Patienten mit assoziativer Agnosie die frühen Prozesse intakt sind, sie aber Schwierigkeiten mit der Mustererkennung haben, die später stattfindet. Üblicherweise wird zwischen früher und später Informationsverarbeitung



■ **Abb. 2.2** Ein Patient mit assoziativer Agnosie war zwar in der Lage, diese Abbildung eines Ankers auf der linken Seite nachzuzeichnen (seine Zeichnung ist in der Abbildung rechts dargestellt), jedoch konnte er das Objekt nicht als Anker erkennen. (Aus Ellis & Young, 1988. Nachdruck mit Genehmigung der Experimental Psychology Society. © 1988 *Quarterly Journal of Experimental Psychology*)

im visuellen Informationsfluss unterschieden; diese Unterscheidung stellt eine der Organisationsgrundlagen dieses Kapitels dar.

■ **Abbildung 2.3** bietet einer Person mit normaler Wahrnehmung Gelegenheit, die Unterscheidung zwischen frühen und späten visuellen Prozessen zu verstehen. Wenn Sie dieses Bild niemals zuvor gesehen haben, werden Sie darin nur ein paar Tintenkleckse sehen. Sie mögen in der Lage sein, die Größe der verschiedenen Kleckse zu beurteilen und sie in der Art und Weise wiederzugeben wie der Patient, der von Ratcliff und Newcombe untersucht wurde; ein Muster werden Sie jedoch nicht erkennen. Wenn Sie allerdings etwas länger darauf blicken, werden Sie plötzlich einen Hund in der Mitte des Bildes sehen. Jetzt war Ihre Mustererkennung erfolgreich, und Sie haben dem Gesehenen eine Interpretation hinzugefügt. Wir beginnen dieses Kapitel mit einer Erörterung, wie visuelle Information vor der Mustererkennung verarbeitet wird, und im Anschluss daran werden wir Prozesse der Mustererkennung behandeln.

➤ **Die visuelle Wahrnehmung kann in eine frühe Phase, in der Formen und Objektaus der visuellen Szenerie extrahiert werden, und eine späte Phase, in der Formen und Objekte erkannt werden, unterteilt werden.**

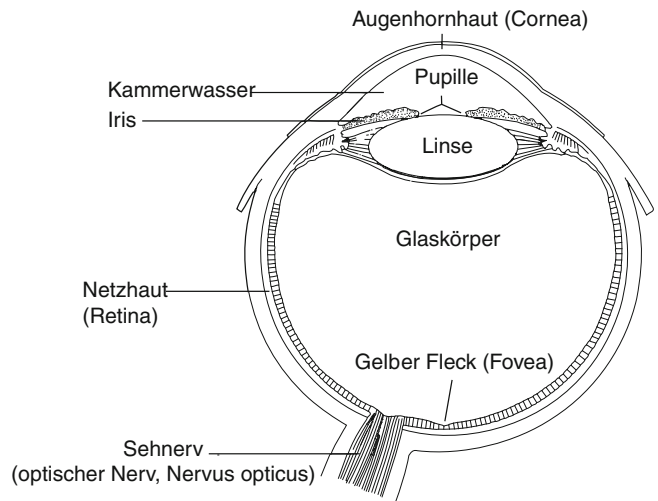
2.1.2 Frühe Prozesse der visuellen Informationsverarbeitung

Man weiß eine ganze Menge über die neuronalen Prozesse, die der frühen visuellen Informationsverarbeitung zugrunde liegen.

■ **Abbildung 2.4** zeigt eine schematische Darstellung des Auges. Das Licht durchquert die Linse und den Glaskörper und fällt auf die Netzhaut an der Rückseite des Auges. Die Retina beherbergt die Photorezeptorzellen, die aus lichtempfindlichen Molekülen



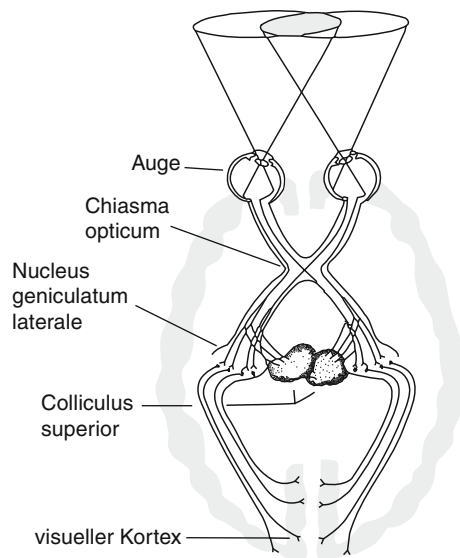
■ **Abb. 2.3** Eine Szenerie, in der wir anfangs lediglich visuelle Details aus frühen Wahrnehmungsprozessen (Kleckse) erkennen können, erst nach längerer Betrachtung tritt der Umriss eines Hundes hervor. (© Ronald James)



■ **Abb. 2.4** Eine schematische Darstellung des Auges. Das Licht tritt durch die Hornhaut (Cornea) ein, durchquert das Kammerwasser, die Pupille, die Linse und den Glaskörper, und trifft anschließend auf die Netzhaut (Retina), die durch das Licht stimuliert wird. (Aus Lindsay & Norman, 1977. Nachdruck mit Genehmigung des Verlags. © 1977 Academic Press)

aufgebaut sind. Werden diese Moleküle Licht ausgesetzt, unterliegen sie strukturellen Veränderungen. Das Licht wird beim Durchqueren des Glaskörpers leicht gestreut, sodass das Bild, das auf der Netzhaut entsteht, nicht völlig scharf ist. Eine der Funktionen der frühen Prozesse der visuellen Informationsverarbeitung besteht darin, diesem Bild Schärfe zu verleihen.

Das Licht wird durch einen photochemischen Prozess in Nervenimpulse umgewandelt. Es gibt zwei verschiedene Typen von Photorezeptoren im Auge – Stäbchen und Zapfen. Die Zapfen sind für das Farbsehen zuständig und erbringen eine hohe Auflösung und Schärfe. Um eine Reaktion der Stäbchen hervorzurufen, ist weniger Lichtenergie notwendig, jedoch ist ihre Auflösung geringer. Daraus ergibt sich, dass sie hauptsächlich für weniger scharfes Schwarz-weiß-Sehen zuständig sind, wie wir es

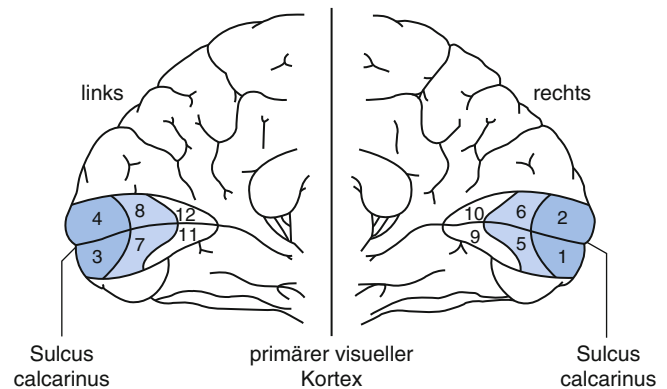
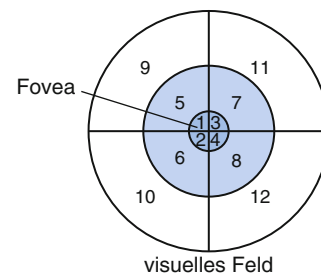


■ **Abb. 2.5** Nervenbahnen vom Auge zum Gehirn. Die optischen Nerven beider Augen treffen sich im Chiasma opticum. Information über den linken Teil des visuellen Feldes gelangt in die rechte Hirnhälfte, und Information über den rechten Teil des visuellen Feldes gelangt in die linke Hirnhälfte. Fasern des optischen Nerven haben synaptische Verbindungen zu Zellen in subkortikalen Strukturen, wie dem Nucleus geniculatus lateralis und dem Colliculus superior. Beide Strukturen sind mit dem visuellen Kortex verbunden. (Aus BIOLOGICAL SCIENCE, THIRD EDITION by William T. Keeton Paula DiSanto Bensaoudon, illus. Copyright © 1980, 1979, 1978, 1972, 1967 by W. W. Norton & Company, Inc. Used by permission of W. W. Norton & Company, Inc.)

zum Beispiel nachts erleben. Eine hohe Dichte an Zapfen findet man vor allem in einem kleinen Bereich der Netzhaut, der **Fovea** (Gelber Fleck) genannt wird. Wenn wir ein Objekt fixieren, dann bewegen wir unsere Augen so, dass das Objekt auf die Fovea fällt. Dies ermöglicht uns, die hohe Auflösung der Zapfen bei der Wahrnehmung des Objekts zu maximieren. Foveales Sehen erfasst feine Details. Der Rest des visuellen Feldes, die Peripherie, ist für die Erfassung eher globaler Information einschließlich des Erkennens von Bewegung verantwortlich.

Die Rezeptorzellen sind synaptisch mit Bipolarzellen und diese wiederum mit Ganglienzellen verbunden, deren Axone aus dem Auge austreten und den optischen Nerv bilden, der zum Gehirn führt. Insgesamt befinden sich etwa 800.000 Ganglienzellen im optischen Nerv jedes Auges. Jede Ganglienzelle encodiert Informationen aus einer kleinen Region der Netzhaut. Die Feuerrate des Axons einer Ganglienzelle encodiert typischerweise den Betrag an Stimulation, der sich durch die Menge des auf diese Region der Netzhaut auftreffenden Lichtes ergibt.

■ **Abbildung 2.5** zeigt die Nervenbahnen von den Augen zum Gehirn. Die optischen Nerven beider Augen treffen sich im Chiasma opticum; dabei überkreuzen sich die Nerven der nasalen Seiten der Netzhäute (die jeweils innen liegende Seite nahe der Nase) und führen zur gegenüberliegenden Seite des Gehirns, während die Nerven der Außenseiten der Netzhäute zur jeweils auf derselben Kopfseite gelegenen Hirnhälfte weiterführen. Dies bedeutet, dass die rechten Hälften beider Augen mit der rechten Gehirnhälfte verbunden sind. Wie ■ **Abb. 2.5** zeigt, bündelt die Linse das Licht so, dass der linke Teil des visuellen Feldes auf die



■ **Abb. 2.6** Die Projektion des visuellen Feldes (oben) auf den visuellen Kortex. Die oberen Bereiche werden unterhalb des Sulcus calcarinus abgebildet, die unteren Bereiche des visuellen Feldes oberhalb dieser Furche. Außerdem wird der linke Teil des in vier Quadranten eingeteilten Gesichtsfeldes auf den rechten Kortex projiziert. Man beachte die überproportional große Kortexfläche für den Bereich des schärfsten Sehens, der Fovea. (Nach Abbildung 29.7 in Kandel et al., 1991)

rechte Hälfte jedes Auges fällt. Dadurch wird Information über den linken Teil des visuellen Feldes an die rechte Gehirnhälfte und entsprechend Information über die rechte Seite des visuellen Feldes an die linke Gehirnhälfte übermittelt. Dies ist ein Beispiel für die in ► **Kap. 1** behandelte allgemeine Tatsache, dass die linke Hirnhälfte Informationen über den rechten Teil der Welt und die rechte Hirnhälfte Informationen über den linken Teil der Welt verarbeitet.

Im Gehirn angelangt sind die Fasern der Ganglienzellen synaptisch mit Zellen verschiedener subkortikaler Strukturen verbunden (subkortikale Strukturen befinden sich unterhalb des Kortex). Zwei dieser Strukturen sind der Nucleus geniculatus lateralis (seitlicher Kniehöcker) und der Colliculus superior (■ **Abb. 2.5**). Man nimmt an, dass der Nucleus geniculatus lateralis ein Bestandteil des neuronalen Pfades ist, der wichtig für das Wahrnehmen von Details und das Erkennen von Objekten ist, während der Colliculus superior mit dem Lokalisieren von Objekten im Raum zu tun hat. Man nennt dies auch die Was-Wo-Unterscheidung. Beide neuronalen Strukturen sind mit dem primären visuellen Kortex verbunden (Brodmann-Areal 17 im Hirnatlas der ► **Farbtafel 1.1**), dem ersten kortikalen Areal, das visuellen Input erhält; allerdings gibt es noch weitere visuelle Areale, wie etwa die Brodmann-Areale 18 und 19, die es umgeben.

■ **Abbildung 2.6** beschreibt die Projektion der sichtbaren Welt im primären visuellen Kortex. Sie zeigt, dass der visuelle Kortex, wie in ► **Kap. 1** erörtert, topologisch organisiert ist. Die Fovea

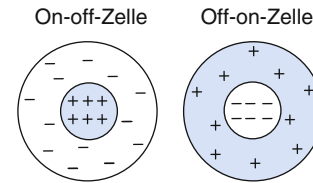
ist in überproportionaler Größe repräsentiert, die Randzone hingegen erscheint verkleinert. ■ Abbildung 2.6 zeigt wie auch ■ Abb. 2.5, dass das linke visuelle Feld im rechten Kortex und das rechte Feld im linken Kortex wiedergegeben wird. Es illustriert auch eine weitere Umkehr der Abbildung – der obere Teil des visuellen Feldes wird im unteren Teil des visuellen Kortex dargestellt und der untere Teil im oberen Bereich.

Wie ■ Abb. 2.1 veranschaulicht, folgen visuelle Informationen – nachdem sie in den primären visuellen Kortex projiziert worden sind – zwei Pfaden. Die Informationen, die sich entlang des Wo-Pfades ausbreiten, gelangen in parietale Hirnregionen, die auf die Verarbeitung räumlicher Informationen und die Koordination des Gesichtssinnes mit Handlungen spezialisiert sind. Informationen, die sich entlang des Was-Pfades ausbreiten, gelangen in den temporalen Kortex, der für das Erkennen von Objekten zuständig ist. Affen mit Läsionen im Wo-Pfad haben Schwierigkeiten dabei, das Erkennen bestimmter Orte zu erlernen, während Affen mit Läsionen im Was-Pfad Schwierigkeiten beim Erlernen des Erkennens von Objekten aufweisen (Pohl, 1973; Ungerleider & Brody, 1977). Andere Forscher (z. B. Milner & Goodale, 1995) argumentieren, dass der Wo-Pfad in Wirklichkeit ein auf Handlungen spezialisierter Informationsverarbeitungspfad sei. Sie wiesen darauf hin, dass Patienten mit einer Agnosie infolge von Schädigungen am Temporallappen, jedoch intakten Parietallappen, oft in der Lage sind, angemessene Handlungen in Bezug auf Objekte auszuführen, die von ihnen nicht erkannt werden können. So konnte eine Patientin (Goodale, Milner, Jakobson & Carey, 1991) zielsicher den Arm ausstrecken und einen Türgriff ergreifen, den sie nicht erkennen konnte.

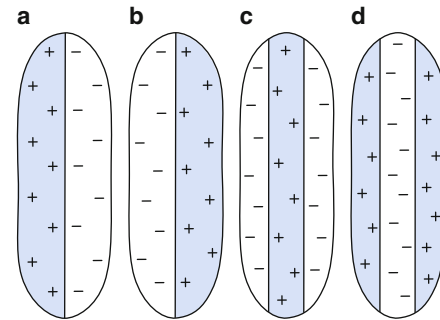
➤ **Lichtenergie wird durch einen photochemischen Prozess in neuronale Aktivität umgewandelt. Diese Information wird über verschiedene neuronale Bahnen zur Sehrinde und von dort entlang der Was- und Wo-Pfade durch das Gehirn weitergeleitet.**

2.1.3 Die Kodierung von Information in den visuellen Zellen

Die Forschungen Kufflers (1953) zeigten, wie Informationen durch die Ganglienzellen enkodiert werden. Diese Zellen feuern im Allgemeinen mit einer Spontanrate, auch wenn kein Licht von den Augen aufgenommen wird. Einige Ganglienzellen zeigen einen Anstieg der Spontanrate, wenn Licht auf eine kleine Region der Netzhaut fällt. Fällt jedoch Licht auf jene Region, die genau dieses empfindliche Zentrum umgibt, sinkt die Spontanrate. Fällt Licht auf Regionen, die weiter von dem empfindlichen Zentrum entfernt sind, so zeigt sich keine Veränderung der Spontanrate. Diese Zellen sind als On-off-Zellen bekannt. Es gibt auch Off-on-Ganglienzellen, bei denen Licht im Zentrum die Spontanrate des Feuerns senkt und Licht in der Umgebung die Rate erhöht. Zellen im Nucleus geniculatum laterale verhalten sich auf die gleiche Art und Weise. ■ Abbildung 2.7 zeigt die rezeptiven Felder dieser Zellen.



■ **Abb. 2.7** On-off- und Off-on-rezeptive Felder von Ganglienzellen und den Zellen im Nucleus geniculatum laterale

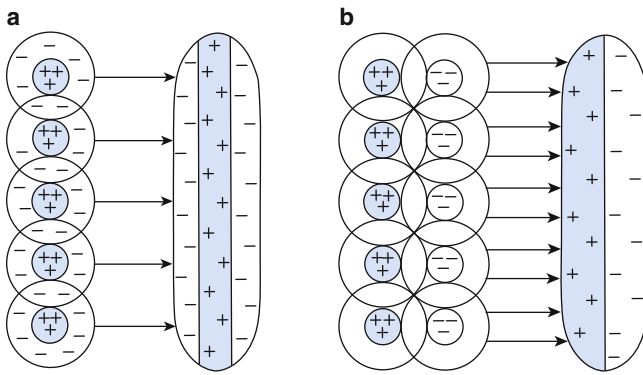


■ **Abb. 2.8** Antwortmuster von Zellen im visuellen Kortex. **a** und **b** sind Kantendetektoren, die positiv auf Licht auf der einen Seite und negativ auf Licht auf der anderen Seite reagieren. **c** und **d** sind Balkendetektoren, sie reagieren positiv auf Licht im Zentrum und negativ auf Licht in der Peripherie beziehungsweise umgekehrt

Hubel und Wiesel (1962) fanden in ihrer Untersuchung über den primären visuellen Kortex der Katze, dass kortikale visuelle Zellen in einer komplexeren Art und Weise antworten als Ganglienzellen und Zellen im Nucleus geniculatum laterale. ■ Abbildung 2.8 zeigt vier Antwortmuster, die bei kortikalen Zellen beobachtet wurden. Man erkennt, dass alle rezeptiven Felder eine längliche Form besitzen, im Gegensatz zu den kreisrunden rezeptiven Feldern der On-off- und Off-on-Zellen. Die Zelltypen, die in den ■ Abb. 2.8a und ■ Abb. 2.8b abgebildet sind, heißen **Kantendetektoren**. Sie reagieren positiv auf Licht an der einen Seite einer Linie und negativ auf Licht an der anderen Seite. Sie zeigen maximale Reaktion, wenn eine Hell-dunkel-Kante entlang der Grenzlinie verläuft. Die Zelltypen der ■ Abb. 2.8c und ■ Abb. 2.8d sind **Balkendetektoren**. Sie reagieren positiv auf Licht im Zentrum und negativ auf Licht in der Peripherie beziehungsweise umgekehrt. Somit reagiert ein Balkendetektor mit einem positiven Zentrum maximal, wenn ein Lichtstreifen genau sein Zentrum abdeckt.

■ Abbildung 2.9 veranschaulicht, wie On-off- und Off-on-Zellen zur Bildung von Balken- oder Kantendetektoren kombiniert sein können. Hierbei ist zu beachten, dass eine einzelne Zelle nicht ausreicht, um einen Detektor zu stimulieren. Vielmehr reagiert der Detektor auf *Muster* von stimulierten Zellen. Wir sehen, dass das Nervensystem sogar auf diesem niedrigen Niveau Information in Form von Mustern neuronaler Aktivität verarbeitet. Dieser Aspekt wurde in ► Kap. 1 eingehend behandelt.

Sowohl Kanten- als auch Balkendetektoren sind spezifisch im Hinblick auf Position, Ausrichtung und Ausdehnung. Das heißt, dass sie nur auf eine Stimulation in einem kleinen Bereich des visuellen Feldes reagieren, auf Balken und Kanten eng

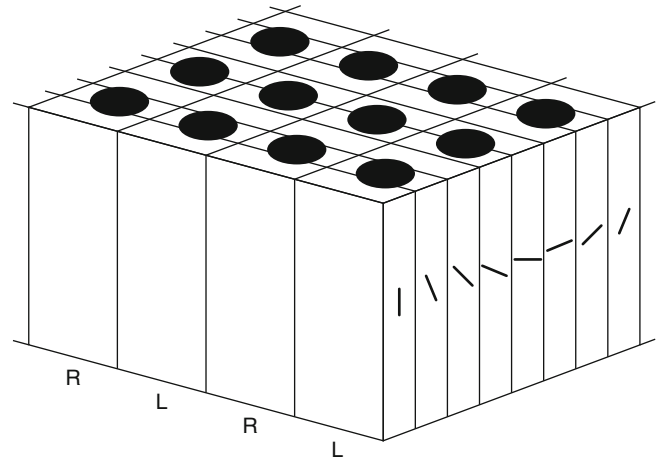


■ **Abb. 2.9** Hypothetische Kombinationen von On-off- und von Off-on-Zellen zum Aufbau von **a** Balkendetektoren und **b** Kantendetektoren

begrenzter Ausrichtung und Ausdehnung. Allerdings scheinen verschiedene Detektoren auf verschiedene Ausdehnungen und Ausrichtungen eingestellt zu sein. Dadurch würde eine Teilmenge von Balkendetektoren durch solch ein Muster maximal stimuliert werden.

■ **Abbildung 2.10** zeigt die Anordnung von Zellen in Form von Hypersäulen im primären visuellen Kortex nach Hubel und Wiesel (1977). Sie fanden heraus, dass der visuelle Kortex in Regionen von 2×2 mm eingeteilt ist, die sie Hypersäulen nannten. Jede Hypersäule repräsentiert eine bestimmte Region des rezeptiven Feldes. Wie wir in ► Kap. 1 festgestellt haben, ist der visuelle Kortex topographisch organisiert, und so finden sich benachbarte Gebiete im visuellen Feld in benachbarten Hypersäulen wieder. ■ **Abbildung 2.10** zeigt, dass jede Hypersäule wiederum einen zweidimensionalen Aufbau aufweist. Entlang der einen Dimension liegen abwechselnde Reihen, die jeweils Input vom linken und vom rechten Auge erhalten. Entlang der anderen Dimension variieren die Zellen im Hinblick auf die Orientierung, für die sie am empfindlichsten sind. Benachbarte Regionen repräsentieren ähnliche Orientierungen. Diese Organisation veranschaulicht uns, wie viel Information des visuellen Feldes enkodiert wird. Hunderte von Raumregionen werden für jedes Auge getrennt repräsentiert, und innerhalb dieser Regionen werden viele verschiedene Orientierungen repräsentiert. Obwohl dies in ■ **Abb. 2.10** nicht zu sehen ist, kommt noch hinzu, dass unterschiedliche Zellen unterschiedliche Größen und Längen von Linien kodieren. Es wird also eine gewaltige Informationsmenge aus dem visuellen Signal extrahiert, bis es die ersten kortikalen Areale erreicht.

Zusätzlich zu dieser reichhaltigen Repräsentation der Orientierung von Linien extrahiert das visuelle System noch andere Informationen aus dem visuellen Signal. Beispielsweise können wir auch die Farbe von Objekten wahrnehmen und erkennen, ob sie sich bewegen. Livingstone und Hubel (1988) vertreten die Ansicht, dass all diese unterschiedlichen Dimensionen (Form, Farbe und Bewegung) voneinander getrennt im visuellen System verarbeitet werden. Es gibt viele unterschiedliche visuelle Pfade und viele verschiedene Areale des Kortex, die der visuellen Verarbeitung dienen (32 nach der Zählung von Van Essen und DeYoe, 1995). Unterschiedliche Pfade besitzen Zellen mit



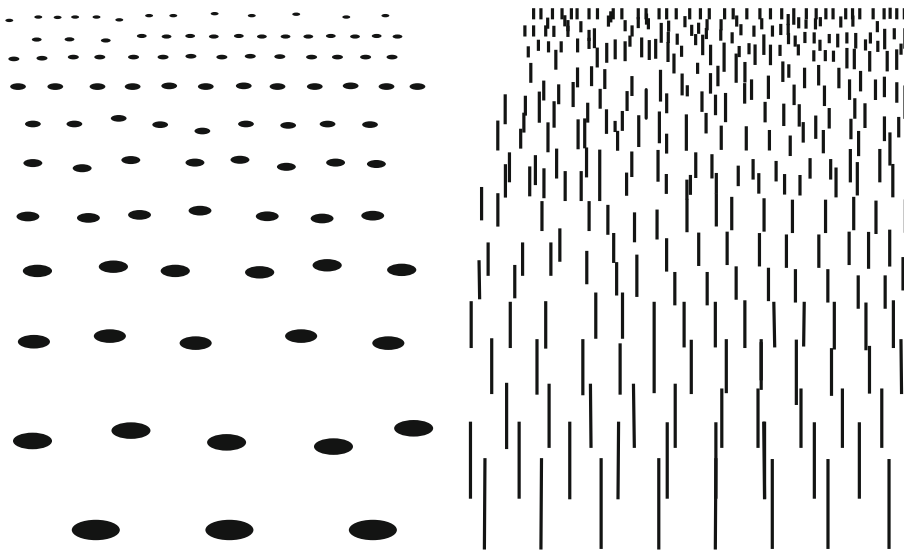
■ **Abb. 2.10** Darstellung einer Hypersäule im visuellen Kortex. Die Hypersäule ist in einer Dimension nach Input vom rechten oder linken Auge organisiert. In der anderen Dimension ist sie nach der Ausrichtung von Linien organisiert, für die die rezeptiven Zellen jeweils am empfindlichsten sind. Benachbarte Regionen repräsentieren ähnliche Orientierungen. (Nach Horton, 1984. Cytochrome oxidase patches: A new cytoarchitectonic feature of monkey visual cortex. *Philosophical Transactions of the Royal Society of London*, 304, 199–253. By permission of the Royal Society. © 1984 The Royal Society)

spezifischen Empfindlichkeiten für Farbe, Bewegung und Orientierung. Es scheint, dass das visuelle System den Reiz in viele unabhängige Merkmale zerlegt und deren Anordnung repräsentiert. Solche räumlichen Repräsentationen von Merkmalen heißen **Merkmalskarten** (*feature maps*; Wolfe, 1994). Wenn sich also ein vertikaler roter Balken an einem bestimmten Ort bewegt, gibt es verschiedene Merkmalskarten, die repräsentieren, dass er rot ist, vertikal ist und dass er sich an diesem Ort bewegt. Diese Karten können sich in unterschiedlichen visuellen Arealen des Gehirns befinden.

➤ **Die Ganglienzellen enkodieren das visuelle Feld mit Hilfe von On-off- und Off-on-Zellen. Diese werden durch höhere visuelle Prozesse verknüpft, um unterschiedliche Merkmale abzubilden.**

2.1.4 Die Wahrnehmung von Tiefe und von Oberflächen

Auch nachdem das visuelle System Kanten und Balken in der Umwelt identifiziert hat, muss noch eine umfangreiche Informationsverarbeitung durchgeführt werden, bevor das System in der Lage ist, die Welt wahrzunehmen. Eines der zu lösenden Probleme betrifft die Feststellung, wo sich diese Kanten und Balken im Raum befinden. Das grundlegende Problem dabei ist, dass die an der Netzhaut bereitgelegte Information von Natur aus zweidimensional ist (2-D), während wir eine dreidimensionale (3-D) Repräsentation der Welt erstellen müssen. Das visuelle System bedient sich einiger Hinweisreize, um auf Entfernungen zu schließen. Einer dieser Hinweisreize ist der Texturgradient,



■ **Abb. 2.11** Beispiele für Texturgradienten. Elemente scheinen mit zunehmender Distanz vom Betrachter dichter gepackt zu sein. (Aus Gibson, 1950. Nachdruck mit Genehmigung des Verlags. © 1950 Houghton Mifflin)

der die Tendenz von Elementen bezeichnet, mit zunehmender Distanz vom Betrachter dichter gepackt zu erscheinen. Gibson (1950) gibt hierfür klassische Beispiele (■ Abb. 2.11). Obwohl sich die Linien und Ellipsen auf einer ebenen Fläche befinden, führt die Veränderung in der Textur zum Eindruck von Tiefe. Ein anderer Hinweisreiz für Tiefe ist die Stereopsis, die Möglichkeit, dreidimensionale Tiefe durch die leicht unterschiedlichen Abbilder der Umwelt in beiden Augen wahrzunehmen. 3-D-Brillen, wie man sie in einigen Kinos und bei anderen Vorführungen findet, beruhen auf diesem Prinzip. Sie filtern das Licht, das von einer einzigen 2-D-Quelle (beispielsweise einer Kinoleinwand) kommt, in einer Weise, durch die unterschiedliche Lichtinformationen auf beide Augen fallen. Die Wahrnehmung einer dreidimensionalen Struktur durch Stereopsis kann sehr eindrucksvoll sein.

Eine dritte Informationsquelle über dreidimensionale Strukturen beruht auf der sogenannten Bewegungsparallaxe. Bewegen sich Punkte, die weiter vom Beobachter entfernt sind, so ergibt dies eine langsamere Bewegung über die Retina hinweg als für nähere Punkte. Das Gleiche gilt, wenn man den Kopf bewegt: Entferntere Objekte wandern langsamer über die Netzhaut, als näher gelegene Objekte dies tun. Eine interessante Demonstration ergibt sich, wenn Sie auf die Blätter eines nahen Baumes oder Busches blicken und dabei ein Auge geschlossen halten. Ohne die stereoptische Information werden Sie den Eindruck eines sehr flachen Bildes haben, wobei es schwierig ist, die Lage der vielen Blätter relativ zueinander zu bestimmen. Bewegen Sie nun den Kopf, erscheint die dreidimensionale Struktur des Baumes plötzlich sehr klar, und Sie können leicht die Lage der Blätter und Zweige relativ zueinander erkennen.

Obwohl es einfach ist, die Bedeutung von Hinweisreizen wie Texturgradient, Stereopsis und Bewegungsparallaxe für die Tiefenwahrnehmung zu demonstrieren, ist es schwierig zu verstehen, wie das Gehirn diese Informationen verarbeitet. Im Bereich konstruktivistischer Ansätze (*computational vision*) wurde viel

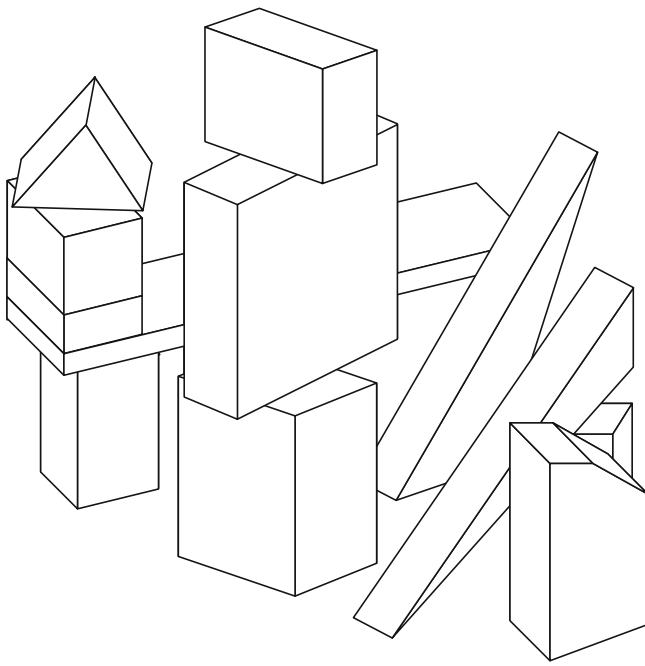
an diesen Vorgängen gearbeitet. David Marr (1982) hatte großen Einfluss mit seinem Vorschlag, dass diese verschiedenen Informationsquellen zusammenwirken, um das zu erstellen, was er eine **2½-D-Skizze** nennt; diese 2½-D-Skizze erlaubt die Bestimmung der relativen Lage von verschiedenen visuellen Merkmalen zum Betrachter. Er erkannte jedoch auch, wie weit diese Repräsentation von einer eigentlichen Wahrnehmung der Welt entfernt ist. Genau genommen bildet eine solche Repräsentation nur Teile von Oberflächen ab und stellt noch nicht dar, wie diese zusammen ein Objekt draußen in der Umgebung bilden (das in ■ Abb. 2.3 angesprochene Problem). Er verwendete den Begriff **3-D-Modell**, um eine Repräsentation von Objekten in der Szenerie zu bezeichnen.

► **Hinweisreize wie Texturgradient, Stereopsis und Bewegungsparallaxe wirken zusammen, um eine Repräsentation der Lage von Oberflächen im dreidimensionalen Raum zu erstellen.**

2.1.5 Objektwahrnehmung

Ein Hauptproblem bei der Berechnung einer Repräsentation der Welt liegt in der Gliederung von Objekten. Zu wissen, wo sich die Linien und Balken im Raum befinden, reicht nicht aus; wir müssen darüber hinaus wissen, welche zusammengehören, um Objekte zu bilden. Betrachten Sie die Anordnung in ■ Abb. 2.12. Viele Linien verlaufen kreuz und quer, aber irgendwie setzen wir sie zusammen, um zur Wahrnehmung einer Reihe von Objekten zu gelangen.

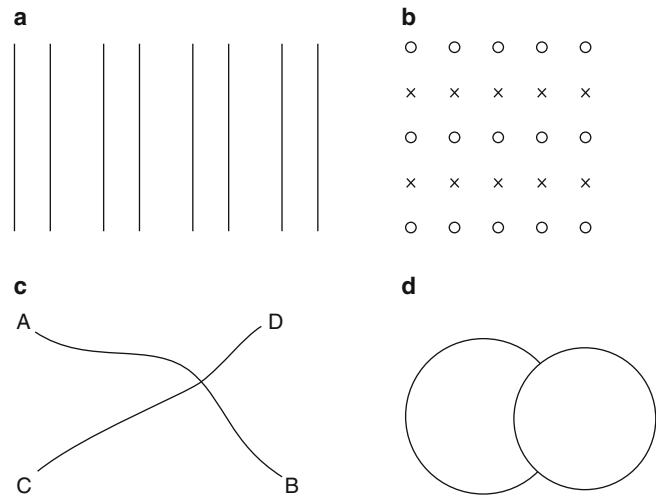
Gewöhnlich folgen wir bei der Ordnung von Objekten zu Einheiten bestimmten Prinzipien. Diese Prinzipien wurden von den Gestaltpsychologen, die sie zuerst formuliert haben (z. B.



▣ **Abb. 2.12** Ein Beispiel dafür, wie wir die Wahrnehmung vieler unterbrochener Linien zur Wahrnehmung zusammenhängender Körper aggregieren. (Aus Winston, 1970. Nachdruck mit Genehmigung des Verlags. © 1970 Massachusetts Institute of Technology)

Wertheimer, 1912), **Gestaltgesetze der Wahrnehmungsorganisation** genannt.¹ Betrachten Sie die verschiedenen Teile in ▣ Abb. 2.13. In ▣ Abb. 2.13a nehmen wir eher vier Paare von Linien und nicht etwa acht einzelne Linien wahr. Hier wird das Prinzip der Nähe veranschaulicht: Nahe beieinander liegende Elemente organisieren sich oft zu Einheiten. ▣ Abbildung 2.13b veranschaulicht das Prinzip der Ähnlichkeit. Wir neigen dazu, diese Anordnung als Reihen von Kreisen zu sehen, die sich mit Reihen von Kreuzen abwechseln. Ähnlich aussehende Objekte werden bevorzugt zu einer Gruppe zusammengefasst. ▣ Abbildung 2.13c veranschaulicht das Prinzip des glatten Verlaufs. Wir nehmen zwei Linien wahr, eine von A nach B und eine andere von C nach D, obwohl es dafür eigentlich keinen Grund gibt. Diese Skizze könnte auch für ein anderes Paar von Linien stehen: eine, die von A nach D führt, und die andere, die von C nach B führt. Die Linie von A nach B weist jedoch einen glatteren Verlauf auf als die stark abgelenkte Linie von A nach D. ▣ Abbildung 2.13d veranschaulicht die Prinzipien der Geschlossenheit und der guten Gestalt. Wir sehen in der Zeichnung einen Kreis, der teilweise einen anderen Kreis verdeckt, obwohl das verdeckte Objekt viele andere Formen haben könnte.

Diese Prinzipien ordnen für gewöhnlich sogar völlig unbekannte Stimuli zu Einheiten. Palmer (1977) untersuchte das Wiedererkennen von Figuren, wie sie ▣ Abb. 2.14 darstellt. Er zeigte den Probanden zunächst Stimuli der Art von ▣ Abb. 2.14a



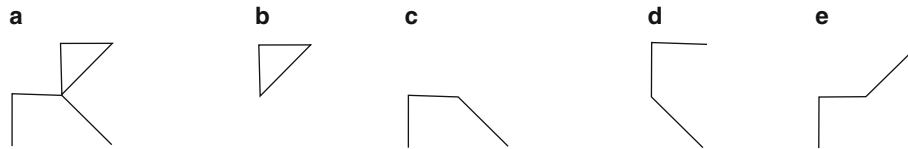
▣ **Abb. 2.13** Veranschaulichung der Gestaltprinzipien: **a** das Prinzip der Nähe, **b** das Prinzip der Ähnlichkeit, **c** das Prinzip des glatten Verlaufs und **d** das Prinzip der Geschlossenheit beziehungsweise der guten Gestalt

und ließ sie dann entscheiden, ob die Teilstücke ▣ Abb. 2.14b bis ▣ Abb. 2.14e einen Bestandteil der Originalfigur darstellen. Der Stimulus in ▣ Abb. 2.14a neigt dazu, sich in ein Dreieck (Geschlossenheit) und in einen verbogenen Buchstaben n (glatter Verlauf) zu gliedern. Palmer konnte zeigen, dass die Wiedererkennung der Teilstücke am schnellsten erfolgte, wenn diese Teilstücke den von den Gestaltprinzipien vorhergesagten Segmenten entsprachen. Die Stimuli der ▣ Abb. 2.14b und ▣ Abb. 2.14c wurden also schneller als jene der ▣ Abb. 2.14d und ▣ Abb. 2.14e wiedererkannt. Wir sehen also, dass das Wiedererkennen entscheidend von der ursprünglichen Gliederung der Figur abhängt. Wenn die Gestaltprinzipien zu einer Gliederung führen, die der tatsächlichen Struktur des Musters widerspricht, wird das Erkennen unter Umständen in starkem Maße beeinträchtigt. ZuMbEiSpIeLiStDiEsErSaTzScHwIeRiGzUIEsEn. Die Gründe für die Schwierigkeiten liegen darin, dass das Gestaltprinzip der Ähnlichkeit es erschwert, benachbarte Buchstaben unterschiedlicher Groß- und Kleinschreibung als Einheiten zu sehen, und dass aus dem Prinzip der Nähe folgende Hinweise nach der Entfernung der Wortzwischenräume wegfallen.

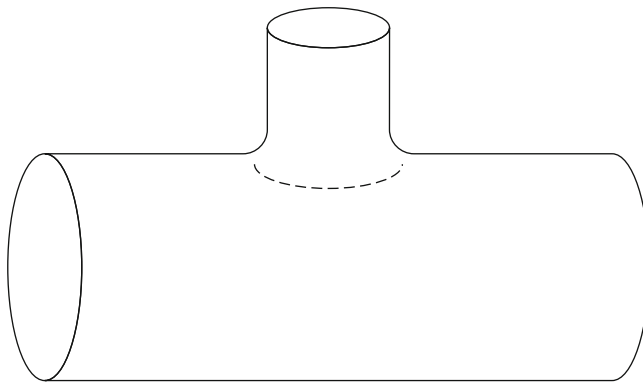
Diese Vorstellungen über die Gliederung von Objekten können auf die Beschreibung der Unterteilung komplexerer dreidimensionaler Strukturen ausgedehnt werden. ▣ Abbildung 2.15 veranschaulicht einen Vorschlag von Hoffman und Richards (1985), wie gestaltartige Prinzipien genutzt werden können, um eine Umrisszeichnung eines Objekts in Teilobjekte zu gliedern. Sie stellten fest, dass sich an der Stelle, an der ein Segment an ein anderes angefügt wird, typischerweise ein konkaver Verlauf der Umrisslinie zeigt. Hier scheint das Gestaltprinzip des glatten Verlaufs herangezogen zu werden: Die Linien an den konkaven Stellen sind keine glatten Verläufe; deshalb können sie die zu verbindenden Teile nicht gruppieren.

Gegenwärtig glaubt man, dass die der Fähigkeit zur Identifikation von Position und Form eines Objektes im dreidimensionalen Raum zugrunde liegenden visuellen Mechanismen zum großen Teil angeboren sind. Bereits im Säuglingsalter scheinen

¹ Ursprünglich als **Gestaltgesetze der Wahrnehmungsorganisation** bezeichnet, hat sich für die Organisationsprinzipien der Wahrnehmung im deutschen wissenschaftlichen Sprachgebrauch der Begriff **Gestaltprinzipien** durchgesetzt. [Anmerkung der Übersetzer]



■ **Abb. 2.14** Beispiele für Stimuli, wie sie Palmer (1977) zur Untersuchung der Gliederung unbekannter Figuren verwendete. **a** zeigt den Originalstimulus, den die Probanden sahen. **b** bis **e** zeigen Teilfiguren zur Wiedererkennung. Die Stimuli in **b** und **c** wurden schneller erkannt als diejenigen in **d** und **e**



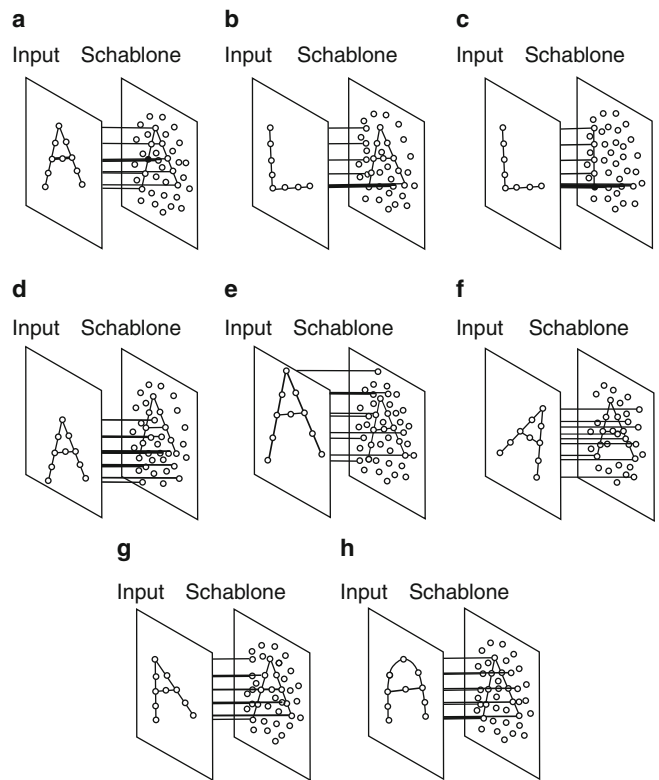
■ **Abb. 2.15** Gliederung eines Objekts in Teilobjekte: Die Teilobjektgrenze kann durch eine Kontur beschrieben werden, die den Stellen mit maximaler konkaver Biegung folgt (gestrichelte Linie). (Aus Stillings et al., 1987. Cognitive Science: An Introduction; nach Hoffman & Richards, 1985. Parts of Recognition. Cognition, 18, pp. 65–96. Nachdruck mit Genehmigung von MIT Press. © 1987 Massachusetts Institute of Technology)

Objekte und Formen erkannt zu werden und Einschätzungen, wo sich diese Objekte im dreidimensionalen Raum befinden, möglich zu sein (z. B. Granrud, 1986, 1987).

➤ **Visuelle Szenen werden mit Hilfe der Gestaltprinzipien in Objekte gegliedert.**

2.2 Visuelle Mustererkennung

Wir sind in unserer Darstellung der visuellen Informationsverarbeitung an jenem Punkt angelangt, an dem wir die visuelle Welt in Objekte zergliedern können. Jedoch sind wir noch weit davon entfernt, die Welt sehen zu können, denn wir müssen die Objekte hierzu auch identifizieren. Dies ist die Aufgabe der Mustererkennung. Ein Großteil der Forschung in diesem Bereich beschäftigt sich mit der Frage, wie wir Buchstaben erkennen können. Beispielsweise stellt sich das Problem, wie wir die Darbietung des Buchstabens A als eine Ausprägung des Musters A erkennen. Wir behandeln zunächst die Mustererkennung in Bezug auf die Buchstabenerkennung, um dann zu einer allgemeineren Darstellung der Objekterkennung überzugehen.



■ **Abb. 2.16** Beispiele für Versuche eines Schablonenabgleichs für die Buchstaben A und L. **a** und **c** sind erfolgreiche Versuche, **b** und **d** bis **h** sind erfolglose Versuche. (Nach Neisser, 1967. © 1967 Appleton)

2.2.1 Schablonenabgleich

Die wahrscheinlich offensichtlichste Art der Mustererkennung bedient sich des **Schablonenabgleichs** (*template matching*). Die Wahrnehmungstheorie des Schablonenabgleichs beruht auf der Annahme, dass ein getreues Netzhautbild des Objekts an das Gehirn übermittelt und dann der Versuch unternommen wird, es direkt mit bereits gespeicherten Mustern zu vergleichen. Diese Muster heißen Schablonen. Nach dieser Grundannahme sollte das Wahrnehmungssystem versuchen, das Bild eines Buchstabens mit allen Schablonen zu vergleichen, die es für jeden Buchstaben besitzt, und daraufhin die Schablone mit der besten Übereinstimmung melden. ■ **Abbildung 2.16** zeigt verschiedene Beispiele für erfolgreichen oder erfolglosen Schablonenabgleich. In allen Fällen wird versucht, eine Übereinstimmung zwischen den vom Buchstaben A stimulierten Netzhautzellen und den für eine Buchstabenschablone spezifizierten Netzhautzellen zu erzielen.

Konsequenzen

Die Trennung von Mensch und BOT

Die besondere Natur der visuellen Wahrnehmung des Menschen wurde bei der Entwicklung von CAPTCHAs (von Ahn, Blum & Langford, 2002) angewendet. CAPTCHA steht dabei als Akronym für *Completely Automated Public Turing Test to Tell Computers and Humans Apart*, also den vollständig automatisierten öffentlichen Turing-Test zur getrennten Ansprache von Computern und Menschen. Motiviert sind die CAPTCHAs durch reale Probleme, wie sie zum Beispiel Yahoo beim öffentlichen

E-Mail-Account antrifft. Das Problem liegt darin, dass BOTs (kleine selbständig arbeitende Computerprogramme) sich in solchen Accounts anmelden und dann Spams versenden. Um zu testen, ob hinter der Anmeldung wirklich ein Mensch steht, öffnet solch ein Überwachungssystem ein Schriftbild wie in **Abb. 2.17**. Aber auch wenn der Schablonenansatz beim Erkennen solcher Schriftzüge scheitern mag, sollte man die Computer nicht abschreiben. Einige Computerwissenschaftler sehen dies als

intellektuelle Herausforderung, und es hat ein „Aufrüsten“ mit immer raffinierten Erkennungsprogrammen und immer anspruchsvolleren CAPTCHAs begonnen. Die dabei erfolgreichen Methoden haben allerdings nur noch wenig mit Schablonenabgleich zu tun, sondern arbeiten mit Merkmalsextraktion, um die Items zu definieren (z. B. Mori & Malik, 2003). Sie können die CAPTCHA-Website besuchen und sich an der Forschung auf <http://www.google.com/recaptcha> beteiligen.



Abb. 2.17 Ein Anmeldeformular mit CAPTCHA zur Erkennung unbefugter Zugriffe durch Computerprogramme. Menschen können die Schriftzüge lesen, aber Computerprogramme haben enorme Schwierigkeiten bei der Worterkennung anhand von Schablonenabgleich. (Nachdruck mit Erlaubnis von Luis von Ahn)

Abb. 2.16a zeigt einen Fall, bei dem eine Übereinstimmung erreicht und ein A erkannt wird. **Abb. 2.16b** ist ein Beispiel für einen Fall, bei dem der Input eines L nicht mit der Schablone für A übereinstimmt. Dagegen stimmt L in **Abb. 2.16c** mit der Schablone L überein. Beim Schablonenabgleich können jedoch sehr leicht Probleme auftreten: **Abb. 2.16d** zeigt das Fehlschlagen eines Abgleichs, weil das Bild auf einen falschen Teil der Netzhaut fällt; im Falle von **Abb. 2.16e** entstehen Probleme, weil das Netzhautbild die falsche Größe besitzt. **Abb. 2.16f** verdeutlicht, was geschieht, wenn das Bild nicht die passende Ausrichtung besitzt, und die **Abb. 2.16g** und **Abb. 2.16h** demonstrieren die Schwierigkeiten, die sich ergeben, wenn die Bilder vom Standard-A abweichen.

Obwohl zahlreiche Probleme beim Verfahren des Schablonenabgleichs existieren, ist dies nach wie vor eine der Methoden, die bei der Verarbeitung visueller Informationen durch Maschinen benutzt werden (Ullman, 1996). So wurden Prozeduren zum Drehen, Dehnen oder anderweitigem Verzerren zum Abgleich eines Bildes mit der Schablone entwickelt. Auch bei bildgebenden Verfahren wie fMRI (► Kap. 1) wird der Schablonenabgleich verwendet. Wenn die Forscher behaupten, dass Hirnregionen (wie die in **Abb. 1.15**) Aktivierungsmuster (wie die in **Abb. 1.16**) zeigen, dann behaupten sie typischerweise,

dass die gleichen Regionen in den Gehirnen ihrer Probanden dieses Muster zeigen. Um festlegen zu können, dass es sich um dieselbe Region handelt, kartieren sie die individuellen Gehirne anhand eines Referenzgehirns mit Hilfe eines komplizierten computerbasierten dreidimensionalen Schablonenabgleichs. Allerdings scheinen den Fähigkeiten von Computern beim Schablonenvergleich Grenzen gesetzt zu sein, wenn es um Mustererkennung geht, wie sie anhand der CAPTCHAs untersucht wurden (siehe Kasten „Konsequenzen“).

► **Beim Schablonenabgleich wird versucht, das Ausmaß der Übereinstimmung zwischen einem Muster und einem Stimulus zu erfassen.**

2.2.2 Merkmalsanalyse

Psychologen haben – zum Teil auch wegen der Schwierigkeiten, die im Zusammenhang mit dem Modell des Schablonenabgleichs aufgeworfen werden – die Betrachtungsweise vorgeschlagen, dass die Mustererkennung auf einer **Merkmalsanalyse** (*feature analysis*) beruht. Nach diesem Modell wird jeder Reiz

Kognitive Psychologie

Anderson, J.R. - Funke, J. (Hrsg.)

2013, X, 362 S. 270 Abb., 159 Abb. in Farbe., Hardcover

ISBN: 978-3-642-37391-6