

Künstliche Retina

Analog-Chips kopieren die Natur

Der Versuch, die Funktionalität der Netzhaut (Retina) des Auges auf künstliche Systeme zu übertragen, begründet sich einerseits im Bestreben, die Sehzentren des Hirns zu modellieren, andererseits aber auch auf der Erkenntnis, dass die Retina künstlichen Bildaufnahme- und -verarbeitungssystemen bis anhin in mancher Hinsicht überlegen ist. So verfügt sie über einen wesentlich grösseren Helligkeitsdynamikbereich, führt erste Datenverarbeitungsschritte effizienter aus und nutzt ihre Datenübertragungsbreite durch Kompression der Bilddaten besser als herkömmliche künstliche Systeme.

Die meisten Informationen über die Umwelt werden vom Menschen und von vielen Tierarten durch das optische System gewonnen. Schon sehr früh hat der Mensch deswegen begonnen, Apparate zu bauen, die ihm behilflich sein können, optische Reize besser wahrnehmen zu können. Anfänglich dienten diese Apparate ausschliesslich der Abbildung. Mit der Einführung der Fotografie im letzten Jahrhundert wurde dann auch die Speicherung von Bildern möglich. Die künstliche Verarbeitung visueller Daten erlebte jedoch erst im Zeitalter der integrierten Schaltungen und Computer ihren Durchbruch. Und vielleicht erst jetzt, nach jahrzehntelangen Anstrengungen in der digitalen Bildverarbeitung, deren Erfolge bislang weit hinter den ursprünglichen Erwartungen zurückgeblieben sind, können wir die Leistung richtig würdigen, die unser Gehirn vollbringt, wenn es aus dem immensen Datenfluss, der durch unsere Augen in das Gehirn strömt, anscheinend mühelos die wesentlichen Informationen extrahiert und speichert. Dies ist noch beeindruckender, wenn man sich die Kompaktheit des Gehirns vergewärtigt sowie seinen geringen Leistungsverbrauch, der etwa so gross ist wie der einer Glühbirne.

Künstliche Systeme zur Aufnahme und Verarbeitung von Bilddaten sind bis heute viel weniger leistungsfähig als biologische. Dies ist insbesondere in natürlichen Umgebungen augenfällig, auf die biologische Systeme infolge ihrer langen Entwicklungszeit optimiert sind. Deshalb überrascht es kaum, dass nicht nur Naturwissenschaftler, sondern auch Ingenieure schon vor geraumer Zeit begonnen haben, sich mit biologischen Methoden auseinanderzusetzen. Als vor-

derstes Glied in der Kette der visuellen Datenverarbeitung hat das Auge natürlich die Aufmerksamkeit zuerst auf sich gelenkt. Die ältesten und am weitesten verbreiteten künstlichen Abbildungsmethoden sind derjenigen des menschlichen Auges prinzipiell sehr ähnlich und mittlerweile durch jahrhundertlange Entwicklung sogar überlegen. Bei der Bildaufnahme jedoch, das heisst bei der Umwandlung von elektromagnetischen Schwingungen in ein elektrisches Signal, können die heutigen Standardsysteme (elektronische Kameras) in vielen Belangen nicht mehr mithalten. Und schon für die ersten Bildverarbeitungsschritte wird typischerweise ein Computer zur Hilfe genommen, der Sekunden bis Minuten braucht, um Bildverarbeitungsoperationen durchzuführen, die von der Netzhaut mit minimalem Leistungsaufwand in einigen Millisekunden realisiert werden, obwohl die biologischen Schaltkreise etwa eine Million Mal langsamer sind als die eines Computers und zudem eine erhebliche statistische Variabilität aufweisen. Aus diesem Grunde ist in den letzten Jahrzehnten die Retina als Bildaufnahme- und -vorverarbeitungssystem vermehrt ins Zentrum des Interesses gerückt.

Simulation oder Emulation?

Die Ursache für diese gewaltige Diskrepanz zwischen rechnergestützter und biologischer Bildverarbeitung liegt hauptsächlich darin, dass – im Gegensatz zu Abbildung und Bildaufnahme – Struktur und Datenrepräsentation solcher Systeme grundverschieden sind. Anders als ein computerbasiertes System ist die Netzhaut topographisch dem Bild ange-

passt, das sie aufnimmt. Die Bildverarbeitungsoperationen, die die Netzhaut ausführt, benötigen nur eine lokale Vernetzung der Datenverarbeitungselemente und ermöglichen deswegen kurze Signalwege und somit eine schnelle parallele Datenverarbeitung auf kleinstem Raum. Voraussetzung dafür ist, dass die Datenverarbeitung und -zwischenspeicherung direkt am Ort der Bildaufnahme geschieht. Ein System, das aus einer Videokamera und einem Computer besteht, kann diese Voraussetzung nicht erfüllen, da Aufnahme, Speicherung und Verarbeitung der Daten konzeptuell und räumlich getrennt sind.

Bildsignale werden als analoge physikalische Grössen von einem Sensor empfangen und in ebensolche Grössen transformiert. Innerhalb der Retina wird diese Repräsentationsform für Signale und gespeicherte Werte beibehalten, während ein konventioneller Prozessor die Sensorsignale erst in digitaler Form verarbeiten kann. Durch die analoge Datenverarbeitung der Netzhaut können physikalische Prozesse viel direkter und deshalb effizienter ausgenutzt werden als von einem

Adresse des Autors

Dr. Jörg Kramer, Institut für Neuroinformatik, Universität und ETH Zürich, Winterthurerstrasse 190, 8057 Zürich

digitalen Prozessor, dessen ineffiziente Kodierungen sich natürlich in Grösse, Leistungsverbrauch und Komplexität der Schaltungen niederschlagen.

Die Reproduktion der Beziehung zwischen einem Eingangs- und Ausgangssignal eines Systems durch ein anderes System wird Simulation genannt. Ein System aus Videokamera und Computer kann also zum Beispiel dazu verwendet werden, die Netzhaut zu simulieren. Wenn aber das ursprüngliche System auf dem Niveau der einzelnen Datenverarbeitungselemente kopiert wird, also ein struktureller und funktioneller Isomorphismus besteht, wird dies Emulation genannt. Aufgrund der vorangehenden Betrachtungen können wir also annehmen, dass eine Emulation der Netzhaut effizienter ist als eine Simulation. Eine Emulation setzt allerdings voraus, dass man die Funktionsweise des ursprünglichen Systems versteht und nicht nur seine Übertragungsfunktion kennt. Bei der Netzhaut ist dies im Gegensatz zu den

meisten anderen Hirnregionen weitgehend der Fall.

Aufbau der Retina

Dass die Netzhaut zu den am besten verstandenen Teilen des Hirns gehört, liegt in erster Linie daran, dass sie sich relativ gut untersuchen lässt: Sie ist peripher angeordnet und daher leicht zugänglich für anatomische und physiologische Analysen. Zudem besteht sie nur aus einer kleinen Anzahl lokal vernetzter Zelltypen, die von anderen Hirnregionen klar abgegrenzt sind. Diese Eigenschaften und die Tatsache, dass typische Eingangssignale gut bekannt und kontrollierbar sind, erleichtern die Untersuchung der Retina.

Die Retinae der meisten Wirbeltiere sind sich sehr ähnlich in Struktur und Signalcharakteristik. Sie enthalten fünf Hauptzelltypen, die in drei Schichten angeordnet sind (Bild 1). Die äussere Schicht besteht aus Photorezeptorzellen, die Lichtreize in elektrische Signale verwandeln, die mittlere aus Bipolar-, Horizontal- und Amakrinzellen, die die Signale analog verarbeiten, und die innere Schicht aus Ganglienzellen, deren Ausgangssignale als Impulse durch den Sehnerv in die höheren Zentren der Signalverarbeitung weitergeleitet werden. Die Photorezeptoren sind auf der dem Lichteinfall abgewandten Seite der Netzhaut angeordnet; die Retina wird also von der Rückseite her beleuchtet.

Es werden zwei Typen von Photorezeptorzellen unterschieden. Die Zapfenzellen sind über die ganze Netzhaut verteilt. In der Fovea centralis, einem Bereich im Zentrum der Netzhaut mit einem Durchmesser von etwa einem halben Millimeter, ist dieser Zelltyp mit einer besonders grossen Dichte vertreten. Der Zellabstand von etwa $2,5\ \mu\text{m}$ in der Fovea centralis limitiert unsere maximale Sehauflösung auf etwa eine halbe Bogenminute. Wegen der geringeren Zapfendichte am Rand der Netzhaut verschlechtert sich die Sehauflösung in der Peripherie des Sehfeldes auf etwa ein Grad. Die Zapfenzellen sprechen auf Reize im Tageslichtbereich an. Sie können in drei Klassen eingeteilt werden, die in verschiedenen – teilweise sich überlappenden – Spektralbereichen empfindlich sind mit mittleren Wellenlängen, die blauem, grünem und orangem Licht entsprechen.

Die Stäbchenzellen, die sich beim Menschen auf der ganzen Retina ausser in der Fovea finden, haben eine breitbandigere spektrale Empfindlichkeit. Sie kommen im Dämmerlicht und nachts zum Einsatz und sättigen bei starker Beleuchtung. Stäbchen sind so lichtempfindlich, dass sie fast einzelne Photonen detektieren können. Die Netzhaut hat etwa 125 Millionen Photorezeptoren, von denen die überwiegende Mehrheit Stäbchen sind. Photorezeptoren wandeln eine Änderung in der einfallenden Lichtintensität logarithmisch in eine Änderung ihres Ausgangspotentials um. Sie haben

eine relativ kurze Ansprechzeit und zeigen eine langsame Adaptierung. Ein Photorezeptor ist über etwa fünf Grössenordnungen in der Beleuchtungsstärke empfindlich; mit Zapfen und Stäbchen zusammen deckt das Auge also etwa zehn Dekaden ab, von Sternenlicht bis zu gleissendem Sonnenlicht. Im Gegensatz zu fast allen anderen Nervenzellen, die auf einen Reiz mit Depolarisation und Ausschüttung von Neurotransmitter an chemischen Synapsen reagieren, werden Photorezeptoren in Wirbeltieren durch Stimulation hyperpolarisiert. Das bedeutet, dass sie in der Dunkelheit aktiv sind und diese Aktivität bei Beleuchtung abnimmt. Benachbarte Photorezeptoren können durch elektrisch resistive Verbindungen (Gap junctions) miteinander gekoppelt sein, so dass eine lokale Mitteilung des Signals entsteht.

Die äussere und die mittlere Zellschicht der Retina werden durch die äussere plexiforme Schicht getrennt, die die Kontaktierungsstellen zwischen Photorezeptor-, Bipolar- und Horizontalzellen enthält.

Bipolarzellen sind direkt an eine oder mehrere Photorezeptorzellen gekoppelt, empfangen aber auf indirektem Weg auch Signale von weiter entfernten Photorezeptorzellen. Der Bereich der Netzhaut, über den Lichtreize eine bestimmte Bipolarzelle beeinflussen, wird rezeptives Feld dieser Zelle genannt. Es gibt zwei Typen dieser Zellart. Man nennt sie On-Zentrum- und Off-Zentrum-Zellen, je nachdem, ob sie depolarisiert (erregt) oder hyperpolarisiert (gehemmt) werden, wenn die direkt an sie gekoppelten Photorezeptorzellen beleuchtet werden. In beiden Fällen wirkt das indirekte Signal dem direkten entgegen, so dass eine On-Zentrum-Zelle erregt wird, wenn ein Lichtreiz auf die Mitte ihres rezeptiven Feldes projiziert wird, und gehemmt wird, wenn der Lichtreiz auf dessen Peripherie fällt. Bei einer Off-Zentrum-Zelle ist das genau umgekehrt. Das rezeptive Feld einer Bipolarzelle besteht in der Fovea centralis aus einigen wenigen Photorezeptorzellen, wächst aber auf einen Durchmesser von bis zu zwei Millimetern in der Peripherie, wo sich die rezeptiven Felder benachbarter Bipolarzellen also zu einem grossen Teil überlappen.

Horizontalzellen stellen über chemische Synapsen laterale Verbindungen zwischen Photorezeptorzellen oder Bipolarzellen her, sind untereinander aber wie die Photorezeptoren über Gap junctions verbunden. Sie sind dünner gesät als Bipolarzellen und reagieren langsamer auf einen Reiz als Photorezeptorzellen. Wie die Photorezeptorzellen werden sie durch

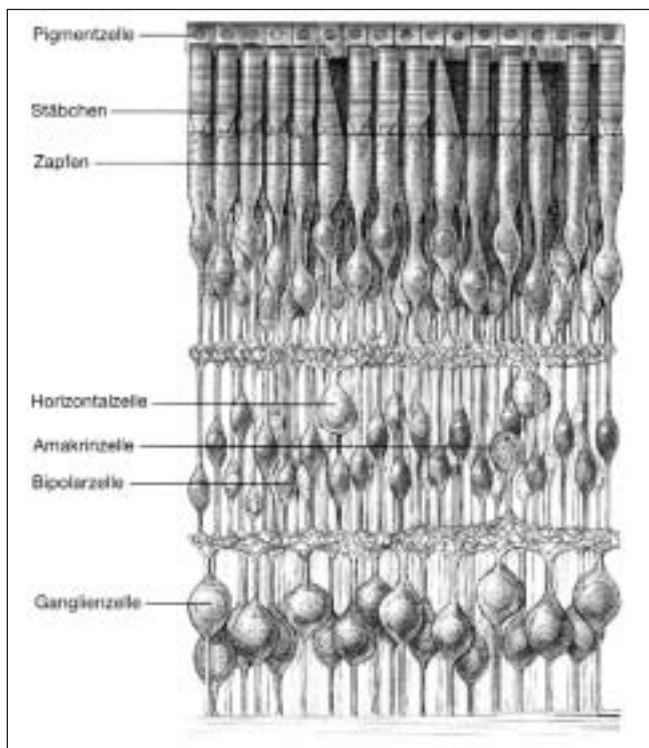


Bild 1 Schematischer Querschnitt durch die Netzhaut

Sie besteht aus drei Zellschichten, die miteinander (vertikal) und jede unter sich (lateral) vernetzt sind. Aus: Eye, Brain and Vision von David H. Hubel. Scientific American Library 1988. Mit freundlicher Genehmigung von W.H. Freeman & Company.

Stimulation hyperpolarisiert. Da die rezeptiven Felder der Bipolar- und Horizontalzellen etwa gleich gross sind, wird angenommen, dass die Horizontalzellen mindestens teilweise für die Wechselwirkungen auf dem indirekten Weg verantwortlich sind. Photorezeptor-, Bipolar- und Horizontalzellen generieren keine Aktionspotentiale wie die meisten anderen Neuronen, sondern leiten ihre Signale durch analoge graduierte Potentialänderungen weiter. Die Signalstärke wird also in der Amplitude der Potentialänderung kodiert. Diese Methode funktioniert bei Nervenzellen nur über kleinere Strecken, da die Verluste und Zeitkonstanten sonst zu gross würden. Die Länge der lateralen Verbindungen innerhalb der Netzhaut entspricht etwa der Grösse der rezeptiven Felder. Die längsten Verbindungen sind also nur etwa einen Millimeter lang.

Zwischen der mittleren und der inneren Zellschicht der Retina liegt die innere plexiforme Schicht, in der Bipolar-, Amakrin- und Ganglienzellen miteinander verbunden sind.

Die Ganglienzellen haben auch einen Zentrum-Umfeld-Aufbau und ähnliche rezeptive Felder wie die Bipolarzellen. Ihre Polarität ist dabei die gleiche wie die der sie innervierenden Bipolarzellen. Neben On-Zentrum- und Off-Zentrum-Zellen existiert auch ein On-Off-Typus, der an On- und Off-Zentrum-Bipolarzellen gekoppelt ist. Im Gegensatz zu Bipolarzellen erzeugen Ganglienzellen Aktionspotentiale, die über den Sehnerv und das Chiasma opticum, wo Signale von beiden Augen zusammenkommen, zum nächsten Sehzentrum im Hirn weitergeleitet werden. Bei niedrigeren Wirbeltieren handelt es sich dabei um das Tectum opticum, bei höheren um den Corpus geniculatum tief im Hirn, der seinerseits mit dem visuellen Cortex im Hinterkopf in Verbindung steht. Ganglienzellen zeigen im Dunkeln eine spontane Aktivität von etwa fünf bis zehn Nervenimpulsen pro Sekunde.

Amakrinzellen stellen laterale Verbindungen zwischen Bipolar- und Ganglienzellen her. Es gibt davon mehr als zwanzig Klassen, die verschiedene Wechselwirkungen vermitteln. Eine Klasse stellt zum Beispiel erregende elektrische Kopplungen zwischen Bipolarzellen her, eine andere bewirkt negative Rückkopplungen auf Bipolarzellen, die dadurch kurz nach ihrer Aktivierung gehemmt werden, also eine zeitliche Hochpasscharakteristik aufweisen. Je nach ihrem zeitlichen Verhalten unterscheidet man zwei Typen von Ganglienzellen: Die einen zeigen eine lang anhaltende Stimulationsantwort, die anderen eine trans-

iente Antwort, entweder wegen Kopplung an Bipolarzellen mit Hochpasscharakteristik oder wegen Hemmung durch Amakrinzellen. Insgesamt gibt es etwa eine Million Ganglienzellen, also etwa hundertmal weniger als Photorezeptorzellen. In der Fovea centralis ist jeder Zapfen nur an eine Bipolarzelle direkt gekoppelt, die wiederum auch nur eine Ganglienzelle innerviert, in der Peripherie kommen mehrere Photorezeptoren – hauptsächlich Stäbchen – auf eine Bipolarzelle und mehrere Bipolarzellen auf eine Ganglienzelle. Dies erhöht die Empfindlichkeit, reduziert aber das räumliche Auflösungsvermögen der Ganglienzelle. Durch die antagonistische Wirkung ihres Zentrum-Umfeld-Aufbaus reagieren Bipolar- und Ganglienzellen nur sehr schwach auf diffuses Licht, das Zentrum und Umfeld gleichermaßen stimuliert und somit bewirkt, dass sich Erregung und Hemmung die Waage halten.

Die Retina ist also trotz ihrer Kompaktheit und lokalen Struktur ein relativ komplexes System mit vielfältigen Verschaltungen, deren Zweck nicht unbedingt immer offensichtlich ist. Dennoch wurden verschiedene Funktionen durch intensive Forschung im Laufe der Zeit aufgeschlüsselt. Eine Auswahl davon wird im nächsten Abschnitt erläutert.

Funktion der Retina

Da wir in der Lage sind, unsere Augen in kurzer Zeit fast beliebig auszurichten, ermöglicht uns die gute Auflösung und Farbdiskrimination in der Fovea centralis, die interessanten Details in der Umgebung genau zu untersuchen, während die schlechter auflösende und weitgehend monochromatische – dafür aber sehr lichtempfindliche – Peripherie dazu dient, diese Details in einem grossen Gesichtsfeld zu finden und unsere Augenbewegungen dorthin zu lenken.

Die logarithmische Charakteristik der Photorezeptorzellen bewirkt, dass die Potentialdifferenz, die ein lokaler Bildkontrast hervorruft, unabhängig von der Helligkeit ist. Diese Strategie trägt der Tatsache Rechnung, dass die meisten Objekte in unserer Umgebung Licht reflektieren oder streuen, das von einer global wirkenden Lichtquelle ausgesandt wird. Da das Streulicht im allgemeinen proportional zur Helligkeit der Lichtquelle ist, hängen die Verhältnisse der Lichtstärken und somit die Kontraste nicht von der Beleuchtung ab. Gerade in natürlichen Umgebungen ist es wichtig, dass die Repräsentation der visuellen Signale weitgehend invariant gegenüber der Helligkeit der Lichtquelle ist, denn

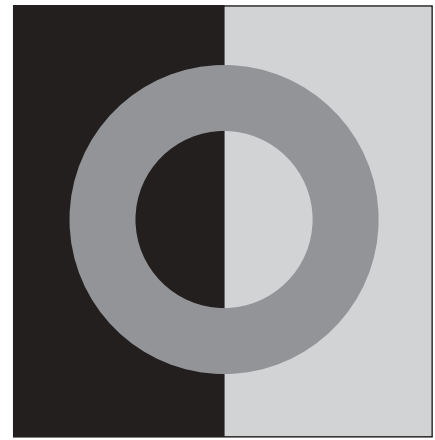


Bild 2 Variables Helligkeitsempfinden

Wenn die Übergangslinie zwischen hellem und dunklem Hintergrund abgedeckt wird, scheint der einfarbige Ring links heller zu sein als rechts.

letzten Endes geht es darum, Objekte unabhängig von der Beleuchtung, die sich ja schon tagsüber um mehrere Grössenordnungen ändern kann, zu erkennen. Durch logarithmische Kompression wird ein riesiger Beleuchtungsstärkenbereich auf Potentialdifferenzen von weniger als hundert Millivolt übertragen. Damit eine Photorezeptorzelle aber trotzdem auf kleine Helligkeitsänderungen empfindlich ist, werden transiente Signale verstärkt und das Potential dann langsam auf seinen Gleichgewichtswert zurückadaptiert. Auf dieser Adaptierungsfähigkeit basiert der enorm grosse Helligkeitsdynamikbereich der Photorezeptorzellen. Eine menschliche Photorezeptorzelle erreicht durch diese Strategie einen Helligkeitsdynamikbereich von fünf Grössenordnungen bei einer Kontrastauflösung von 2%, die weitgehend unabhängig von der Beleuchtung ist. Der Preis, den wir für diese Datenkodierung zahlen, ist unsere vergleichsweise schlecht ausgebildete Fähigkeit, absolute Helligkeiten abzuschätzen. So erscheint uns ein schwarzes Objekt im Sonnenlicht dunkler als ein weisses Objekt bei Zimmerbeleuchtung, auch wenn uns Lichtmessgeräte eines Besseren belehren (Bild 2).

Hinter der räumlichen Informationsverarbeitung stecken ähnliche Prinzipien wie hinter der zeitlichen. Der Bandpassfilter, der durch laterale Erregung und Hemmung implementiert wird, unterdrückt hohe Raumfrequenzen, um – auf Kosten der Auflösung – das Signal-Rausch-Verhältnis zu verbessern, und niedrige Raumfrequenzen, um das Signal zu dekorrelieren und langwellige Beleuchtungsvariationen herauszufiltern. So ergibt sich auch hier eine Kontrastkodierung, was im wesentlichen einer Kantenextraktion entspricht. Auch bei computer-

gestützter Bildverarbeitung ist die Kantenextraktion normalerweise die erste und wichtigste Operation.

Die Auftrennung der Bipolar- und Ganglienzellen in On-Zentrum- und Off-Zentrum-Typen hat vermutlich damit zu tun, dass es für das Signal-Rausch-Verhältnis und den Energieverbrauch günstiger ist, wenn die Zellen im Grundzustand weitgehend inaktiv sind, als wenn sie auf einem mittleren Aktivitätsniveau gehalten werden müssen, um positive und negative Signale kodieren zu können.

Die analogen graduierten Potentiale, mit denen Photorezeptor-, Bipolar- und Horizontalzellen Signale repräsentieren und kommunizieren, haben den Vorteil, dass sie kleine Werte besser kodieren als Aktionspotentiale dies tun und dass sie eine höhere Informationsübertragungsrate erlauben. Die wertkodierte Darstellung der visuellen Daten ermöglicht zudem – im Gegensatz zur zeitkodierten Repräsentation durch Nervenimpulse – die elegante Implementation einer Signalausbreitungsfunktion, die mit der Entfernung durch resistive Verbindungen graduell abnimmt, ohne dass die Zellen individuell weiträumig gekoppelt sein müssen.

Die Retina kann also in erster Näherung als nichtlinearer zeitlicher und räumlicher Bandpassfilter für die Abbildung der Umwelt betrachtet werden. Ihre Funktion ist einerseits, im zentralen Bildbereich eine gute Auflösung und Farbdiskrimination zu erreichen, und andererseits, das normalerweise stark korrelierte Bild zu komprimieren, also die Bildinformation, die mehr in räumlichen und zeitlichen Veränderungen als in konstanten Grauwerten liegt, effizienter zu kodieren. Dadurch kann die limitierte Datenübertragungsbandbreite des Sehnervs besser genutzt und eine Überflutung der höheren Sehzentren im Gehirn durch redundante Eingangssignale verhindert werden.

Implementationen künstlicher Retinae

Wie wir gesehen haben, benutzt die Netzhaut also eine Fülle von Strategien, um Bilddaten möglichst effizient aufzunehmen und zu kodieren. Die meisten davon können auch in künstlichen Systemen angewandt werden. Es sollte dabei aber nicht vergessen werden, dass solche Systeme aus anderen Materialien aufgebaut sind und sich deswegen nicht alle Verarbeitungsstufen gleich gut emulieren lassen. Zudem sind auch nicht alle Elemente im Datenverarbeitungsprozess der Netzhaut hinreichend analysiert worden, abgesehen davon, dass gewisse Unter-

schiede zwischen den Retinae verschiedener Wirbeltierarten bestehen. So werden Netzhautemulationen immer einen gewissen Grad an Abstraktion und Vereinfachung beinhalten. Dabei ist es natürlich von Vorteil, Substrate zu wählen, bei denen physikalische Vorgänge ausgenutzt werden können, die ähnliche Charakteristika aufweisen wie die natürliche Retina.

Wir wollen im folgenden Möglichkeiten und Limitationen von Netzhautemulationen anhand einiger Systeme, die zur Entwicklung der künstlichen Retina wesentliche Beiträge geleistet haben, aufzeigen. Alle diese Systeme haben gemein, dass die Retina lediglich auf elektrischem Niveau emuliert wurde und kein Versuch unternommen wurde, biochemische Vorgänge nachzuvollziehen.

Ein frühes System [1] wurde aus diskreten elektrischen Standardkomponenten wie Photosensoren, Widerständen, Kondensatoren, Differenzverstärkern und Lampen zusammengebaut. Es enthielt 700 Photosensoren, die an zwei resistive Netzwerke angekoppelt waren, welche die erregende und hemmende laterale Wechselwirkung emulierten, die durch die Gap junctions der Photorezeptorzellen beziehungsweise der Horizontalzellen vermittelt wird (Bild 3). Die beiden Netzwerke waren als hexagonales Gitter (Bienenwabengeometrie) ausgelegt, wobei jeder Knoten nur mit seinen sechs nächsten Nachbarn direkt verbunden war. Der Einfluss des Potentials an einem Knoten auf andere Knoten fällt bei einem solchen Netzwerk etwa exponentiell mit der Distanz ab. Die Widerstandsparameter waren dabei so gewählt, dass die Signale im erregenden Netzwerk stärker abfielen als im hemmenden, wie das beim biologischen Vorbild auch der Fall ist. Die Ansprechzeiten der beiden Netzwerke wurden durch Kondensatoren an den entsprechenden Knoten bestimmt. Jedem Photorezeptor wurde je ein Knoten in den beiden Netzwerken zugeordnet. Die Signale dieser beiden Knoten wurden subtrahiert in einer Schaltung, die eine Bipolar- oder Ganglienzelle

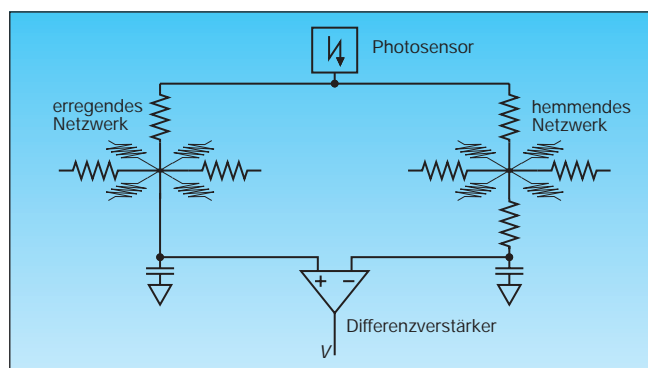
emulierte und deren Ausgangssignal eine Lampe ansteuerte. Auf der zweidimensionalen Lampenanzeige konnte dann eine Zentrum-Umfeld-Antwort beobachtet werden mit einer kürzeren Ansprechzeit für die Erregung als für die Hemmung, wie sie für solche Zellen typisch ist (Bild 4). Die räumliche Zentrum-Umfeld-Charakteristik wurde also durch die Differenz zweier radial etwa exponentiell abfallender Funktionen erzeugt.

Durch die Verwendung von Widerstandsnetzwerken funktionierte das ganze System parallel und in Echtzeit ohne Iterationen und Konvergenzprobleme, die mit Simulationen rückkoppelter Schaltungen einhergehen würden, und vor allem – wie auch unser Hirn – ohne taktgebende Uhr, die ja das Herz eines heutigen Computers bildet.

Knapp zwanzig Jahre später wurde damit begonnen, Retina-Emulationen als integrierte Schaltungen herzustellen. Hierbei hat sich Silizium als sehr geeignetes Substrat erwiesen. Dies liegt zunächst einmal daran, dass die Siliziumtechnologie schon sehr ausgereift ist und man sich mittlerweile preisgünstigen Zugriff dazu verschaffen kann, ohne aufwendige Apparaturen kaufen zu müssen. Entscheidend ist aber auch, dass sich mit Standardsiliziumprozessen gute Photosensoren herstellen lassen, deren spektrale Empfindlichkeit derjenigen der Stäbchenzellen recht ähnlich ist. Diese Photosensoren lassen sich mit elektrischen Schaltungen auf einem Siliziumchip integrieren. Integrierte Schaltungen sind jedoch im wesentlichen zweidimensional, und so müssen die Schaltungselemente, die Zellen aus verschiedenen Schichten der Retina entsprechen, in dieselbe Ebene geklappt werden (Bild 5). Da die Netzhaut jedoch nur aus wenigen Zellschichten besteht und für die lateralen Wechselwirkungen nur nächste Nachbarn miteinander verdrahtet werden müssen, lässt sich dies – natürlich mit reduzierter Auflösung – durchaus bewerkstelligen. Neben ihren offensichtlichen Vorteilen wie ihrer mechanischen

Bild 3 Schematischer Aufbau einer künstlichen Netzhaut [1]

Das Signal jedes Photosensors interagiert lateral über je ein erregendes und ein hemmendes hexagonales Widerstandsnetzwerk mit denjenigen benachbarter Photosensoren. Das Ausgangssignal V ergibt sich aus der Signaldifferenz zwischen den lokalen Knoten beider Netzwerke.



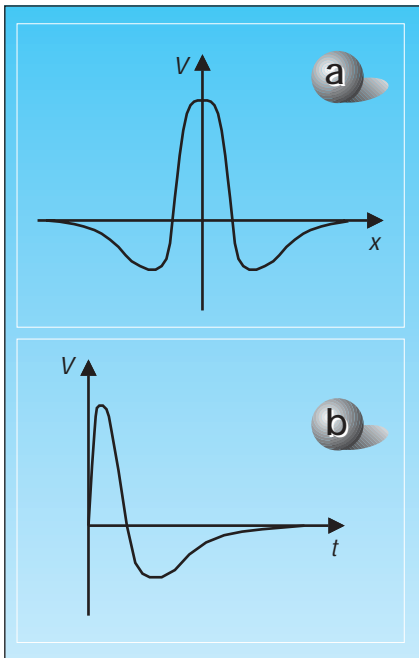


Bild 4 Impulsantworten der künstlichen Netzhaut von Bild 2

- a Interpolierte räumliche Verteilung des Ausgangssignals entlang einer Achse bei Beleuchtung eines einzigen Photosensors
- b Zeitlicher Verlauf des Ausgangssignals an einer Stelle bei grossflächiger Stimulation mit einem kurzen Lichtblitz. Der Nulldurchgang des Signals kommt durch die grössere Zeitkonstante des hemmenden Widerstandsnetzwerkes zustande.

Robustheit, geringen Grösse und geringem Leistungsverbrauch haben integrierte Emulationen allerdings einen wesentlichen Nachteil. Es handelt sich dabei um die Tatsache, dass sich bei kleinen Strukturen im allgemeinen und bei integrierten Analogschaltungen im besonderen die Inhomogenität des Fabrikationsprozesses bemerkbar macht und somit das Problem von systematischen räumlichen Unregelmässigkeiten und eventuell sogar Fehlstellen inhärent ist. Ein Problem, das auch von biologischen Verarbeitungsstrukturen gelöst werden muss. Es ist deshalb eine Herausforderung, beim Entwurf retinaler Schaltungen und Datenverarbeitungsalgorithmen Methoden zu finden, die gegenüber Variabilität der Schaltungselemente tolerant sind und die auch den Ausfall einzelner Elemente verkraften können. Letzteres ist schon durch die lokale Kopplung und parallele Datenverarbeitung einer künstlichen Netzhaut teilweise gewährleistet, denn ein Schaltungsdefekt beeinflusst dort nur eine Umgebung von der Grösse des rezeptiven Feldes um den entsprechenden Bildpunkt. Bei rechnergestützter serieller Bildverarbeitung kann ein einzelner defekter Transistor unter Umständen den ganzen Verarbeitungsprozess lahmlegen.

Die erste integrierte Retina auf Siliziumbasis [2] besass Photorezeptorzellen mit logarithmischer Charakteristik über vier bis fünf Grössenordnungen in der Beleuchtungsstärke. Ein hexagonales Netzwerk von nichtlinearen regelbaren Widerständen mit relativ hohen Streukapazitäten übertrug die hemmende Wirkung der Horizontalzellen. In Bipolarzellen wurde dann das Potential eines Knotens von demjenigen des dazugehörigen Photorezeptors subtrahiert, so dass sich eine Zentrum-Umfeld-Charakteristik ergab und zeitliches Hochpassverhalten für ausgedehnte Lichtreize. Durch das Fehlen eines erregenden Netzwerkes war kein räumlicher Tiefpassfilter zur Rauschunterdrückung vorhanden und das Zentrum der räumlichen Charakteristik nur einen Bildpunkt gross. Infolge der logarithmischen Kompression des Helligkeitssignals wurde durch die Subtraktion in den Bipolarzellen der Logarithmus des Kontrastes kodiert, und das Ausgangssignal war somit in erster Näherung unabhängig von der Bildhelligkeit. Alle Schaltungselemente wurden mit sehr kleinen Strömen betrieben, so dass die ganze Schaltung mit einer Auflösung von 48×48 Bildpunkten eine Leistung von nur einem Zehntel Milliwatt verbrauchte.

Da bei dieser Retina kleine Signale im oben beschriebenen räumlichen Rauschen untergingen, benutzte eine spätere Fassung [3] einen Rückkopplungsmechanismus, durch den das System bei homogener Beleuchtung kalibriert werden konnte. Zur Speicherung der Kalibration

wurde eine Methode angewandt, die sonst für reprogrammierbare digitale Speicherchips benutzt wird, nämlich die Speicherung elektrischer Ladung auf einer unter normalen Bedingungen isolierten Leiterplatte, die kapazitiv das Verhalten einer Ausleseschaltung beeinflusst. Der Speicherinhalt kann durch gezieltes Anlegen von Spannungen oder durch Beleuchtung mit ultravioletter Strahlung verändert werden. In vorliegendem Fall wurde die zweite Methode verwendet. Das räumliche Rauschen konnte damit bei der kalibrierten Beleuchtungsstärke drastisch reduziert werden. Da aber auch der Verstärkungsfaktor der einzelnen Elemente statistischen Fluktuationen unterworfen war, musste die Kalibration bei wesentlichen Änderungen der Beleuchtungsstärke wiederholt werden.

In einem weiteren Iterationsschritt [4] wurde deshalb die Siliziumretina mit einem adaptiven Photosensor ausgestattet, der eine sehr ähnliche Charakteristik aufwies wie biologische Photosensoren und durch einen gut kontrollierbaren transienten Verstärkungsmechanismus das Signal-Rausch-Verhältnis entscheidend verbesserte. So konnte auch ohne Kalibration ein akzeptables Ausgangssignal generiert werden. Bild 6 zeigt das transiente Ausgangssignal einer Weiterentwicklung dieser Retina.

Die oben beschriebene integrierte Emulation und ihre Modifikationen verwendeten Spannungen zur Repräsentation der Signale. Leider ist die Imple-

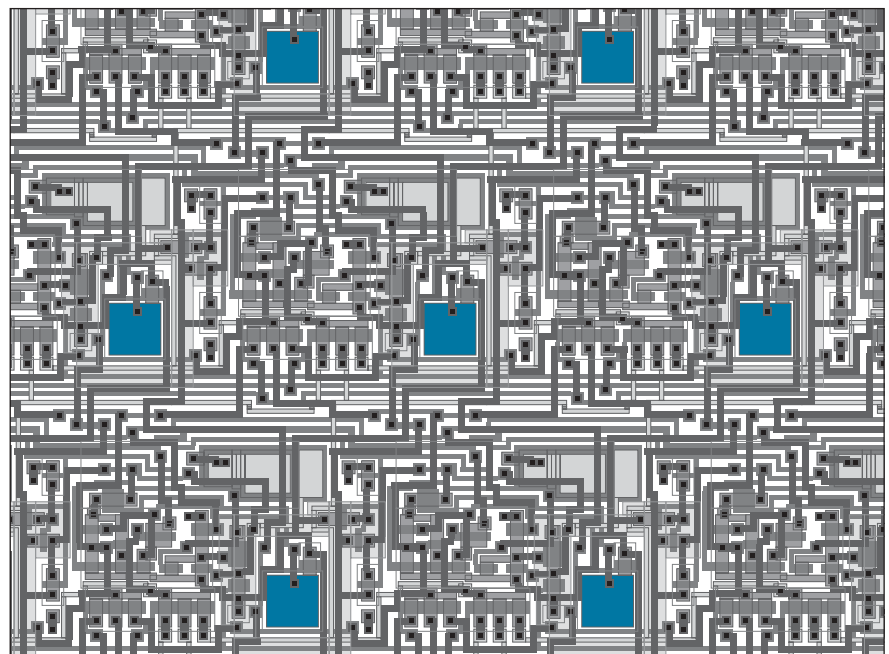


Bild 5 Teil eines Maskenentwurfs für eine integrierte künstliche Netzhaut auf Siliziumbasis

Die Schaltung hat eine hexagonale (bienenwabeförmige) Struktur. Die lichtempfindlichen Bereiche eines Photosensors und seiner sechs nächsten Nachbarn sind als quadratische Flächen erkennbar.

mentierung von einigermaßen linearen Widerstandsnetzwerken für grössere Spannungssignale in Standardsiliziumprozessen recht aufwendig. Das Problem liegt darin, dass in integrierten Schaltungen sehr kleine Ströme fliessen und deshalb hochohmige Widerstände erforderlich sind, um vernünftige Spannungsgradienten zu erzeugen. Ein Standardsiliziumprozess verfügt aber nur über niederohmige resistive Schichten, weshalb man sich zur Implementation von solchen Widerstandselementen mit relativ aufwendigen Transistorschaltungen behelfen muss. Ein System mit zwei hexagonalen Widerstandsnetzwerken erfordert sechs Widerstandselemente pro Photosensor, wodurch die erforderliche Schaltung pro Bildpunkt natürlich relativ gross wird. Wie sich aber herausstellte, kann ein regelbares, mehr oder weniger lineares resistives Netzwerk mit einem einzigen Transistor pro Widerstandselement hergestellt werden, wenn Ströme anstelle von Spannungen zur Datenrepräsentation verwendet werden.

Basierend auf dieser Erkenntnis wurde in der Folge eine kompakte Emulation der Wechselwirkung von Photorezeptor- und Bipolarzellen mit je einem erregenden und einem hemmenden hexagonalen Widerstandsnetzwerk implementiert [5], dessen Ausgangsströme eine Zentrum-Umgebung-Charakteristik aufwiesen, die den Kontrast relativ unabhängig von der Beleuchtungsstärke kodierte. Das zeitliche Verhalten der biologischen Retina wurde dabei jedoch nicht modelliert. Durch das Fehlen der transienten Verstärkung trat das räumliche Rauschen wieder vermehrt in Erscheinung. Es konnte jedoch durch Glättung des Signals über mehrere Photorezeptoren mittels erregender Interaktion reduziert werden. Weil nur etwa zehn Transistoren pro Photorezeptor für die Bildverarbeitung benötigt wurden, konnten relativ grosse Netzwerke monolithisch integriert werden. Auf einem Chip mit einer Fläche von einem Quadratzentimeter wurde in einem konservativen Siliziumprozess eine Auflösung von bis zu 230×210 Bildpunkten erreicht. Mit einem modernen Prozess könnte die Auflösung um etwa einen Faktor drei pro Dimension gesteigert werden. Für Videoauflösung würde also weniger als ein Quadratzentimeter benötigt.

Wie bei der biologischen Netzhaut ist die verfügbare Datenübertragungsbandbreite bei integrierten Schaltungen beschränkt. Zwar sind die Schaltungen auf Siliziumchips etwa eine Million Mal schneller als die biologischen, dafür ist die Verdrahtung des Chips mit der Aussenwelt viel gröber als im Sehnerv, so

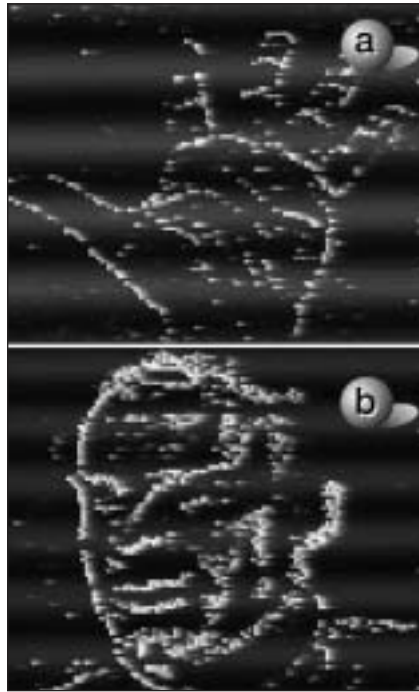


Bild 6 Signalverteilungen am Ausgang einer integrierten künstlichen Netzhaut mit 125×94 Bildpunkten in hexagonaler Anordnung

Durch die hemmende Wechselwirkung zwischen benachbarten Bildpunkten und die Verstärkung zeitlicher Änderungen in den einzelnen Photosensoren werden sich bewegende Konturen im Bild hervorgehoben.

dass nur einige wenige Kanäle zur Verfügung stehen. Die effiziente Bildkompression, die die Netzhaut durchführt, ist deswegen bei Emulationen mit integrierten Schaltungen ebenso wichtig wie in der biologischen Netzhaut.

Spätere Versionen der obigen Schaltung [6] wurden deshalb mit einem zusätzlichen Element ausgestattet, das den Ausgangsstrom wie eine Ganglienzelle in Aktionspotentiale mit Impulsfrequenzadaptation umformte. Die Impulse aller Zellen wurden dann auf einen asynchronen digitalen Datenübertragungsbus transferiert, wobei in jedem Impuls die Position der sendenden Zelle binär kodiert wurde, so dass ein Empfängersystem die Topographie der Zellstruktur rekonstruieren konnte. Andere Emulationen auf Siliziumbasis [7] konzentrierten sich auf die topographische Organisation der Netzhaut mit einer hochauflösenden Fovea centralis in der Mitte und radial nach aussen abnehmender Auflösung. Eine solche Topographie schränkt aber die Möglichkeiten zur Datenverarbeitung in einer zweidimensionalen Ebene stark ein, weil in der Mitte zwischen den Photosensoren kaum für weitere Strukturen Platz ist. Eine solche Anordnung kann ausserdem ihr Potential nur in einem aktiven System voll entfalten, das

durch Ausrichtung der optischen Achse des Abbildungssystems interessante Objekte auf die Fovea centralis projizieren kann. Wenn dies der Fall ist, können gewisse Bildverarbeitungsalgorithmen effizient implementiert werden wie zum Beispiel massstabsinvariante Objektrepräsentationen.

Die Aufnahme von Farbinformation durch drei verschiedene Typen von Zapfenzellen wird von herkömmlichen Videokameras, die mit einem Mosaik aus drei verschiedenen Farbfiltern versehen sind, bereits gut emuliert. Standardsiliziumprozesse können aufgrund der wellenlängenabhängigen Eindringtiefe der optischen Strahlung auch zur Aufnahme von Farbsignalen verwendet werden [8], allerdings mit relativ schlechter Diskriminationsfähigkeit.

Die in diesem Abschnitt beschriebenen Schaltungen bilden nur eine kleine Auswahl der künstlichen Systeme, die Aspekte der Bildverarbeitung durch die Netzhaut emulieren. Keines dieser Systeme implementiert alle bekannten wichtigen Funktionen der Retina, ganz zu schweigen von den Funktionen, die noch nicht ausreichend analysiert wurden. Welche Funktionen für eine Emulation wichtig sind, hängt von den anvisierten Anwendungen ab, die von der Modellierung biologischer Prozesse bis zu industriellen Systemen reichen.

Anwendungen künstlicher Retinae

Durch die vielfältigen Vorteile der Netzhaut gegenüber traditionellen künstlichen Bildaufnahme- und -verarbeitungssystemen ergibt sich eine grosse Palette von praktischen Anwendungen, bei denen es attraktiv ist, solche Systeme durch Netzhautemulationen zu ersetzen. Dabei muss aber den strukturellen und funktionalen Eigenheiten der künstlichen Retinae auch bei der Konzeption des Gesamtsystems Rechnung getragen werden. So eignen sie sich wegen ihrer limitierten Bildauflösung am besten zum Einsatz in aktiven Systemen, von denen sie auf interessante Bildregionen gerichtet werden können.

Ein interessantes Anwendungsgebiet wird einmal die Medizin werden, die versucht, degenerierte natürliche Retinae durch künstliche zu ersetzen. Diese Bestrebungen werden im vorliegenden Heft im Artikel von Alfred Stett und Hugo Hämmerle ab Seite 11 näher beleuchtet.

Viele Möglichkeiten eröffnen sich auch in der Robotik. Hier könnten künstliche Retinae in autonomen Systemen

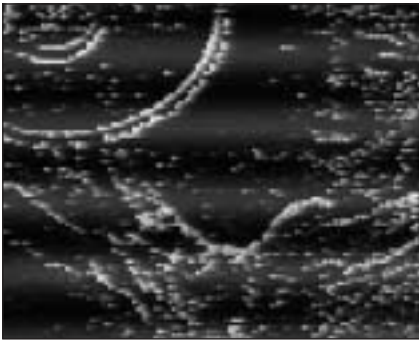


Bild 7 Signalverteilung am Ausgang einer integrierten künstlichen Netzhaut (vgl. Bild 5) bei grosser Helligkeitsdynamik in der optischen Abbildung

Obwohl die Glühbirne (links oben) tausendmal heller ist als das Gesicht und die Hand, werden die Umrisse von allen Objekten im gleichen Bild extrahiert. Eine herkömmliche Kamera hat schon mit einer Helligkeitsdynamik von einer Dekade Mühe, wie sie in Bildern mit starken Schatten oder mit schwarzen und weissen Objekten zu finden ist.

eingesetzt werden, die sich durch Lernprozesse eine Repräsentation ihrer Umwelt aufbauen. Wie beim Menschen spielen dabei visuelle Eingangsdaten im allgemeinen eine zentrale Rolle. Künstliche Retinae sind für deren Erfassung geeignet, weil sie wichtige Anforderungen solcher Systeme erfüllen wie Echtzeitbildverarbeitung, Kompaktheit, Robustheit und niedriger Leistungsverbrauch. Ein illustratives Beispiel für die Wichtigkeit dieser Anforderungen sind extraterrestrische Roboter. Man denke dabei an das Mars-Fahrzeug «Pathfinder», das aufgrund seiner ungenügenden Autonomie von der Erde aus gesteuert werden musste. Die grosse Entfernung zwischen Erde und Mars verursachte lange Kommuni-

kationszyklen und erlaubte daher nur langsame Bewegungen des Fahrzeuges. Hinzu kam ein grosser Energieverbrauch, der durch die relativ schwache Sonneneinstrahlung auf dem Mars nicht ausreichend durch Sonnenkollektoren abgedeckt werden konnte.

In der Verkehrstechnik interessiert sich vor allem die Automobilindustrie für die potentiellen Fähigkeiten der künstlichen Retinae. Hier sind unter anderem kostengünstige Produktionsmöglichkeiten gefragt. Da integrierte Retinachips im Gegensatz zu Videokamerachips mit billigen Standardsiliziumprozessen hergestellt werden können, wäre diese Anforderung bei einer Massenproduktion zweifellos erfüllt. Der Hauptvorteil der Retinachips in einer solchen Anwendung ist aber die grosse Helligkeitsdynamik. Wenn eine Kamera zur Bildaufnahme für ein elektronisches System eingesetzt werden soll, das im Auto dem Fahrer unter anderem helfen soll, die Spur und den Abstand zum Vordermann zu halten und Verkehrszeichen zu erkennen, muss sie in der Lage sein, unter allen natürlichen Bedingungen gleichzeitig alle wichtigen Details zu erkennen. Unregelmässige Beleuchtungsverhältnisse, zum Beispiel bei Tunneldurchfahrten, Lichtreflexionen auf nasser Fahrbahn oder bei Nachtfahrten dürfen die Funktion eines solchen Systems nicht beeinträchtigen (Bild 7). Wie wir gesehen haben, könnte eine künstliche Retina mit solchen Problemen allenfalls durch Blendenregulierung mit globalen Helligkeitsvariationen zurechtkommen.

Ein grundlegendes Projekt unseres Institutes befasst sich mit der Emulation verschiedener Teile des Cortex mit integrierten Schaltungen, welche über einen asynchronen digitalen Datenbus durch Impulse kommunizieren [9] und sensorische Eingangssignale (z.B. von einer Retina) verarbeiten können. Im Rahmen eines solchen modularen Systems können biologische Datenverarbeitungs- und Lernprozesse studiert werden, die vielleicht in nicht allzu ferner Zukunft auch ihren Weg in praktische Anwendungen finden werden. Wenn nur ein kleiner Teil der Leistungen des visuellen Cortex in einem solchen System repliziert werden könnte, wäre dies schon ein grosser Fortschritt gegenüber existierenden künstlichen Systemen. Allein schon ein Gesichtserkennungssystem, das ähnlich gut funktioniert wie unser biologisches, hätte ein riesiges Anwendungspotential.

Zusammenfassung

Wir haben gesehen, dass wir bei der Aufnahme und Verarbeitung visueller Information aus unserer Umgebung mit dem eigenen Sehsystem allen künstlichen Systemen, die wir bisher gebaut haben, in den meisten Hinsichten weit überlegen sind. Bei der Konzeption künstlicher Systeme lohnt es sich deshalb, biologische Lösungsansätze in Betracht zu ziehen. Wie wir anhand von künstlichen Retinae aufgezeigt haben, können biologisch inspirierte Systeme effizient implementiert werden und sind in vielen Bereichen einsetzbar.

Literatur

- [1] K. Fukushima, Y. Yamaguchi, M. Yasuda, S. Nagata: An Electronic Model of the Retina. Proc. IEEE 58(1970)12, 1950-1951.
- [2] C.A. Mead, M.A. Mahowald: A Silicon Model of Early Visual Processing. Neural Networks 1(1988), 91-97.
- [3] C.A. Mead: Adaptive Retina. In: Analog VLSI Implementation of Neural Systems. Boston: Kluwer Academic 1989, pp. 239-246.
- [4] M.A. Mahowald: Silicon Retina with Adaptive Photoreceptors. Proc. SPIE 1473(1991), 52-58.
- [5] K.A. Boahen, A.G. Andreou: A Contrast Sensitive Silicon Retina with Reciprocal Synapses. In: Adv. Neural Information Processing Systems 4. San Mateo, CA: Morgan Kaufmann 1992, pp. 764-772.
- [6] K.A. Boahen: A Retinomorph Vision System. IEEE Micro 16(1996)5, 30-39.
- [7] J. Van der Spiegel, G. Kreider, C. Claeys, I. Debusschere, G. Sandini, P. Dario, F. Fantini, B. Belluti, G. Soncini: A Foveated Retina-like Sensor Using CCD Technology. Norwell, MA: Kluwer Academic 1989.
- [8] P. Seitz, D. Leipold, J. Kramer, J.M. Raynor: Smart Optical and Image Sensors Fabricated with Industrial CMOS/CCD Semiconductor Processes. Proc. SPIE 1900(1993), 30-39.
- [9] S.R. Deiss, R.J. Douglas, A.M. Whatley: A Pulse-coded Communications Infrastructure for Neuromorphic Systems. In: Pulsed Neural Networks. Cambridge, MA: MIT 1999, pp. 157-178.

Rétine artificielle

Copie nature sur puce analogique

En termes de faculté d'interprétation des images de son environnement, l'homme devance largement tout système artificiel. Une raison majeure de ce fait réside dans les méthodes efficaces que notre rétine utilise pour capter et filtrer l'image, alors que les divers systèmes concoctés par l'intelligence des hommes en divergent encore largement. Une nouvelle perspective s'ouvre par le développement d'une rétine artificielle qui s'oriente à la structure et au fonctionnement de la rétine naturelle. Produites par le biais de processus siliceux standards sur monolithes, de telles rétines artificielles sont peu onéreuses, compactes, robustes, économiques en énergie. La grande dynamique de luminosité et la qualité du prétraitement et du codage des données qui caractérisent le modèle fourni par la nature ont pu être reproduites avec de tels circuits. Or, déjà par ces possibilités, la rétine artificielle serait promise à une foule d'applications en recherche et dans l'industrie. Mais au-delà, elle est capable de générer les signaux d'entrée de circuits dont le but est la modélisation d'autres parties du traitement des données visuelles, en temps réel, de l'organe biologique. On espère ainsi pénétrer davantage la façon dont le système visuel de l'homme traite les données acquises. D'autre part, des applications pratiques variées s'offrent à l'exploitation de ces nouvelles connaissances scientifiques.