

Max-Planck-Institut für biologische Kybernetik Tübingen

1958 wurde am Max-Planck-Institut für Biologie in Tübingen eine „Forschungsgruppe Kybernetik“ gegründet, die 1960 als selbständige Abteilung Reichardt weitergeführt werden konnte. Aus dieser Abteilung ging 1968 das Max-Planck-Institut für biologische Kybernetik hervor. Die gemeinsamen Interessengebiete der Arbeitsgruppen reichen von den Primärprozessen in Lichtsinneszellen und der Informationsverarbeitung in visuellen und akustischen Systemen über die Struktur und Funktion von Neuronenverbänden bei Insekten und Wirbeltieren bis zur Psychophysik der Sehvorgänge beim Menschen. Die Untersuchungen werden mit neuroanatomischen, histochemischen, elektrophysiologischen, verhaltensphysiologischen, verhaltensgenetischen und psychophysischen Methoden durchgeführt.

Geschäftsführender Direktor:

Prof. Dr. Karl Georg Götz (bis 31. 7. 1994)

Prof. Dr. Kuno Kirshfeld (seit 1. 8. 1994)

Kollegium, Wissenschaftliche Mitglieder:

Prof. Dr. Valentin Braitenberg

(em. seit 1. 7. 1994)

Prof. Dr. Heinrich Bülthoff

Prof. Dr. Karl Georg Götz

Prof. Dr. Kuno Kirshfeld

60 Mitarbeiter, davon 19 Wissenschaftler; 7 Drittmittelbeschäftigte, davon 6 Wissenschaftler; dazu 11 wissenschaftliche Gäste und Stipendiaten.

Fachbeirat/Kuratorium: Prof. Dr. Dr. h. c. mult. Hermann Haken, Stuttgart (Vorsitzender); Bergassessor a. D. Klaus Haniel, Icking; Prof. Sir Bernhard Katz, London; Prof. Dr. Hubert Markl, Konstanz; Prof. Dr. Randolph Menzel, Berlin; Prof. Dr. Heinz Wässle, Frankfurt/M.

Arbeitsgruppe Braitenberg

Arbeitsgebiete

Kleinhirn. Theorie der Kleinhirnfunktion und der Bewegungssteuerung. Quantitative Anatomie der Kleinhirnrinde und Elektrophysiologie an Kleinhirn-Schnittpräparaten und Kleinhirnkulturen. – *Kortex.* Anatomische Untersuchungen zur Konnektivität der Großhirnrinde am Schnittpräparat im Hinblick auf die Hebb'sche Theorie der cell assemblies. Vergleichend anatomische Studien am Kortex verschiedener Säuger. Computersimulationen zur Konnektivität des Kortex. Psychophysische Untersuchungen am Menschen bezüglich der Orientierungsspezifität im visuellen Kortex.

Arbeitsgruppe Bülthoff

Arbeitsgebiete

Raum- und Formwahrnehmung. Wie berechnet unser Sehsystem die Form und Bewegung von Objekten aus den zweidimensionalen Abbildungen auf unserer Retina. Unter Einsatz von Computer-Animationstechnik konnten wir zeigen, daß die Relativbewegung zwischen bewegten Objekten und ihrem Schattenwurf besonders gut geeignet ist, Mehrdeutigkeiten in der Bildinterpretation aufzulösen und beim Betrachter den Eindruck von Tiefe und Bewegung zu erzeugen (*I. Bülthoff, Kersten, Mamassian*). – *Navigation und kognitive Karten.* Bei der Exploration von Labyrinth durchläuft ein Beobachter eine Folge von Korridoren und Verzweigungspunkten, aus deren Kenntnis er eine Repräsentation des Labyrinths aufbauen kann, die dann doch u. a. zielgerichtete Navigation ermöglicht. Ordnet man jedem Korridor zwei Ansichten zu (eine für

jede Richtung), so kann man als Repräsentation den Graphen aus diesen Ansichten benutzen. Diese ansichtsbasierte Navigation wurde in einem künstlichen neuronalen Netzwerk simuliert und auf einem Miniaturroboter in einem realen Labyrinth erprobt. Parallel dazu wurde mit interaktiver Computergrafik ein virtuelles Labyrinth implementiert, in dem sich Versuchspersonen „bewegen“ können. In psychophysischen Experimenten werden damit Lernen und Gebrauch von kognitiven Karten durch Versuchspersonen untersucht (Dartsch, Distler, Huber, Mallot, Schölkopf). – *Höhere Mechanismen des Farbsehens*. Während die frühen Stadien der Farbwahrnehmung sehr gründlich untersucht sind, ist über höhere, kortikale Mechanismen relativ wenig bekannt. In psychophysischen Experimenten untersuchen wir die Eigenschaften dieser Mechanismen und Interaktionen bei der gemeinsamen Verarbeitung von Farbe und anderen elementaren Reizattributen (Form und Bewegung). Die Ergebnisse der Psychophysik werden mit neurophysiologischen Daten aus Einzelzellableitungen in Beziehung gesetzt (Gegenfurtner und Mitarbeiter).

Aktueller Forschungsschwerpunkt

Erkennung von Objekten im Raum

Bei der Verarbeitung visueller Information unterscheidet man gewöhnlich drei Ebenen: Die erste Verarbeitungsebene (*early vision*) umfaßt überwiegend reizabhängige, parallel und ohne besondere Aufmerksamkeitsmechanismen ablaufende Operationen. Sie liefert bildhafte „Karten“ aus lokalen Messungen von Helligkeit, Farbe, Bewegung, Kantenorientierung und Tiefe, sowie einfache Einteilungen der Bildfläche in zusammengehörige Bereiche (Bildsegmentierung, Figur-Grund-Unterscheidung). In der mittleren Verarbeitungsebene (*mid level vision*) werden die nach Ort und Merkmalsart getrennt vorliegenden lokalen Informationen zu einer einheitlicheren Repräsentation integriert, um eine möglichst fehlerfreie Interpretation der dreidimensionalen Welt

aus zweidimensionalen Abbildungen auf unserer Retina zu gewährleisten. Die höchste Verarbeitungsebene (*high level vision*) bilden die Erkennungsleistungen, z. B. die Entscheidung über die Gleichheit eines gesehenen Objekts mit einem bereits bekannten, im Gedächtnis abgespeicherten Objekt. An dieser Stelle werden also Gedächtnisleistungen zum ersten Mal bedeutsam. Während die frühe Bildverarbeitung im Roboter-Sehen wie auch in der neurophysiologischen und psychophysischen Wahrnehmungsforschung vergleichsweise gut verstanden ist, liegen über die höheren Mechanismen nur wenig Kenntnisse vor. Im folgenden wird auf ein wichtiges Teilproblem, nämlich die Erkennung von Objekten im Raum, näher eingegangen. Dabei zeigt sich, daß die Definition von Erkennung als Feststellung der Gleichheit mit einem bekannten Objekt unscharf ist: sie enthält keine Aussage darüber, welche Informationen über das Objekt im Gedächtnis gespeichert sind und somit als Grundlage für den Vergleich herangezogen werden können. Grundsätzlich kann man davon ausgehen, daß die gespeicherte Information (die Objektrepräsentation) im Hinblick auf den tatsächlich zu erwartenden „Entscheidungsbedarf“ optimiert sein wird. Ziel der Informationsverarbeitung ist nicht das Ablegen von Informationen in Speichern, sondern die situationsgerechte Steuerung von Verhalten (*purposive vision*).

Raumkonstanz: „Was ist wo?“

Bewegt man ein Objekt an eine andere Stelle im Bild oder verändert man die Beleuchtung, so verändert das Objekt nicht seine Identität (Abb. 1). Diese „Objektkonstanz“ der dreidimensionalen Szene ist aus Bildern nicht ohne weiteres zu entnehmen. Tatsächlich können Drehungen im Raum oder veränderte Beleuchtung das Bild ein und desselben Objekts bis zur Unkenntlichkeit verändern. Die Feststellung der Objektkonstanz ist daher ein zentrales Problem der visuellen Objekterkennung, nach dessen Lösung der verbleibende Klassifikationsschritt, d. h. die Entscheidung, um

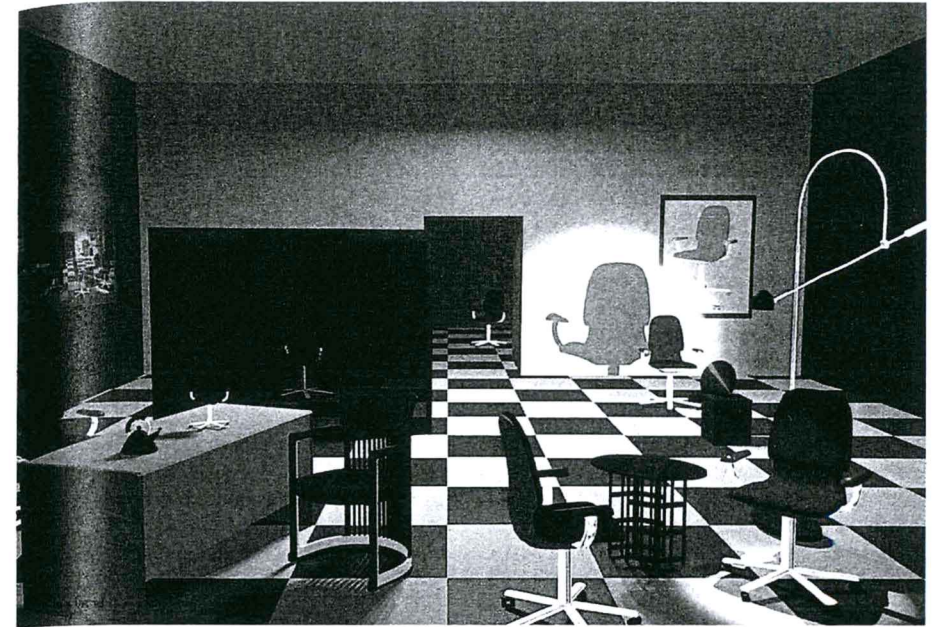


Abb. 1: Illustration verschiedener Probleme, die natürliche und künstliche Gehirne bei der Objekterkennung lösen müssen. *Rauschen*: Der Stuhl hinter der halbdurchsichtigen spanischen Wand ist schwerer zu erkennen als der gleiche Stuhl rechts im Bild. *Verdeckung*: Hinter dem Tisch wird ein Stuhl gesehen, obwohl nur Teile davon im Bild wirklich vorhanden sind. *Größenkonstanz*: Wahrgenommene Größe ist eine Funktion des vom Objekt eingenommenen Sehwinkels und des gesehenen Abstandes. Das Stuhlmodell auf dem Schreibtisch und der echte Stuhl im Nebenzimmer sind im Bild gleich groß, erscheinen aber unterschiedlich, in Übereinstimmung mit ihrer wirklichen Größe in der dreidimensionalen Welt. Umgekehrt erscheint der Stuhl rechts genauso groß wie der im Nachbarzimmer, obwohl sein Bild wesentlich größer ist. *Lagekonstanz*: Verschiedene Ansichten der Stühle werden jeweils als gleiche Objekte erkannt. *Kontext*: Das Stuhlmodell auf dem Schreibtisch wird schlechter erkannt als der Stuhl hinter dem Schreibtisch, der an einer sinnvollen Stelle steht. Dies äußert sich vor allem in der Reaktionszeit des Betrachters. *Beleuchtung*: Unter unterschiedlicher Beleuchtung können die Bilder gleicher Objekte sehr verschieden aussehen, werden jedoch mühelos erkannt. Auch der Schatten an der Wand wird als solcher erkannt und nicht mit einer Sitzgelegenheit verwechselt. *Funktionalität*: Sucht man in der dargestellten Szene eine Sitzgelegenheit, so wird das Stuhlmodell auf dem Schreibtisch und das Bild an der Wand nicht mehr in die gleiche Objektklasse eingeordnet wie die wirk-

chen Stühle. Dagegen gehören jetzt auch der Designer-Stuhl in Bildmitte und u. U. auch die Tischplatte dazu.

welche Art von Objekt es sich handelt, meist keine großen Schwierigkeiten mehr bereitet. Als Spezialfall betrachten wir hier die Raumkonstanz, genauer die Erkennung von Objekten aus unterschiedlichen Blickwinkeln. Nach einer gängigen Definition bedeutet „Sehen“, aufgrund von optischen Eindrücken zu erkennen, was sich wo befindet; in diesem Sinne ist die Raumkonstanz die Lösung des „was?“-Problems. Eine Übersicht über verschiedene weitere Probleme der Objekterkennung gibt Abbildung 1.

Zwei Verfahren zur Erzeugung der Raumkonstanz

Wir betrachten eine Situation, in der ein Objekt oder eine Anzahl von Ansichten dieses Objekts bekannt ist. Die Aufgabe besteht dann darin, von einer weiteren Ansicht zu entscheiden, ob sie dasselbe oder ein anderes Objekt zeigt. Diese Auf-

gabe ist gleichbedeutend damit, aus den bekannten Ansichten eine neue zu erzeugen, die der fraglichen Ansicht möglichst ähnlich ist. Gelingt dies durch Transformationen, die der Drehung des Objekts im Raum entsprechen, so handelt es sich tatsächlich um eine andere Ansicht des bereits bekannten Objekts.

Modellbasierter Ansatz. Ein denkbarer Mechanismus für die Raumkonstanz besteht darin, die Frage nach dem „was?“ vollständig vom „wo?“ zu trennen. Dazu muß man zu einem gegebenen Bild eines Objekts eine Transformation finden, die das Bild in eine invariante Beschreibung dieses Objekts überführt. Für die zweidimensionale Mustererkennung verschiebt man die zu klassifizierende Grauwertverteilung z. B. so, daß der Schwerpunkt des Musters in Bildmitte zu liegen kommt. Für die Kompensation von Drehungen in zwei oder drei Dimensionen sind ähnliche Vorschläge unter dem Namen „mentale Rotation“ gemacht worden: Die Idee ist hierbei, daß man zunächst ein dreidimensionales Modell des Objekts erzeugt, welches man dann „im Geiste“ in eine Standardposition drehen kann. Bei derartigen Transformationsansätzen führt man also zwischen dem Bild (den Rohdaten) und der Klassifikationsinstanz eine weitere Repräsentation ein, die *invariant* gegen die betrachtete Streuung der abgebildeten Szene ist. Auf der durch die mentale Rotation erreichten Repräsentationsstufe ist die Information über die ursprüngliche Lage des Objekts nicht mehr vorhanden.

Bildbasierter Ansatz. Der zweite Ansatz geht von der Überlegung aus, daß einerseits die für die mentale Rotation erforderlichen räumlichen Objektmodelle schwer zu konstruieren sind und daß andererseits die vollständige Abtrennung der Orientierungsinformation von der Objekterkennung biologisch wenig plausibel ist. Will man z. B. ein Objekt nicht nur benennen, sondern es ergreifen oder mit ihm eine Handlung ausführen, so muß man über beide Informationen in Kombination verfügen. Diese Anforderung erfüllt der bildbasierte Ansatz, bei dem Objekte durch eine Anzahl von An-

sichten aus verschiedenen Blickwinkeln repräsentiert sind. Die Vorhersage neuer Ansichten erfolgt hier durch eine Interpolation; bei der Klassifikation eines Objekts wird das Bild mit bekannten Ansichten verglichen, indem man z. B. feststellt, ob und wie gut es aus ihnen durch diese Interpolation rekonstruiert werden kann. Ergebnis des bildbasierten Verfahrens ist nicht die Erkennung eines Objekts, sondern die eines Objekts unter einem bestimmten Blickwinkel.

In Zusammenarbeit mit T. Poggio (MIT, Cambridge, USA) werden in der Arbeitsgruppe zwei Varianten des bildbasierten Verfahrens genauer untersucht:

1. *Pixelweise Interpolation der Grauwertverteilungen in einem dreischichtigen, vorwärtsgekoppelten neuronalen Netzwerk.* Hierbei faßt man die Ansichten als Vektoren in einem hochdimensionalen Merkmalsraum auf. Jeder Knoten der mittleren Netzwerkschicht ist auf eine bestimmte Ansicht „abgestimmt“ und reagiert auf eine gezeigte Ansicht umso stärker, je ähnlicher sie der von ihm repräsentierten Ansicht ist. Die Interpolation erfolgt durch einfache Überlagerung der in den Netzwerkknoten gespeicherten Ansichten, gewichtet mit der Ähnlichkeit des Ausgangsbildes mit dieser Ansicht, d. h. mit der Erregung des Netzwerkknotens.

2. *Bestimmung eines Verschiebungs-Vektorfeldes zwischen den beiden zu vergleichenden Bildern mit anschließender Interpolation korrespondierender Punkte.* Interpoliert man nach dem oben genannten Verfahren etwa die Frontalansicht und das Halbprofil eines Gesichtes, so wird das Ergebnis u. a. zwei Abbilder der Nase enthalten, so, als hätte man beide Bilder mit zwei Diaprojektoren überlagert. In diesem Fall ist es günstiger, vor der Interpolation die lokale Verschiebung der Nase (und anderer Bildmerkmale) zu bestimmen und dann ein Bild davon an die mittlere Position zu setzen. Interpoliert werden hier also nicht nur die Bildintensitäten, sondern auch die Positionen der Bildmerkmale. Die Positionsdifferenzen der Bildmerkmale muß man dazu vorher bestimmen, z. B. mit Hilfe eines Korrespon-

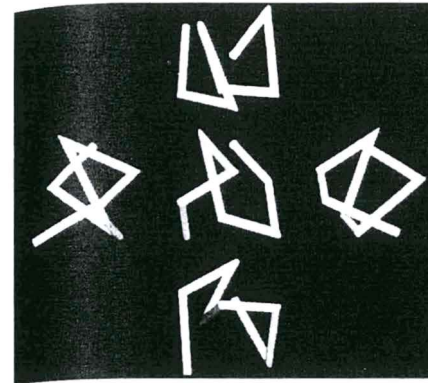


Abb. 2: Blickpunktabhängigkeit der Objekterkennung. Aufgabe: Sind hier unterschiedliche Objekte dargestellt, oder handelt es sich nur um verschiedene Ansichten des gleichen Objekts? Lösung: Die Abbildung zeigt fünfmal dasselbe Objekt, das jeweils um die vertikale und horizontale Achse gedreht wurde. Um die Aufgabe zu lösen, haben Sie vermutlich eine „mentale Rotation“ durchgeführt, für die Sie mehrere Sekunden benötigt haben. Dieser zeitaufwendige Mechanismus ist für schnelle Objekterkennung ungeeignet und wird deshalb vom Menschen nur in Spezialfällen eingesetzt.

denz-Algorithmus, wie er aus dem Stereo- oder Bewegungssehen bekannt ist. Dabei fließt durch die Feststellung der Verschiebungen (Disparitäten) implizit auch Information über die räumliche Struktur des Objekts ein, doch wird daraus kein explizites dreidimensionales Modell konstruiert.

Auch bei Kenntnis der räumlichen Struktur eines Objekts werden neue Ansichten schlecht erkannt

Ein experimenteller Zugang zu der Frage, ob die Raumkonstanz bei der Objekterkennung auf der Konstruktion eines dreidimensionalen Modells für die mentale Rotation beruht, ergibt sich aus folgender Überlegung: Mentale Rotation setzt die Konstruktion eines dreidimensionalen Objektmodells voraus. Wenn also das gezeigte Bild hinreichende Information über die dreidimensionale Struktur des Objekts enthält (z. B. durch stereoskopische Darbietung oder durch Ausnutzung des kineti-

schen Tiefeneffekts), so sollten aus der Rekonstruktion eines dreidimensionalen Modells beliebige zweidimensionale Ansichten vorhergesagt werden können. Die experimentelle Frage lautet also, ob die Kenntnis der dreidimensionalen Struktur ausreicht, um neue Ansichten eines Objekts zu erschließen.

Für die psychophysischen Experimente wurden unbekannte und für die Versuchspersonen bedeutungslose Objekte erzeugt, indem zufällig im Raum verteilte Punkte durch eine Art geknicktes Rohr miteinander verbunden wurden (Abb. 2). Versuchspersonen betrachteten diese Objekte aus einer gegebenen Ansicht, wobei das Objekt in einem Winkelbereich von ($\pm 15^\circ$) aus der Bildebene gedreht wurde, um einen räumlichen Eindruck zu erzeugen (kinetischer Tiefeneffekt). Danach wurden für kurze Zeit (< 15 sec) stärker verdrehte Ansichten des gleichen Objekts zusammen mit neuen Objekten gezeigt und die Versuchsperson sollte entscheiden, ob das bekannte oder ein neues Objekt vorlag. Die Ergebnisse zeigen, daß neue Ansichten nur in einem Winkelbereich von $\pm 30^\circ$ erkannt werden; weiter gedrehte Ansichten können nicht mehr zuverlässig von Bildern neuer, ähnlicher Objekte unterschieden werden. Dieses Ergebnis ist mit den oben skizzierten bildbasierten Ansätzen zur Raumkonstanz besser vereinbar als mit der Annahme eines dreidimensionalen Objektmodells und der mentalen Rotation.

Symmetrie und virtuelle Ansichten

Natürliche Objekte weisen oft Symmetrien auf, die für die Erzeugung der Raumkonstanz nutzbar gemacht werden können. Hat man zum Beispiel eine Ansicht eines bilateral symmetrischen Objekts, so kann man zeigen, daß das Spiegelbild dieser Ansicht wieder eine mögliche Ansicht desselben Objekts ist. Da man das Spiegelbild in der Bildebene, d. h. ohne Kenntnis der räumlichen Struktur des Objekts erzeugen kann, kennt man mit jeder Ansicht eines symmetrischen Objekts automatisch weitere, „virtuelle“ Ansichten. Die Information für diese

virtuelle Ansicht basiert dabei gewissermaßen auf der Annahme einer räumlichen Symmetrie des Objekts. Bilaterale Symmetrie kommt in der Technik wie in der Natur häufig vor und zeichnet dabei gerade besonders bedeutsame „Objekte“ wie Tiere und Menschen aus. Die Rolle der Symmetrie bei der Erkennung von Gesichtern wird z.Zt. in psychophysischen Experimenten untersucht.

Verwendet man in dem oben beschriebenen Extrapolationsexperiment symmetrische Rohr-Objekte, so steigt die Erkennungsrate im Bereich der virtuellen Ansicht an, obwohl diese nicht gezeigt wurde. Dieses Ergebnis unterstützt die bildbasierte Theorie der Raumkonstanz und zeigt, daß die zusätzliche Information in symmetrischen Objekten ausgenutzt wird.

Bevorzugte Ansichten bekannter Objekte

Bei den bisher betrachteten „Zufallsobjekten“ wurde darauf geachtet, daß alle gezeigten Ansichten mehr oder weniger gleichwertig waren. Die Punktwolken, welche die Knickpunkte der Rohrobjekte definieren, waren z.B. so gewählt, daß die Punktdichte in allen Ansichten möglichst homogen war und keine Verdeckungen von Punkten auftraten. Im allgemeinen ist eine solche Gleichwertigkeit der Ansichten nicht gegeben. So können verschiedene Ansichten desselben Objekts ganz unterschiedlich viel Information enthalten, je nachdem, welche Teile gerade verdeckt sind. Zwei wichtige Unterscheidungen, von denen die erste geometrisch, die zweite dagegen empirisch motiviert ist, sind:

1. „Stabile“ versus „instabile“ Ansichten: Die Oberfläche von Objekten läßt sich in der Regel in glatte Flächenstücke zerlegen, die von Kanten begrenzt sind; Beispiele sind die Facetten eines Polyeders oder die Innen- und Außenseite sowie der Henkel einer Tasse. Dreht man nun ein solches Objekt um einen kleinen Winkel, so kann sich das Bild u.U. qualitativ verändern, indem nämlich eines dieser Flächenstücke neu sichtbar wird oder ganz verschwindet. Ansichten, bei denen alle (oder die meisten)

Drehungen zu solchen qualitativen Änderungen führen, heißen *instabil*. Sie bilden nur einen kleinen Anteil aller möglichen Ansichten, und ihr Auftreten in natürlichen Bildern ist in diesem Sinne unwahrscheinlich. Ansichten, die sich bei kleinen Drehungen nur wenig (d.h. stetig) verändern, heißen *stabil*.

2. *Kanonische (mustergültige) Ansichten*: In psychophysischen Experimenten wählen Versuchspersonen reproduzierbar bestimmte Ansichten als besonders typisch oder treffend zur Charakterisierung eines Objekts aus. Bei diesen Ansichten handelt es sich in der Regel um stabile Ansichten im Sinne der obigen Unterscheidung; darüberhinaus zeigen sie die für die Erkennung des Objekts besonders wichtigen Details.

Im Wahrnehmungsexperiment scheint die „Mustergültigkeit“ einer Ansicht A_0 von ihrem Informationsgehalt für die Erkennungsaufgabe abzuhängen. Ein einfaches Modell für die Beurteilung dieses Informationsgehaltes benutzt einen linearen Klassifikator, um – ausgehend von einer vorgegebenen Ansicht – benachbarte Ansichten des gleichen Objekts von Ansichten anderer Objekte zu unterscheiden. Der (für den Klassifikator nutzbare) Informationsgehalt der ursprünglichen Ansicht ist dann durch die zugehörige Erkennungsleistung, d.h. die Allgemeingültigkeit dieser Ansicht definiert. Abbildung 3 zeigt das Ergebnis anhand von Alltagsobjekten. Die Ansicht, die die beste Generalisierung von einer gegebenen zu neuen Ansichten erlaubt, ist in der Abbildung rechts unten dargestellt. Psychophysische Experimente zur Bestimmung der kanonischen Ansichten der gleichen Objekte werden zur Zeit durchgeführt.

Greifbewegungen

Das Erkennen von Objekten wurde bisher als die Feststellung der Gleichheit mit einem bekannten Objekt definiert. Im realen Verhaltenskontext ist dies jedoch nur ein kleiner Teil des Problems. Will man z.B. ein Objekt ergreifen, so ist dazu weniger die Klassifikation als die Einschätzung der Form und die Auswahl von Greifpunkten

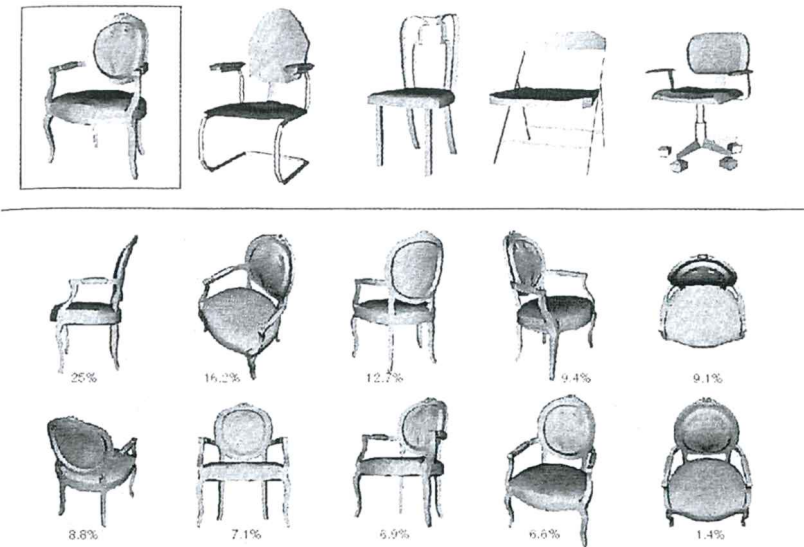


Abb. 3: Informationsgehalt verschiedener Ansichten eines Stuhls. Obere Reihe: Einige Beispiele für die verwendeten Stuhltypen, unter denen jeweils der Barockstuhl von einem automatischen Erkennungsalgorithmus erkannt werden soll. Insgesamt wurden 24 verschiedene Stuhlmodelle verwendet. Unten: Verschiedene Ansichten des Barockessels wurden entsprechend der Fehlerrate (Prozentangaben) geordnet, die der automatische Erkennungsalgorithmus macht, wenn er nur mit einer Ansicht (umrahmt, oben) trainiert worden ist. Die in der Abbildung ganz rechts unten dargestellte Ansicht erlaubt die beste Erkennungsrate (Fehlerrate 1,4%).

erforderlich. In Zusammenarbeit mit A. Blake (Oxford University) und M.A. Goodale (University of Western Ontario, Canada) wurde das Ergreifen flacher, unregelmäßig berandeter Scheiben mit zwei Fingern untersucht. Ein Kriterium für die Güte eines solchen Zwei-Finger-Griffs ist das dabei erzeugte Drehmoment: Der Griff ist stabil, wenn die von den beiden Fingern ausgeübten Kräfte diametral entgegengesetzt und auf den Schwerpunkt gerichtet sind. In psychophysischen Experimenten wurde diese einfache Hypothese zunächst experi-

mentell überprüft. Dabei zeigt sich, daß Versuchspersonen tatsächlich in der Lage sind, anhand der Objektansicht zwei geeignete Greifpunkte auszuwählen und die Hand gezielt zu diesen Punkten zu bewegen. Zusätzlich zu den oben genannten objektabhängigen Auswahlkriterien wird dabei noch die erforderliche Drehung des Handgelenks und des Unterarms berücksichtigt. Die Leistungen von zwei Patienten mit Hirnverletzungen im Okzipitallappen wurden mit denen von gesunden Versuchspersonen verglichen. Dabei zeigte sich eine interessante „Dissoziation“ der Leistungsfähigkeit für verschiedene Aufgaben: während ein Patient die Objekte richtig greifen, nicht aber die Gleichheit oder Ungleichheit zweier Objekte anhand ihrer Bilder beurteilen konnte (Agnosie), war es in einem zweiten Fall gerade umgekehrt (Ataxie). Diese Ergebnisse zeigen, daß die neuronalen Verarbeitungszentren für verschiedene Teilaufgaben bis zu einem gewissen Grad separat operieren. Dies bedeutet im Hinblick auf das Studium der visuellen Informationsverarbeitung, daß – abhängig von

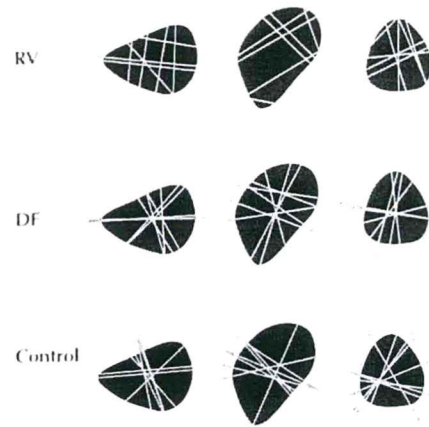


Abb. 4: Dissoziation von Handlung und Wahrnehmung. Wie greift man am besten flache Scheiben mit Daumen und Zeigefinger? Die optimalen Greifpunkte (dargestellt durch die Verbindungslinien der Kontaktpunkte) stehen in etwa senkrecht auf dem Rand und verlaufen in der Nähe des Schwerpunktes (Control). Ein neurologischer Patient (RV) greift die Scheiben an ungeeigneten Stellen (optische Ataxie), obgleich er durchaus in der Lage ist, die Scheiben in einem Erkennungsexperiment zu unterscheiden. Ein anderer Patient (DF) ergreift die Objekte zuverlässig, kann sie aber erstaunlicherweise nicht unterscheiden (visuelle Agnosie).

den jeweiligen Aufgaben (hier Greifen und Erkennen) – unterschiedliche Verarbeitungen durchgeführt werden und unterschiedliche Repräsentationen vorhanden sind.

Schlußbemerkung

Das Erkennen von Objekten im Raum ist ein für das Verständnis der menschlichen Wahrnehmung wie auch für technische Anwendungen außerordentlich wichtiges Problem der visuellen Informationsverarbeitung. Die dargestellten Ergebnisse zeigen, daß die bisher übliche Zerlegung des Problems in ein (sehr schwieriges) Konstanzproblem und ein (fast triviales) Klassifikationsproblem bei der visuellen Wahrnehmung des Menschen nicht verwirklicht ist. Vielmehr tritt die Erkennung schon auf dem Bildniveau ein, so daß das Konstanzproblem gar nicht in voller Allgemeinheit

gelöst werden muß. Neben der größeren Praktikabilität, die es auch für automatische Sehsysteme (Robotik) interessant macht, hat dieses Vorgehen den Vorteil, ohne aufwendige Modelle der räumlichen Struktur der Objekte auszukommen. Es fügt sich somit in einen verhaltensorientierten Ansatz der visuellen Wahrnehmung, der das Ziel der Informationsverarbeitung nicht im Anlegen von Repräsentationen, sondern in der Unterstützung von Handlungen sieht. Zukünftige Experimente müssen zeigen, für welche Verhaltensleistungen räumliche Objektmodelle wirklich erforderlich und im visuellen System realisiert sind (Blanz, H. Bülthoff, Opitz, O'Toole, Troje, Vetter).

Arbeitsgruppe Götz

Arbeitsgebiete

Neuronale Grundlagen des Orientierungsverhaltens der Fliege. Bewegungsvorgänge im Gesichtsfeld der Fliege erzeugen in den bewegungsrichtungsempfindlichen Schichten des Nervensystems landkartenartige Erregungsmuster, die an identifizierbaren Einzelzellen des visuellen Systems in vivo sichtbar gemacht werden können. Als „Prozessrechner mit Parallelstruktur“ ist dieser Hirnbereich an der Lauf- und Flugsteuerung beteiligt. – Suchstrategien von normalen oder erblich gestörten Fliegen in der Laufarena und im Flugsimulator können Aufschluß über die Datenverarbeitung bei der Orientierung im Raum und bei der Auswahl bevorzugter Zielobjekte geben. – Die meisten der identifizierbaren Einzelzellen des visuellen Systems reagieren auf Eigenbewegungen um bestimmte Körperachsen. Bei der Regelung der Kopfstellung und Körperhaltung wird die visuell wahrgenommene Eigenbewegung mit entsprechenden Signalen der mechanosensorischen Systeme kontextabhängig verrechnet. – Die Suche nach den höheren Funktionen des Insektenhirns konzentriert sich auf zwei prominente Bereiche: Wird die Entwicklung der beiden „Pilzkörper“ im Larvenstadium

unterbunden, verliert die Fliege ihre chemosensorische Lernfähigkeit. – Fliegen mit erblich gestörtem Laufverhalten weisen vergleichsweise häufig Strukturanomalien im „Zentralkomplex“ auf. Die Beteiligung dieses Hirnbereichs an der Aktivierung und Modulierung der vom Bauchmark gesteuerten Beintätigkeit ist die Grundlage einer umfassenden Funktionsanalyse.

Arbeitsgruppe Kirschfeld

Arbeitsgebiete

Primärprozesse und Phototransduktion in Photorezeptoren von Insekten. Ionenregu-

lation im Nervensystem von Insekten. Vergleichende Physiologie der Verarbeitung visueller Information bei Insekten und Wirbeltieren. Verarbeitung akustischer Signale im Nervensystem von Schleiereulen. Einfluß von Allgemeinanästhetika auf das Nervensystem von Insekten und Wirbeltieren. Psychophysik des Seh- und Hörsystems vom Menschen.

Anschrift

Spemannstrasse 38
72076 Tübingen
Tel. 07071/60 15 80
Telefax 07071/60 15 75