

# Blickkontrolle in neuraler Architektur für Aktive Sehsysteme

*A. Schierwagen*

Universität Leipzig, Institut für Informatik

Augustusplatz 10-11, 04109 Leipzig

E-mail: [schierwagen@aix550.informatik.uni-leipzig.de](mailto:schierwagen@aix550.informatik.uni-leipzig.de)

## **Zusammenfassung**

Im Unterschied zu traditionellen Methodologien in der Künstlichen Intelligenz wird das Sehen im Rahmen des neuen Paradigmas "Aktives Sehen" im Verhaltenskontext des Systems untersucht. Aktive Sehsysteme (ASS) weisen Merkmale auf, die der Biologie entstammen (Binokularität, Foveae, Blickkontrolle mit hoher Geschwindigkeit). Dieser Bericht analysiert Berechnungsstrategien der visuomotorischen Kontrolle, die von biologischen Sehsystemen verwendet werden, um sie für künstliche ASS nutzbar zu machen. Neurale Karten, Populationskodierung und Informationsverarbeitung in geschichteten Strukturen werden als universelle Prinzipien analoger neuraler Berechnungen gekennzeichnet, die auch der Blickkontrolle zugrunde liegen. Auf der Basis dieser Prinzipien schlagen wir ein Modell vor, das *kartierte neurale Felder* verwendet, um die raum-zeitliche Dynamik in der Motorkarte des Colliculus superior der Säugetiere darzustellen, eines sensomotorischen Transformationszentrums im Mittelhirn, das entscheidend an der Blickkontrolle beteiligt ist.

## **1. Aktive Sehsysteme**

Lebewesen sind künstlichen Systemen hinsichtlich Orientierungs- und Erkundungsverhalten deutlich überlegen. Der Hauptgrund dafür mag die Tatsache sein, daß die traditionellen Methodologien in der Künstlichen Intelligenz Aspekte des Verhaltens weitgehend ignoriert haben. Im Bereich des Sehens beispielsweise ist die bislang wichtigste Berechnungstheorie von Marr [1] essentiell über passives Sehen. Das bedeutet, daß die Effekte des Verhaltens, das der Beobachter ausführt, ausgeschlossen wurden. Ebenfalls ausgeschlossen wurden spezielle Merkmale des menschlichen Sehens, wie sein komplexes Blickkontrollsystem, das schnelle Augenbewegungen (Sakkaden) einschließt, und die Fovea mit ihrer stark vergrößerten Auflösung.

Ein neues Paradigma, das als Animaten- oder Aktives Sehen [2,3] bezeichnet wird, argumentiert dafür, daß Sehen am besten im Verhaltenskontext des Systems verstanden werden kann. Aktive Sehsysteme (ASS) haben biologische Merkmale wie Binokularität und Foveae. Die wichtigste Eigenschaft, die quasi das Gebiet des Aktiven Sehens definiert, ist die mit hoher Geschwindigkeit arbeitende Blickkontrolle. In Abhängigkeit von der Aufgabe kann ein ASS den visuellen Sensor während der Datenaufnahme ausrichten. Verglichen mit passiven Systemen zeigen ASS, daß visuelle Berechnungen viel weniger umfangreich sind, wenn sie im Verhaltenskontext betrachtet werden.

Um die Spezifik des Aktiven Sehens darzustellen, soll die Organisation des sakkadischen Augenbewegungssystems hier kurz skizziert werden. Die Sinnesfläche im menschlichen Auge (die Retina) unterscheidet sich von der üblicher elektronischer Kameras dadurch, daß sie eine Region höher Auflösung nahe der optischen Achse besitzt (die Fovea). In diesem kleinen Bereich (weniger als 0,01 % der Retina) ist die Auflösung eine Größenordnung besser als in der Peripherie. Dieser retinale Aufbau bewirkt, daß ein großer Teil des visuellen Felds auf die Retina abgebildet wird, wobei hohe Sehschärfe nur in der Fovea vorliegt. Das menschliche visuelle System benutzt also Sakkaden, um die Fovea auf räumliche Ziele in unterschiedlicher Position zu lenken. Daher ist adaptive Blickkontrolle von grundsätzlicher Bedeutung für das menschliche Sehen.

Gegenwärtig untersuchen wir im Rahmen eines multidisziplinären Forschungsprojekts biologische Sehsysteme. Von besonderem Interesse dabei sind Berechnungsstrategien der

visuomotorischen Kontrolle. Biologische Untersuchungsergebnisse bilden den Ausgangspunkt für die Entwicklung von Computermodellen, um ein ASS zu erzeugen.

## **2. Blickkontrolle: Neurobiologische Resultate und Modelle**

Eine zentrale Rolle bei der Blickkontrolle spielt der Colliculus superior (CS), ein sensomotorisches Transformationszentrum im Mittelhirn der Säuger. Der CS ist eine aus 7 Faser- bzw. Zellschichten aufgebaute Struktur. Aufgrund von anatomischen, physiologischen und Verhaltensstudien wird der CS in zwei Hauptteile eingeteilt - die oberflächigen und die tiefen Schichten.

Neurone in den oberen CS-Schichten antworten nur auf visuelle Reize. Sie sind über die CS-Fläche in einem topografischen, kartenartigen Format verteilt, d.h., die Lage eines Neurons in dieser Karte wird durch die Position des Zentrums seines rezeptiven Felds bestimmt. Diese retinotopie Repräsentation erstreckt sich über die CS-Fläche von  $0^\circ$  bis zu  $80^\circ$  des kontralateralen visuellen Halbfelds.

Zellen in den tiefen Schichten entladen mit hochfrequenten Bursts von Spikes unmittelbar vor Beginn von schnellen Augenbewegungen (Sakkaden) [4]. Der Bereich möglicher Sakkaden, vor denen ein gegebenes Neuron jeweils feuert, bestimmt sein Bewegungsfeld (BF). Das BF eines Sakkadenneurons hat ein Zentrum, das zu maximaler Entladung gehört, und einen Gradienten der Antwort, der abfällt für Sakkadenvektoren, die vom Optimum abweichen. Damit feuert ein bestimmtes Neuron für einen spezifischen, aber großen Bewegungsbereich. Neurone in den tiefen Schichten sind gleichfalls in topografischer Anordnung organisiert: die BF-Zentren sakkadischer Neurone bilden eine motorische Karte (Übersicht in [5]). Das bedeutet, daß die Lage aktiver Neurone den motorischen Fehler repräsentiert, der durch den Blickänderungsvektor (Amplitude und Richtung) definiert wird, der noch von den Augen ausgeführt werden muß, um das Ziel zu erreichen. Die Topografie der Motorkarte ist derart, daß der horizontale Motorfehler i.w. längs der rostro-caudalen Achse kodiert ist, während der vertikale Motorfehler längs der medio-lateralen Achse kodiert ist. Je mehr in caudaler (medio-lateraler) Richtung das Zentrum neuronaler Aktivität lokalisiert ist, desto größer ist der horizontale (vertikale) Motorfehler.

Eine Population sakkadischer Neurone beginnt zu feuern, wenn ein Motorfehler auftritt; dabei hat die Aktivitätsverteilung innerhalb der Motorkarte die Form eines Gauß-Hügels.

Die neuronalen Feuerraten kodieren die Bewegungsgeschwindigkeit (z.B. [6]). Sakkaden zu verschiedenen Teilen des visuellen Felds gehen Aktivitätsgipfel in verschiedenen Regionen der Karte voraus. In neueren Tracerstudien [4, 7] konnte gezeigt werden, daß die oberen, visuellen Schichten derart mit den tiefen Schichten verbunden sind, daß beide Karten in räumlicher Übereinstimmung sind. Die Aktivität in der Motorkarte bringt dann das Auge von seinem aktuellen Fixationspunkt zu dem Ziel, das das rezeptive Feld des entsprechenden visuellen Neurons aktiviert hat.

Elektrophysiologische Ableitungen haben gleichfalls gezeigt, daß Neurone in der 0°-Amplitudenregion am rostralen Pol der Motorkarte aktiv sind während Perioden aktiver Fixation, und inaktiv vor Sakkaden (sog. Fixationsneurone). Die lokale Interaktion von Sakkaden- und Fixationsneuronen in der colliculären Motorkarte scheint an der neuronalen Dynamik beteiligt zu sein: im CS der Katze wandert ein Aktivitätsgipfel während einer Sakkade über die Motorkarte von seiner Ausgangsposition zur Fixationszone [8, 9]. Die jeweilige momentane Gipfelposition auf der Karte bestimmt den verbleibenden Motorfehler. Die Sakkade endet, wenn der wandernde Aktivitätsgipfel die Fixationszone erreicht, wo er erlischt. Dieses Ergebnis verdeutlicht, wie die räumliche Repräsentation der Aktivität auf der Motorkarte transformiert wird in einen Zeitcode (Frequenz und Dauer der Entladung), wie er von Motoneuronen benötigt wird. Die raum-zeitliche Transformation wird durch colliculäre kontinuierliche dynamische Kontrolle der Blickänderung realisiert. Bei der Katze (andere Tiere müssen noch untersucht werden) kodiert die Trajektorie der Aktivität über die Motorkarte räumlich das momentane Motorfehler-Signal. Die reziproke Beziehung der Aktivität zwischen Sakkaden- und Fixationsbereichen der CS-Motorkarte ist im Sinne einer wechselseitigen Inhibition der beiden Gebiete interpretierbar.

Die Natur der Blickkontrolle durch die colliculäre Motorkarte ist umstritten. Einige Autoren nehmen an, daß bei Primaten der colliculäre Motorfehler erzeugt wird, indem die Differenz zwischen der corticalen Zielrepräsentation und der schon von den Augen überstrichenen Strecke gebildet wird. Frühere CS-Modelle ordneten dem CS einen Platz oberhalb der Rückkopplungsschleife zu. Das weiter oben betrachtete dynamische Fehlerkodierungsschema wird allerdings als starkes Argument für eine Position des CS innerhalb der postulierten Schleife angesehen.

Nahezu alle Blickkontroll-Modelle basieren auf der Rückkopplung, um die Genauigkeit der Sakkaden zu gewährleisten. Das bedeutet, daß die Motorkarte z.B. Rückkopplung vom Hirnstamm erhält, um die CS-Aktivität anzupassen. Ein jedes CS-Modell auf Feedback-Grundlage muß erklären, wie spatio-temporale (abwärtige) und temporal-spatiale (auf- oder rückwärtige) Transformationen ausgeführt werden können, um damit Sakkaden und den dynamischen Motorfehler berechnen zu können. Im Unterschied zu der gut verstandenen Kodierung der Sakkaden in der komputatorischen Motorkarte des CS ist die Idee von Rückprojektionen zu dieser Karte, die die nötigen zeitlich-räumlichen Transformationen realisieren, eher spekulativ. Zwar konnten Projektionen von Hirnstammstrukturen zum CS anatomisch nachgewiesen werden, jedoch ließen sich keine Hinweise auf die Mechanismen finden, mit denen das Blickkontroll-System die Aktivität der Sakkaden-Neurone anpassen könnte. Wir gehen daher von der Möglichkeit aus, daß die colliculäre dynamische Fehlerkodierung realisiert werden könnte ohne globale Rückkopplung, sondern allein aufgrund der nichtlinearen, raum-zeitlichen Dynamik neuraler Populationen in der Form kartierter neuraler Felder. Die Berechnungskonzepte, auf denen das Modell beruht - neurale Karten und kontinuierliche neurale Felder - sollen als nächstes erläutert werden.

### ***3. Neurale Karten als repräsentationale Strukturen***

Ein Vergleich künstlicher neuraler Netze mit ihren biologischen Gegenspielern offenbart die Diskrepanz zwischen beiden. Will man realistische Neuronennetze modellieren, muß man eine Reihe komplexer Strukturprinzipien einbeziehen, die in biologischen Neuronennetzen auftreten. Eines der bekanntesten Architekturmerkmale ist die neurale Kartierung. Im weiteren ignorieren wir vollständig das Problem der Kartenentstehung durch Selbstorganisation und betrachten Karten so "wie sie sind" und ihre Beziehung zur nichtlinearen, raum-zeitlichen Dynamik in neuronalen Schichten.

Eine neurale Karte besteht aus einer Schicht von Neuronen, deren komputatorische Eigenschaften mit ihrer räumlichen Position variieren. Grundsätzlich sind Karten durch die rezeptiven Feldeigenschaften ihrer Neurone bestimmt, die ihrerseits durch Parameter wie Position (Schwerpunkt), Linienorientierungspräferenz, Bewegungsrichtungsselektivität usw. beschrieben werden (vgl. [10] für Diskussion).

Die Position eines Neurons in der Karte bestimmt weitgehend (1) welchen Teil des Inputs es empfängt, (2) wie dieser Input verarbeitet und (3) zu welchem Ziel das Ergebnis übertragen wird. In Übereinstimmung mit diesen Aspekten der Kartierung lassen sich Karten als (1) afferente, rezeptotop oder sensorische Karten, (2) interne oder Verarbeitungskarten und (3) efferente oder Motorkarten bezeichnen.

Die Neurone einer rezeptotopen Karte stellen ein Feld breit abgestimmter Filter dar, die das Inputsignal parallel verarbeiten. Auf diese Weise wird ihr Input augenblicklich in eine räumlich kodierte Verteilung neuraler Aktivität überführt, wobei die Werte der abgebildeten Parameter als Lokalisationen und Höhen von Aktivitätsgipfeln dargestellt werden. Auf diese räumlich kodierte Information kann durch Prozessoren höherer Ordnung bequem zugegriffen werden (letztere können gleichfalls Karten sein), unter Verwendung relativ einfacher Verknüpfungsschemata. Beispiele rezeptotoper Karten sind retinotop, tonotop und somatotop Karten.

Wird "Berechnung" als eine (nichttriviale) Transformation der repräsentierten Information definiert, ergibt sich, daß rezeptotop Karten nicht komputatorisch sind - sie bilden nur Sinnesflächen ab unter Beibehaltung räumlicher Beziehungen im peripheren Sensorium. Die bisher untersuchten komputatorischen Karten sind zumeist an der Verarbeitung sensorischer Information beteiligt, etwa die schon erwähnten Karten im visuellen System, aber auch im auditorischen System, und an der Programmierung von Bewegungen (vgl. [11]).

Eine der bekanntesten komputatorischen Karten, die an der Bewegungsprogrammierung beteiligt ist, ist die Motorkarte oder Blickrichtungsänderungskarte im CS. Motorkarten repräsentieren systematische Veränderungen in den Bewegungsparametern (Amplitude und Richtung) topografisch auf der neuralen Schicht. Der Berechnungscharakter dieser Karten ist offensichtlich: das topografisch repräsentierte Bewegungskommando muß in zeitliche Erregungsmuster der Motoneurone transformiert werden; dabei definiert der Aktivitätsschwerpunkt die Merkmale der Transformation. Dieses Verarbeitungsprinzip komputatorischer Motorkarten kann als invers zu dem komputatorischer sensorischer Karten gekennzeichnet werden: während sensorische Karten das Ergebnis ortsvarianter Berechnungen repräsentieren, stellen Motorkarten den Quellcode für systematisch variierende Berechnungen dar.

#### **4. Kartierte neurale Felder**

Die Bedeutung, die der synchronen Aktivität großer Neuronenpopulationen (etwa in Form von Oszillationen und wandernden Aktivitätswellen) zukommt, ist bekannt. Die große Anzahl der Neurone, die an diesen Phänomenen beteiligt sind, legt die Behandlung kontinuierlicher neuraler Feldmodelle anstelle diskreter neuraler Netzwerkmodelle nahe, mit Neuronen als Punkte in einem räumlichen Kontinuum. Dieser Zugang ist besonders gut geeignet für neurale Systeme wie CS und Kortex, wo Neurone eines Typs in relativ homogenen Schichten angeordnet sind, die möglicherweise ein gleichförmiges Konnektivitätsmuster vom Typ der lateralen Inhibition aufweisen. Die Frage ist, ob auf dieser Grundlage ein Modell aufgestellt werden kann, das die in der Motorkarte des CS beobachtete raum-zeitliche Dynamik aufweist.

Eine Kategorisierung der Dynamik in 1-dimensionalen Varianten solcher neuraler Felder wurde von Amari [12] erarbeitet. In einschichtigen Feldern konnten 5 Dynamik-Typen nachgewiesen werden, die i.a. multistabil sind. Bestimmte Felder ermöglichen die Aufrechterhaltung eines lokalisierten Erregungsmusters an der Stelle, wo ein Reiz einwirkte, auch nach seinem Verschwinden. Dieses Verhalten ist mit dem Kurzzeitgedächtnis in Verbindung gebracht worden.

Erregungsmuster können auch interagieren: zwei erregte Gebiete können einander anziehen, so daß sie zu einer lokalen Erregung verschmelzen. In anderen Fällen können sie einander abstoßen, oder sie können unabhängig voneinander bestehen. Zweischichtige Felder gestatten oszillatorische und Wanderwellen-Lösungen. Diese Ergebnisse übertragen sich auf den uns interessierenden 2-dimensionalen Fall. Allerdings folgt auch, daß die Amari-Gleichungen keine isolierten, 1-dimensionalen Trajektorien der Aktivitätshügel erlauben, wie sie in der Motorkarte des CS beobachtet wurden. Dies wird durch die Eigenschaft beliebiger kontinuierlicher neuraler Felder (homogen oder nicht) verursacht, daß die Erregungsausbreitung in Form einer Wanderwelle richtungsunabhängig ist. Um ein Modell mit den gewünschten Merkmalen zu erhalten, bieten sich bestimmte diskrete neurale Netze mit räumlich inhomogenen Kopplungen an, wie kurz dargelegt werden soll.

Aus mathematischer Sicht wird eine aus verschiedenen homogenen Schichten bestehende neurale Struktur durch ein System von Integralgleichungen mit räumlicher und zeitlicher Faltung beschrieben. Die allgemeine Theorie solcher nichtlinearer Integralgleichungen ist

nicht sehr weit entwickelt. Eine bestimmte Klasse von Faltungskernen gestattet die Transformation der Integralgleichungen in partielle Differentialgleichungen PDG [13]. Amari-Gleichungen gehören glücklicherweise zu diesem Typus, so daß wir die Resultate auf sie anwenden können. Nichtlineare PDG werden im Rahmen der Theorie erregbarer Medien untersucht. Eine herausragende Rolle spielt hierbei das FitzHugh-Nagumo (FHN)-Modell, vor allem wegen seiner Einfachheit und der qualitativen Übereinstimmung mit komplizierten Modellen [14].

Ein wichtiges Ergebnis der Untersuchungen zum 2-dimensionalen FHN-Modell ist das folgende: im Gegensatz zu kontinuierlichen, 2-dimensionalen FHN-Modellen mit beliebigen Kopplungen hängt in diskreten Modellen mit inhomogenen Kopplungen der Ausbreitungserfolg einer Erregung von der Richtung ab. Keener [15] konnte zeigen, daß ein diskretes 2-dimensionales Feld mit inhomogenen Kopplungen isolierte 1-dimensionale Trajektorien von Aktivitätshügeln unterstützt, wobei diese Gipfel sich nicht in benachbarte Feldregionen ausbreiten, wenn die Kopplungskoeffizienten in den zwei Raumrichtungen hinreichend unterschiedlich sind.

## **5. *Schlußfolgerungen***

Erregbare Medien sind schon verschiedentlich als Mittel für analoge Berechnungen vorgeschlagen worden [16]. Diese Idee beruht auf der Möglichkeit, ein bestimmtes Problem auf die Dynamik gewisser erregbarer Medien abzubilden, so daß die Lösung sich ergibt, indem man die Entwicklung des erregbaren Mediums verfolgt. Der vorliegende Bericht hat gezeigt, daß neurale Felder vom kontinuierlichen und diskreten Typ ein reiches, z.T. aber unterschiedliches dynamisches Verhalten aufweisen, das benutzt werden kann, um die Idee der analogen Berechnung zu verwirklichen. Dieses Verarbeitungsschema muß ergänzt werden durch afferente topografische Abbildungen entweder von Sinnesflächen (rezeptotope Karten) oder von gewissen Parametern wie Motorfehler im CS (komputatorische Karten). Die resultierenden kartierten neuronalen Felder verkörpern ein allgemeines Schema der Repräsentation und Verarbeitung, das besonders an die Koordinatentransformationen angepaßt ist, die bei der sensomotorischen Kontrolle auftreten. Das Hauptanliegen unseres Forschungsprojekts besteht darin, diese Ideen auf das Problem der Blickkontrolle anzuwenden. Die colliculäre Motorkarte wird dabei durch ein diskretes, inhomogenes, 2-dimensionales und zweischichtiges neurales Feld modelliert, dem eine

retinotopie Karte überlagert wird. Dieses Modell soll für die Kontrolle künstlicher Aktiver Sehsysteme eingesetzt werden.

## **Literatur**

- [1] Marr, D.: Vision. Freeman, San Francisco 1982.
- [2] Aloimonos, Y. (ed.): Active Perception. L. Erlbaum Assoc., Orlando FL, 1993.
- [3] Ballard, D.H. and Brown, C.M.: Principles of animate vision. CVGIP: Image Understanding 56 (1992) 3-21.
- [4] Grantyn, R.: Gaze control through superior colliculus: Structure and function. In: Neuroanatomy of the Oculomotor System (J. Büttner-Ennever, ed.) Elsevier, New York 1988, pp. 273-333.
- [5] Sparks, D.L. and Mays, L.E.: Signal transformations required for the generation of saccadic eye movements. Ann. Rev. Neurosci. 13 (1990) 309-336.
- [6] Berthoz, A., Grantyn, A. and Droulez, J.: Some collicular neurons code saccade velocity. Neurosci. Lett. 72 (1987) 289-294.
- [7] Hilbig, H. and Schierwagen, A.: Interlayer neurons in the superior colliculus of the rat - a tracer study using DiI/Di-ASP. NeuroReport 5 (1994) 477-480.
- [8] Munoz, D.P. and Guitton, D.: Fixation and orientation control by the tecto-reticulospinal system in the cat whose head is unrestrained. Rev. Neurol. (Paris) 145 (1989) 567-579.
- [9] Munoz, D.P., Pelisson, D. and Guitton, D.: Movement of neural activity on the superior colliculus motor map during gaze shifts. Science 251 (1991) 1358-1360.
- [10] Mallot, H.A., von Seelen, W. and Giannakopoulos, F.: Neural mapping and space-variant processing. Neural Networks 3 (1990) 245-263.
- [11] Knudsen, E.I., du Lac, S. and Esterly, S.D.: Computational maps in the brain. Ann. Rev. Neurosci. 10 (1987) 41-65.
- [12] Amari, S.: Dynamics of pattern formation in lateral-inhibition type neural fields. Biol. Cybernetics 27 (1977) 77-87.
- [13] van der Heiden, U.: Analysis of Neural Networks. Springer-Verlag, Berlin Heidelberg New York 1980.
- [14] Schierwagen, A.K. and Francu, J.: The continuum approach as applied to wave phenomena in physiological systems. In: Nonlinear Waves in Active Media (J. Engelbrecht, ed.) Springer-Verlag, Berlin Heidelberg New York Tokyo 1989, pp. 185-191.
- [15] Keener, J.P.: On the formation of circulating patterns of excitation in anisotropic excitable media. J. Math. Biol. 26 (1988) 41-56.
- [16] Mikhailov, A.S.: Foundations of Synergetics I. Springer-Verlag, Berlin etc. 1990.