

Diss. ETH No. 14681

**A Robot using a Modular Navigation
System:
Attentional Manoeuvring based on an
Egocentric Spatial Representation**

A dissertation submitted to the
SWISS FEDERAL INSTITUTE OF TECHNOLOGY
(ETH ZURICH)

for the degree of
Doctor of Natural Sciences

presented by
REGINA MUDRA
Dipl.-Ing. for Electrical Engineering, TU Berlin
born January 6, 1970
citizen of Berlin, Germany

accepted on the recommendation of:
Prof. Dr. Rodney J. Douglas, examiner
Prof. Dr. Rolf Pfeifer, co-examiner

2002

Abstract

The navigation systems of rodents are based on the correlation of different representations containing information about their intention, sensory awareness, physical response and the environment at each moment during movement. Their navigation systems provide them with the ability to move through known and unknown environments. An allocentric representation of the environment seems to have an important role within the navigation system. The firing patterns of the place cells in the rodent hippocampus (CA3) are interpreted as containing such a representation. To date, mainly attractor networks have been used to model such a representation of the environment. So far it is unknown how this representation is built and updated.

In this work an artificial modular navigation system has been developed in which representations of visual input and the environment are combined to influence the behaviour of a robot. The visual representation is built on the basis of external sensory information, while the spatial representation is an internal representation. These representations are encoded within recurrently coupled neuronal network maps that are spanned by pointer neurons. The vector output of the recurrent pointer networks can encode the position of the nearest object or of a specific location in the environment. The robot manoeuvres through the environment in relation to the attended position. Based on its egocentric spatial representation, the robot's position is correlated with the other representations and a motor map is activated. Motor feedback is used to update the egocentric spatial representation by changing the strengths of nearest neighbour connections within the neuronal map.

The accuracy of the robot's navigation based on the modular navigation system including an egocentric spatial representation is tested using three different approaches. Firstly, an avoidance task is implemented by combining the visual representation with the motor map. Secondly, the use of the egocentric spatial representation and its update in combination with the motor map is tested by having the robot move towards specified locations within the environment. Thirdly, by correlating the visual representation and the egocentric spatial representation, the robot's accuracy in manoeuvring is improved.

The results show that the robot can use a vector encoding of the position of an object within its visual field, and so attend to the nearest object in performing an avoidance task. Apart from this qualitative test, the accuracy of the vector encoding (which is very important for the use of the egocentric spatial representation) has also been tested. The robot had to reach a pre-defined goal position that was represented as activity within an egocentric spatial representation. The robot could attend to the goal and find its way to the goal. The representation of the goal was updated during the task. The accuracy of reaching a specific goal position has been improved in comparison with a robot using a path integration algorithm alone. In a statistical analysis, the results of a goal task performed by a robot on the basis of the egocentric spatial representation within the modular navigation system are compared with the case where the robot uses an implementation

of the path integration algorithm (the control task). The results show a better performance of the robot in repeatedly moving towards a defined goal position. The angular error is reduced to a quarter of that in the control task, while the error in distance is not reduced. In the third experiment, in which the modular navigation system using a combination of egocentric spatial representation and visual representation has been tested in the goal task, the errors in the end position are again reduced. The angular error is reduced to an eighth of that in the control task. The advantage of this modular navigation system is the representation independent position encoding, for example in the encoding of the nearest object in the visual field and of the goal position in the egocentric spatial representation. In the last task these positions are one and the same and so a direct combination of the vector outputs is possible. This has been realized by using recurrent pointer networks to build the different representational maps. The 'where' information of the vector outputs encoding positions could be combined to improve the navigation abilities of the robot. The egocentric spatial representation within the modular navigation system allows the correlation of the sensory information in relation to the body of the robot, called embodiment. The major advantage is that the egocentric spatial representation provides the robot with a sense of space and provides information about egocentric bearings and distances to objects. This is necessary for the robot to start building a map based on features of the explored environment. In the future it would be possible to extend the modular navigation system by integrating further sensory representations and environmental representations. Furthermore, the encoding principle in the representations could be modified so that more than one object or location can be handled. This property is important if the robot should maintain an allocentric space representation of an environment during an exploration task.

Keywords: robotics, recurrent neural pointer networks, attention, navigation, egocentric spatial representation, embodiment

Zusammenfassung

Das Navigationssystem von Nagetieren basiert auf der Korrelation von verschiedenen Repräsentationen, die zum einen Wissen über das Ego und seine Absichten, sowie Kenntnisse der sensorischen Wahrnehmungen und physische Reaktionsfähigkeit zu jedem Zeitpunkt der Bewegung innerhalb einer Umgebung beinhalten. Das Tier kann sich in bekannter und unbekannter Umgebung zurechtfinden. Die aktuelle Hirnforschung zeigt, dass der Hippocampus von Nagetieren eine allozentrische Repräsentation der Umgebung im CA3 beinhaltet, in der die aktuelle Eigenposition des Tieres als Zentrum höchster Aktivität innerhalb des neuronalen Verbandes gekennzeichnet ist. Als mathematische Modelle dienen wiederkehrende (recurrent), neuronale Netzwerke, die die Position höchster neuronaler Aktivität vektoriell kodieren. Der aktuelle Wissenstand liefert noch keine schlüssige Erklärung wie Repräsentationen der Umgebungen von den Tieren gebildet oder genutzt werden.

In dieser Dissertation wurde ein künstliches modulares Navigationssystem entwickelt, in dem eine Repräsentation der visuellen Wahrnehmung mit einer räumlichen Vorstellung der Umgebung überlagert und das Verhalten eines Roboters bestimmt wird. Dargestellt sind die verschiedenen Repräsentationen als neuronale Karten innerhalb wiederkehrender, neuronaler Netzwerke, deren Ausdehnung und Form durch Zeigerneurone festgelegt werden. Der Ausgabevektor des wiederkehrenden neuronalen Zeigernetzwerks kodiert die Position des nächsten Hindernisses oder einer bestimmten Position in der Umgebung. In Relation zu der mit Aufmerksamkeit belegten Position manövriert dann der Roboter innerhalb der Umgebung. Auf der Basis der räumlichen Repräsentation, die in der Form einer egozentrischen Raumkarte dargestellt ist, geschieht die Korrelation mit den anderen Repräsentationen und bestimmt die Motorenansteuerung. Die Messungen der jeweils zurückgelegten Distanz werden verwendet, um die egozentrische Karte abzugleichen, indem die Gewichtungen zwischen den Kartenneuronen verändert werden, wodurch sich die Position höchster Aktivität entgegen der ablaufenden Bewegungsrichtung des Roboters verschiebt.

Die Vektoren der verschiedenen Repräsentationen wurden innerhalb des modularen Navigationssystems korreliert und in drei verschiedenen Kombinationen getestet: Zum ersten wurde im Ablauf des Ausweichverhaltens die Kombination von Repräsentation visueller Wahrnehmung und Motorkarte getestet. In der zweiten Versuchsreihe wurde die egozentrische Repräsentation in Verbindung mit der Motorkarte getestet, um den Roboter zu definierten Positionen in der Umgebung fahren zu lassen. Im dritten Versuch wurde die Zielgenauigkeit des Roboters durch die Kombination von visueller und egozentrischer Karte verbessert. Im Anschluss wird das Zusammenspiel der verschiedenen Repräsentationen im modularen Navigationssystem anhand von einzelnen Fallbeispielen diskutiert.

Die Resulte zeigen, dass der Roboter die vektorkodierte Position eines Objektes innerhalb des visuellen Sichtfeldes nutzt, in dem er in Bezug zu diesem wahrgenommenen Objekt eine Pfad-

korrektur vornehmen kann. Neben diesem qualitativen Test wurde auch die Genauigkeit der Vektorkodierung im Zusammenhang mit der egozentrischen Umweltrepräsentation geprüft. Der Roboter musste mehrfach zu bestimmten Positionen im Raum fahren, die innerhalb der egozentrischen Umweltrepräsentation durch gezielte Erregung eines Zielneurons vorgegeben wurden. Der Roboter nutzt die Vektorrepräsentation des markierten Ziels um den Weg zum Ziel zu finden. Im Ablauf der Roboter Bewegung wurde fortwährend, entsprechend der zurückgelegten Distanz, die Zielrepräsentation verschoben, bis der Roboter das Ziel erreicht hat. Die Zielgenauigkeit eines Roboters auf der Basis der egozentrischen Umweltrepräsentation hat sich gegenüber einem Roboter, der Pfadintegration benutzt, verbessert. In der statistischen Auswertung wurden die Navigationsergebnisse des Roboters unter Anwendung der egozentrischen Umweltrepräsentation im modularen Navigationssystem mit den Ergebnissen eines Kontrollversuchs, Implementation von Pfadintegration verglichen. Die Ergebnisse belegen eine verbesserte Navigationsfähigkeit des Roboters beim Erreichen einer Zielposition. So konnte der Fehler im Winkel auf ein Viertel reduziert werden gegenüber der Winkelabweichung im Kontrollversuch, während sich der Fehler in der Distanz nicht minimieren liess. Diese Fehler konnten in der dritten Versuchsreihe, wo das Zusammenspiel von egozentrischer Umweltrepräsentation und visueller Repräsentation im gleichen Navigationsexperiment getestet wurde, nochmals minimiert werden. Der Mittelwert des Winkelfehlers konnte auf ein Achtel des Fehlers des Pfadintegrationsversuchs gesenkt werden und auch für die Distanz ergab sich eine Verbesserung, da der Mittelwert der erreichten Distanz nun der vorgegeben Zieldistanz entspricht.

Der Vorzug des modularen Navigationssystems besteht in der von der Repräsentationsart unabhängigen Positionskodierung, z. B. die Kodierung des nächsten Hindernisses im Sichtbereich oder des Ziels innerhalb der egozentrischen Umweltrepräsentation. Innerhalb der letzten Aufgabenstellung sind diese Positionen ein und dieselbe und somit lassen sich die Vektoren direkt kombinieren. Die Versuchsergebnisse zeigen, dass wiederkehrende neuronale Zeigernetzwerke für die unterschiedlichen Repräsentationen benutzt werden können. Deren Ausgabevektoren kodieren alle Positionen, d.h. eine ortsbezogene Information, unabhängig vom dargestellten Inhalt. Diese örtlichen Informationen können verrechnet werden und somit lässt sich die Navigationsfähigkeit des Roboters verbessern. Die Korrelation der verschiedenen Sensorinhalte erfolgt auf der Basis der egozentrischen Umweltrepräsentation und in Relation zum Körper des Roboters, genannt Embodiment. Ein weiterer Vorteil ist, dass die egozentrische Umweltrepräsentation dem Roboter ein räumliches Vorstellungsvermögen vermittelt und sich somit seine Ausrichtung und Distanz zu Objekten bestimmen lässt. Das sind notwendige Informationen für einen Roboter, der während der Erkundung einer Umgebung eine Karte von derselben zu erstellen hat.

In Zukunft wird es möglich sein, das modulare Navigationssystem um weitere sensorische und auch weitere Umgebungsrepräsentationen zu ergänzen. Desweiteren ist es möglich das Prinzip der Repräsentation so zu verändern, dass mehr als ein Objekt oder eine Position dargestellt werden kann. Dies spielt auch eine Rolle, wenn allgemeingültige Raumkarten von Umgebungen vom Roboter während eines Erkundungsganges erstellt werden sollen.

Suchbegriffe: Robotik, wiederkehrende (recurrent) neuronale Zeiger Netzwerke, Aufmerksamkeit, Navigationssystem, egozentrische Umweltrepräsentation, Embodiment