Orte - Routen - Karten

Aufbau, Struktur und Gebrauch von Ortsrepräsentationen in der Raumkognition

der Fakultät für Biologie
der EBERHARD KARLS UNIVERSITÄT TÜBINGEN

zur Erlangung des Grades eines Doktors
der Naturwissenschaften

von

Kai Basten

aus Lich

vorgelegte
Dissertation
2010

Tag der mündlichen Prüfung:	17.Mai 2010
Dekan:	Prof. Dr. Hanspeter A. Mallot
1. Berichterstatter:	Prof. Dr. Hanspeter A. Mallot
2. Berichterstatter:	Prof. Dr. Heinrich H. Bülthoff

Kurzzusammenfassung

Für das Leben und Überleben von Tieren und Menschen sind nur wenige kognitive Mechanismen so fundamental wie die Fähigkeit sich in der Umwelt zu orientieren. Diese Orientierungsfähigkeit bildete den Rahmen der vorliegenden Arbeit, die sich auf die Erkennung und die kognitive Repräsentation von Orten konzentrierte. Dass einzelne Orte bereits mit geringer sensorischer und verarbeitender Kapazität erkannt und unterschieden werden können, konnte in einer Computersimulation zum Orientierungsverhalten von Wüstenameisen gezeigt werden. Einfache visuelle Merkmale, wie sie in der natürlichen Umgebung der Tiere vorkommen, sind dabei für die Orientierung ausreichend. In einem Verhaltensexperiment beim Menschen konnte nicht nur der Aufbau von einzelnen Ortsrepräsentationen, sondern auch deren Integration in Routen- und Kartenrepräsentationen beobachtet werden. Dabei zeigte sich, dass abhängig davon, ob vermehrt lokale oder globale Strukturen zur Orientierung genutzt wurden, auch verstärkt eine Routen- oder Kartenrepräsentation aufgebaut wurde. Die verschiedenen Raumrepräsentationen selbst stehen in Verbindung mit unterschiedlich komplexen Verhaltensmustern und der Interaktion zwischen Arbeits- und Langzeitgedächtnis. Auf Grundlage dieser Zusammenhänge wurde ein allgemeiner Rahmen formuliert, in dem Orientierungsverhalten, ökologische Voraussetzungen und neuronale Mechanismen diskutiert wurden. In einem weiteren Experiment konnte gezeigt werden, dass der Abruf von Ortswissen aus dem Langzeitgedächtnis in perspektivischer egozentrischer Form erfolgt, wobei eine bevorzugte Perspektive nicht nur von den Strukturen eines Ortes bestimmt wurde, sondern auch situations- und aufgabenabhängig war.

Die einzelnen Arbeiten der vorliegenden kumulativen Dissertation zeigen die unterschiedlichen Stufen von einfacher Ortserkennung über den Aufbau einer Raumrepräsentation bis hin zu Verwendung solcher Repräsentationen aus dem Langzeitgedächtnis. Dabei wurden sowohl bestehende Konzepte der Raumkognition diskutiert und ergänzt, als auch weitere Grundlagen für die Entwicklung theoretischer Modelle und künstlicher Systeme geschaffen.

Manuskripte/Artikel der Dissertation:

K. Basten, H. A. Mallot (accepted), *Simulated visual homing in desert ant natural environments: efficiency of skyline cues*, Biological Cybernetics

R. Hurlebaus, K. Basten, H. A. Mallot, and J.M. Wiener (2008), *Route learning strategies in a virtual cluttered environment.* In: C. Freksa et al., (Ed.), Spatial Cognition (SC) 2008, number 5248 in Lecture Notes in Artificial Intelligence, 104-120.

K. Basten, H. A. Mallot (in prep.), *Trajectory discretisation and sequence analysis for wayfinding experiments in spatial cognition*

H. A. Mallot, K. Basten (2009), *Embodied spatial cognition: Biological and artificial systems*, *Image and Vision Computing*, 27, 1658-1670

K. Basten, T. Meilinger, H. A. Mallot (submitted), *Mental Travel Primes Place Orientation in Spatial Recall - a Novel Demonstration of Situatedness in Spatial Cognition*

Inhaltsverzeichnis

Einleitung

Aufbau der Arbeit

Die vorliegende Arbeit wird in Form einer kumulativen Dissertation eingereicht und ist in fünf Abschnitte gegliedert. Jeder Abschnitt entspricht einem Manuskript, von denen drei bereits veröffentlicht bzw. für die Veröffentlichung akzeptiert sind und zwei weitere zur Veröffentlichung vorbereitet bzw. eingereicht wurden. In diesen Artikeln sind folgende Projekte dargestellt:

1. Im ersten Artikel wird ein Modell zur visuellen Landmarkenorientierung bei Wüstenameisen vorgestellt. Während viele Verhaltensexperimente zeigen, dass sich die Ameisen in ihren natürlichen Umgebungen auch visuell orientieren können, ist noch nicht vollständigt geklärt, welche visuellen Merkmale dabei genutz werden. Dieser Frage wurde mittels einer Computersimulation nachgegangen, in der die Umgebung der Ameisen anhand von Kartierungen aus Feldversuchen modelliert wurde. In dieser Simulation wurde die visuelle Verarbeitung der Ameise modelliert und damit zwei Verfahren zur Ortserkennung entwickelt und getestet.

2. Der Erwerb von Routenwissen beim Menschen wurde im zweiten Projekt der Arbeit untersucht. Am Computer mussten Versuchspersonen den Weg zwischen zwei Zielpunkten finden und lernen. Die dazu präsentierte virtuelle Umgebung bestand aus vielen sehr ähnlichen Hindernissen, die nicht nur eine Route, sondern viele mögliche Routen zuließen. Die Versuchspersonen konnten sich dabei sowohl an lokalen, als auch an globalen Landmarken orientieren. Mit diesem Experiment wurde untersucht, welche Landmarken beziehungsweise Orientierungsstrategien beim Routenlernen in solchen komplexen Umgebungen welche Rolle spielen.

3. Im dritten Projekt wird ein allgemeines Verfahren zur Auswertung von Laufspuren und Routen vorgestellt. Im Kern des Verfahrens werden in einer beliebigen Umgebung diejenigen Orte bestimmt, an denen Navigationsentscheidungen getroffen werden. Routen aus unterschiedlichen Versuchsdurchläufen oder von unterschiedlichen Individuen können damit auf Grund lokaler Navigationsentscheidungen verglichen werden.

4. Im vierten Artikel der Arbeit wird ein generelles Konzept zur Einbettung von Umgebungswissen in ein Ortsgedächtnis dargestellt. Hierbei sind einzelne Orte durch eine Graphenstruktur miteinander verbunden. Teile des Graphen können wiederum durch Regionen in größere Bereiche zusammengefasst werden. Ortsspezifische Ansichten, lokale und globale metrische Informationen, Aktions- oder Bewegungsanweisungen und andere Eigenschaften können an den Graph beziehungsweise an Orte und deren Verbindungen assoziiert werden.

5. In welcher Struktur oder Orientierung liegt Ortswissen im Langzeitgedächtnis vor und welchen Einfluss auf den Abruf dieses Ortswissen hat eine aktuelle Aufgabe, wie z.B. eine Routenplanung? Dieser Frage wurde im letzten Teil dieser Arbeit nachgegangen. Versuchspersonen skizzierten dabei einen bekannten Platz ihres Wohnortes, hier den Tübinger Holzmarkt. Untersucht wurde dabei, in welcher Orientierung die Skizzen auf dem Fragebogen gezeichnet wurden. In einem zweiten Versuch wurde getestet, ob sich die Orientierung des Ortsgedächtnisses beziehungsweise die Orientierung der Skizzen in Abhängigkeit von einer räumlichen Aufgabe ändert. Dazu sollten die Versuchspersonen diesmal vor dem Zeichnen gedanklich einer Route folgen, die den gefragten Platz quert. Die Orientierungen der Skizzen aus den unterschiedlichen Versuchsbedingungen wurden anschließend miteinander verglichen.

Diesen fünf Artikeln ist zunächst eine allgemeine Einführung zu den Themen Navigation und zu grundlegenden Orientierungsmechanismen vorangestellt. Am Anfang jeden Kapitels findet sich vor den einzelnen Artikeln eine kurze Zusammenfassung des jeweiligen Projekts einschließlich einer Darstellung des eigenen Anteils an der entsprechenden Arbeit. Am Ende der Arbeit findet sich abschließend noch eine kurze Zusammenfassung der Ergebnisse.

Einführung: Orientierung und Navigation

Für das Leben und Überleben von Tieren und Menschen sind nur wenige kognitive Fähigkeiten so fundamental wie die Fähigkeit, sich im Raum zu orientieren. Menschen und Tiere bewegen sich in ihrer Umgebung, um Nahrung, Partner und Schutz zu suchen, oder andere Ressourcen zu nutzen. Ebenso müssen Orte, die mit Gefahren verbunden sind, z.B. durch Fressfeinde, gemieden werden. Eine zufällige Bewegung im Raum ist für die meisten Organismen keine vorteilhafte Strategie, da Zeit, Nahrung und Ausdauer nicht in beliebiger Menge vorhanden sind. Daher ist eine koordinierte und zielgerichtete Bewegung in der Umwelt erforderlich. Allgemein kann dieses Verhalten mit dem Begriff 'Navigation' beschrieben werden.

Geprägt wurde der Begriff in der Nautik und ist vom lateinischen Wort *navigare* (dt. ein Schiff führen) abgeleitet. Die Navigationsleistungen der Seefahrer zeigen dabei bereits, wie vielfältig Navigationsstrategien sein können. So werden beispielsweise seit langer Zeit zur Navigation auf See der Sonnenstand und Sternenbilder genutzt, wie schon Homer aufschlussreich in seiner Odyssee beschreibt:

> „*Und der [Odysseus] saß und steuerte kunstgerecht mit dem Ruder, und seine Lieder befiel kein Schlaf, wenn er nach den Pleiaden [Siebengestirn] Ausschau hielt und dem späteintauchenden Sternbild Bootes und der Bärin, die auch mit dem Namen Wagen benannt wird. Die kreist um sich selber und späht zum Orion hinüber, sie, als einzige nicht ins Bad des Okeanos eintaucht. Sie zur Linken zu halten, wenn er das Meer überquere, hatte ihm angeraten die hehre Göttin Kalypso.*“ (Homer, ca. 8.Jhd. v. Chr.)

Sonne und Sternenbilder können genutzt werden, um einen Kurs zu halten. Sie dienen allerdings nur für eine grobe Positionsschätzung. Zur genaueren lokalen Positionsbestimmung wurden und werden in der Seefahrt teilweise noch Strömungen und Wassertiefen gemessen, Temperatur und Färbung des Wassers beurteilt und vor allem Küstenlinien, Inseln und Riffe zur Orientierung verwendet (Johnson and Nurminen (2008)). Vor allem, wenn große Distanzen zurückgelegt werden, können auch Flora und Fauna (z.B. bestimmte Vogelarten) nützliche Hinweise auf die eigene Position geben.

Antike Seefahrer nutzten nicht nur solche Informationen, sondern hielten sie auch schriftlich in einer so gennanten *Periplus* fest ($\pi\epsilon\rho\iota\pi\lambda o\upsilon\varsigma$, gr. Küstenfahrt, Umsegelung). Periploi beschrieben Seerouten in Form einer Liste aufeinanderfolgender Landmarken und enthielten teilweise zusätzlich Schilderungen über ortsübliche Handelswaren oder sonstige lokale Eigenheiten (vgl. Periplus Hannonis, Bayer (1993)).

Aus dem Inselgebiet des heutigen Mikronesien ist bekannt, dass zur Navigation auf See sogar die Form von Wellenmustern und Dünungen genutzt werden kann, die durch Reflexion und Brechung der Wellen an Inseln und Atollen zustande kommen (Spennemann (2005)). Dieses Wissen wurde (und wird teilweise noch) vorwiegend zu Lehrzwecken in speziellen Karten, sogenannten Stockkarten, dargestellt (Johnson and Nurminen (2008), s. Abb. 1).

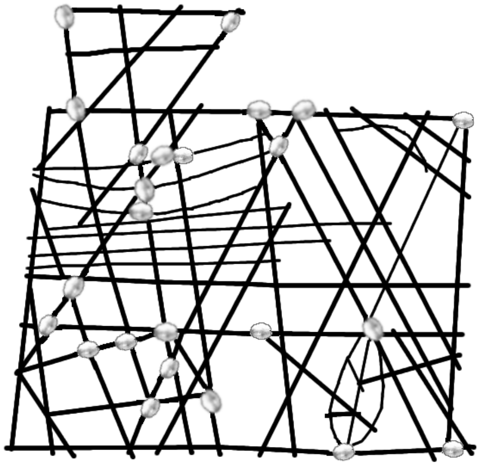


Abbildung 1: Stockkarte von den Marschalinseln. Palmrippen verbunden mit Kokosfasern zeigen Dünungs- und Strömungsmuster. Aufgesteckte Muscheln geben die (relativen) Positionen einzelner Inseln zueinander an. (Graphik nach Johnson and Nurminen (2008), S. 21)

Etwa seit dem 11. Jahrhundert n. Chr. kam in der Seefahrt vermehrt der magnetische Kompass zur Anwendung (Aczel (2005)). Im 18. Jahrhundert wurde der Sextant entwickelt, mit dessen Hilfe der Winkelabstand zwischen dem Horizont und Gestirnen - der Höhenwinkel - gemessen werden konnte. Die eigene Position wurde dann mit diesen Werten mittels Tabellen aus einem Nachschlagewerk (z.B. *Nautical Almanac*) bestimmt. Dies verbesserte die Navigation nach Sternenbildern deutlich. Denn es konnte nicht nur ein Kurs festgelegt werden, sondern durch mehrfache Messungen konnte die Eigenposition mit einer Genauigkeit von einer Bogenminute (1/60 Grad) bzw. einer Seemeile (1.85 km) bestimmt werden.

Seit Ende der 1960er Jahre sind durch Satellitennavigationssysteme letztendlich nicht nur in der Schifffahrt weltweit eindeutige Standortbestimmungen möglich.

Das Beispiel der Seefahrt zeigt, wie auf vielfältige Weise die Eigenposition oder zumindest eine Kursrichtung mittels natürlicher Landmarken bestimmt werden kann. Die Navigation beinhaltet dabei mehrere Aspekte, die aus Sicht des Navigierenden für das Weitere in folgenden Grundfragen formuliert werden können (Levitt and Lawton (1990)):

- Wo bin ich?
- Wo sind andere Orte relativ zu meiner Position?
- Wie gelange ich an diese Orte?

Navigations- und Orientierungsstrategien

Im Tierreich ist die Fähigkeit, sich im Raum orientieren zu können, weit verbreitet. Allerdings finden sich dabei ganz unterschiedliche Komplexitätsstufen zwischen einzelnen Navigationsfähigkeiten. Zur Orientierung gehört genauso der einfache Reiz-Reaktions-Mechanismus, wie auch der Aufbau eines komplexen Ortsgedächtnisses oder die Planung von neuen Wegen.

In dieser Einführung gebe ich einen Überblick über einzelne Navigationsmechanismen und Navigationsstrategien. Im Folgenden beginne ich bei einfachen, weit verbreiteten Fähigkeiten und beschreibe dann zunehmend komplexere Mechanismen, bis hin zur Verwendung von (Straßen-) Karten, die nur beim Menschen Anwendung finden.

Einfache Orientierungsmechanismen

In diesem Abschnitt werden zunächst die einfachsten Mechanismen zur Orientierung beschrieben, beginnend bei der Taxis (s.u.), die schon bei Einzellern zu finden ist. Neben der Taxis kann ein Individuum auch einfach einer Richtung (z. B. der eines Magnetkompasses) folgen. Ein weiterer Mechanismus einfacher Art ist die Wegintegration. Sie befähigt ein Individuum, während einer Exkursion kontinuierlich die Eigenbewegung so zu integrieren oder zu berechnen, dass ständig die Richtung und Distanz zum Ausgangsort bekannt ist. Ganz unabhängig von äußeren Einflüssen ist die systematische Suche. Durch sie kann ein Gebiet effizient nach einem Ziel abgesucht werden.

Taxis

Mit dem Begriff Taxis bezeichnet man ein Orientierungsverhalten, bei dem sich ein Lebewesen zu einer bestimmten Reizquelle (wie z. B. eine Licht- oder Wärmequelle) oder einem Reizgradienten hinwendet (positiv) oder abwendet (negativ) (vgl. Murlis et al. (1992); Paola et al. (2004)).

Viele Ameisenarten zeigen eine Chemotaxis, indem sie Pheromonspuren nutzen, mit denen sie Wege und Routen zu einer Nahrungsquelle markieren (Wilson (1963)). Dabei kann durch Modulation der Stärke der Spur die Qualität einer Nahrungsquelle kodiert werden (Sumpter and Beekman (2003); Jackson and Chaline (2007)).

Die Pheromonspuren können auch Verzweigungen aufweisen, vor allem, wenn unterschiedliche Nahrungsquellen vorhanden sind. Die Pharaoameise *Monomorium pharaonis* entwickelt dabei ein ausgeklügeltes Netzwerk von Pheromonspuren. Hier gibt der Winkel zwischen aufeinandertreffenden Pfaden Aufschluss darüber, in welcher Richtung die Nahrung und in welcher Richtung das Nest zu finden ist (Jackson et al. (2004)). Diese Ameisen nutzten sogar verschiedene Pheromone, um zum Beispiel ein Stoppsignal zu setzen, das die Nutzung eines nicht mehr ertragreichen Pfades beendet (Robinson et al. (2005)).

Der Beriff der Taxis allein wird dem Verhalten der Ameisen nicht gerecht, da einer Pheromonspur auch gefolgt wird, ohne dass diese Spur in sich einen Gradienten aufweist. Daher ist die Taxis hier noch mit einem Folgeverhalten kombiniert, das im nächsten Abschnitt erläutert wird.

Der an sich einfache Mechanismus der Taxis kann ein (zumindest von außen betrachtet) sehr komplexes Verhalten hervorbringen, vor allem dann, wenn sich mehrere Taxien überlagern. Eindrucksvoll wurde dies von Braitenberg (1984) mit Hilfe einfacher Vehikel (d.h. kleiner Fahrzeuge) deutlich gemacht. Die Antriebsmotoren der einzelnen Räder des Vehikels werden dabei direkt über die Stärke der Signale an den unterschiedlichen Sensoren gesteuert. Ein Sensor kann dabei mit steigender Signalstärke den Motor eines Antriebsrads je nach Verschaltung beschleunigen oder auch abbremsen. Werden zwei Sensoren beispielsweise vorne links und rechts am Vehikel positioniert und mit dem Motor auf ihrer Seite verbunden, so kann sich je nach beschleunigender oder bremsender Verschaltung das Vehikel von einer Reizquelle abwenden und sich dem Reiz zuwenden. Werden nun mehrere Sensoren mit den Antriebsmotoren verbunden, so überlagern sich die einzelnen Eingänge und können dadurch ein komplexeres Verhalten erzeugen, das durch einen einzelnen Sensor nicht erklärbar wäre. So könnte zum Beispiel mit zwei einfachen Sensoren und einfacher Verschaltung gleichzeitig Licht gemieden und Wärme bevorzugt werden.

Die Navigationsmöglichkeiten mittels Taxis sind vor allem dadurch begrenzt, da die Fähigkeit fehlt, Informationen räumlich oder zeitlich zu integrieren und zu speichern. Somit ist dieser Mechanismus unabhängig vom Arbeits- und Langzeitgedächtnis.

Richtungsfolgen

Für diese Art der Navigation muss ein Individuum in der Lage sein, seinen Kurs in eine bestimmte Richtung auszurichten. Dazu kann zum Beispiel eine Pheromonspur (wie oben beschrieben) oder das Magnetfeld der Erde genutzt werden, das viele Tiere, darunter Mollusken, Arthropoden und Vertebraten, wahrnehmen können (Wiltschko and Wiltschko (2005)). Der angestrebte Zielort selbst muss dabei nicht zwingend direkt wahrnehmbar sein. Das Richtungsfolgen führt allerdings nur erfolgreich zum Ziel, solange sich das Individuum auf dem zielführenden Pfad befindet. Ein Versatz parallel dazu führt in der Regel zum Verfehlen des Zielortes, wenn nicht weitere Navigationsmechanismen genutzt werden können.

Wegintegration

Die Wegintegration befähig ein Individuum während einer Exkursion, die eigene Position relativ zum Startpunkt zu bestimmen. Dabei werden Rotationen und Translationen entlang eines Pfades kontinuierlich in einem Vektor verrechnet, der die Richtung und Entfernung zum Ausgangspunkt angibt. Bei vielen Tieren beginnen Exkursio-

Wegintegration beim Menschen		
Aufgabenstellung	Modus	Referenz
Laufen zu einer zuvor präsentierten LM + ohne Sicht	a	Rieser et al. (1990)
Laufen zu einer zuvor präsentierten LM + LM Feedback + ohne Sicht	a	Philbeck and O'Leary (2005)
Dreiecksvervollständigung oder -reproduktion + ohne Sicht; auch Patienten†	a	Péruch et al. (2005)
Dreiecksvervollständigung + ohne Sicht	a	May and Klatzky (2000)
Dreiecksvervollständigung + ohne Sicht	a	Loomis et al. (1993)
Dreiecksvervollständigung + blinde VP	a	
Dreiecksvervollständigung + ohne Sicht	a	Klatzky et al. (1999)
Dreiecksvervollständigung		Kearns et al. (2002)
+ ohne Sicht (HMD)	a	
+ optischer Fluss (HMD)	a	
+ optischer Fluss (HMD)	p	
Weg mit unterschiedlich vielen Segmenten + Leinwand	p	Wiener and Mallot (2006)
einfache Drehungen und Strecken o. Dreiecksvervollständigung + Leinwand	p	Riecke et al. (2002)

Tabelle 1: Beispiele zu unterschiedlichen Versuchsbedingungen und Aufgabenstellungen in Wegintegrationsstudien beim Menschen (a: aktives Laufen, p: keine Eigenbewegung, HMD: head-mounted display, LM: Landmarke, VP: Versuchsperson, †: Morbus Meniére + unilaterale vestibuläre Neurotomie)

nen (z.B. Nahrungssuche) an ihrem Nest. Daher hat sich für diesen Vektor auch die Bezeichnung 'Heimvektor' etabliert.

Die Fähigkeit zur Wegintegration ist weit verbreitet und wurde unter anderem nachgewiesen bei Insekten (Wehner (1992); Durier and Rivault (1999); Collett and Collett (2000); Wehner (2003)), bei Spinnen (Moller and Görner (1994); Nørgaard et al. (2003)) und bei Säugetieren (Mittelstaedt and Mittelstaedt (1980); Etienne et al. (1996); Bures et al. (1998); Seguinot et al. (1998); Cattet and Etienne (2004); Etienne and Jeffery (2004)), einschließlich des Menschen (siehe Tabelle 1).

Zur Bestimmung der Eigenbewegung können unterschiedliche sensorische Quellen genutzt werden, wie beispielsweise das Vestibularsystem und Propriozeption, also die Stellung und Bewegung der Gelenke und Muskulatur. Viele Insekten hingegen bestimmen Körperdrehungen beziehungsweise die eigenen Körperausrichtung anhand des Sonnenstandes und dem Polarisationsmuster des Himmels (Wehner (2003)).

Fluginsekten wie die Biene nutzen zur Distanzschätzung das durch Eigenbewegung induzierte optische Flussfeld (Esch and Burns (1995); Srinivasan et al. (2000)). Auch Menschen können das optische Flussfeld zur Wegintegration nutzen (Bremmer and Lappe (1999); Kearns et al. (2002); Riecke et al. (2002)). Dies ist vor allem für Navigationsexperimente in virtuellen Umgebungen am Computerbildschirm wichtig, da hier vestibuläre und propriozeptive Informationen fehlen und die Versuchspersonen daher ihre (virtuelle) Eigenbewegung am präsentierten optischen Fluss abschätzen müssen.

Die Wegintegration stellt allerdings kein perfektes System zur Bestimmung der Heimrichtung dar, sondern weist charakteristische Fehler auf. Allgemein zeigt sich die Tendenz, kleine Distanzen und Drehungen zu überschätzen, hingegen große Distanzen und Drehungen eher zu unterschätzen. Die Entfernung, beziehungsweise der Rotationswinkel, bei der die Überschätzung in eine Unterschätzung übergeht, variiert je nach Versuchsbedingung und Aufgabenstellung und liegt beim Menschen etwa bei einer Distanz von 5-10 Metern und einem Drehwinkel von $30° - 120°$ Grad (s.a. Loomis et al. (2001); Durgin et al. (2009)).

Die Ungenauigkeit des Wegintegrators ist dabei als Summe mehrer Fehlerkomponenten zu betrachten. Dazu zählen die sensorischen Eingänge, der Integrationsprozess und auch die Reproduktion beziehungsweise die Wiedergabe des berechneten Vektors durch die Motorik.

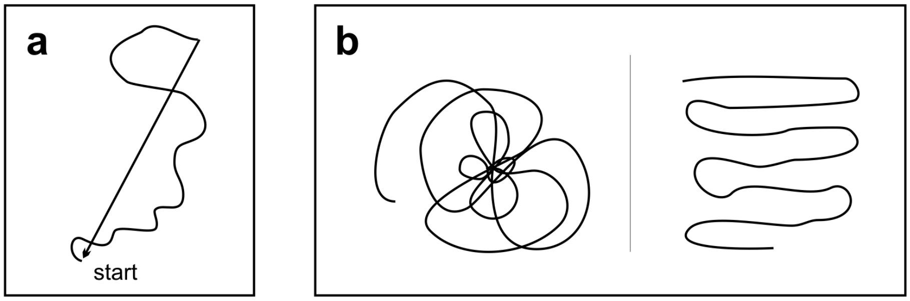


Abbildung 2: (**a**) Nach einer gewundenen Exkursion erlaubt die Wegintegration eine direkte Rückkehr zum Ausganspunkt. (**b**) Systematische Suchmuster folgen z.B. immer größer werdenden Schleifen (*links*) oder einem mäanderatigen Absuchen eines Gebiets (*rechts*), siehe Text.

Systematische Suche

Neben der Wegintegration weisen auch andere Navigationsmechanismen oft Ungenauigkeiten auf. Durch diese Ungenauigkeiten kann beispielsweise bei der Rückkehr von einer Exkursion der Ausgangspunkt, z.B. das Nest, verfehlt werden. Um das verfehlte Ziel dennoch zu finden, hilft das systematische Absuchen der Umgebung. Die systematische Suche wird dabei immer erst dann ausgeführt, wenn andere Strategien keine weiteren Orientierungsinformationen mehr liefern können. Damit stellt sie ein wichtiges Sicherungssystem dar, das unabhängig von den vorhandenen Strukturen der Umgebung ausgeführt werden kann. Der Verlauf einer Suche kann dabei einer Folge von immer größer werdenden Spiralen ähneln. Jede Spirale wird dabei zwischendurch unterbrochen und das Individuum kehrt zum Startpunkt der Suche zurück. Von Zeit zu Zeit ändert sich der Drehsinn der folgenden Spirale (s. Abb. 2b; Hoffmann (1983); Müller and Wehner (1994)). Neben der spiralförmigen Suche wird in begrenzten Umgebungen beobachtet, dass während einer Suche den Wänden beziehungsweise Begrenzungen gefolgt wird (Horev et al. (2007); Yaski et al. (2009)).

Ein Gebiet kann auch durch Mäandern, ähnlich dem Prinzip des Rasenmähens, sehr effizient abgesucht werden. Diese Strategie wird beispielsweise bei der Suche nach Lawinenverschütteten angewendet (Munter (2003) und s. Abb. 2b).

Navigation mit Landmarken

Die Nutzung von Landmarken zur Orientierung erweitert das Navigationsvermögen. Allerdings ist es schwer, den Begriff der Landmarke in seiner Bedeutung scharf einzugrenzen. Im Allgemeinen kann eine Landmarke als ein (z.B. visuelles oder akustisches) Merkmal in der Umgebung betrachtet werden, das der Orientierung hilft.

Dabei können Landmarken verschiedene sensorische Modalitäten haben. So kann der Duft eines Baumes oder der holprige Kopfsteinbelag eines Platzes genauso als Landmarke dienen wie das Rauschen eines Wasserfalls oder der Anblick eines Gebäudes. Nicht nur aus Sicht des Menschen nehmen visuelle Landmarken eine besondere Bedeutung bei der Orientierung ein. Sie sind in der Regel örtlich sehr stabil und vor allem oft bereits aus größerer Entfernung wahrnehmbar. Die Nutzung von mehreren Landmarken ermöglicht es, auf Veränderungen in der Umgebung flexibler zu reagieren, da das Fehlen einer Landmarke durch die anderen Landmarken kompensiert werden kann. Auch können Veränderungen der Landmarken gelernt werden, was im Vergleich zu einfachen Orientierungsmechanismen, wie der Taxis und der Wegintegration, zu einem anpassungsfähigen und komplexeren Navigationsverhalten führt.

Zielführung

Im einfachsten Fall bewegt sich ein Individuum zur Zielfindung auf eine einzelne, weithin sichtbare Landmarke zu, wie zum Beispiel einen Berg oder Leuchtturm. Diese einfache Zielführung mit Hilfe eines einzelnen Zieles wird in der Literatur auch mit dem Begriff *aiming* beschrieben (Franz and Mallot (2000), s. Abb. 3a).

Ein Ziel, das nicht direkt durch eine einzelne Landmarke markiert ist, kann aber

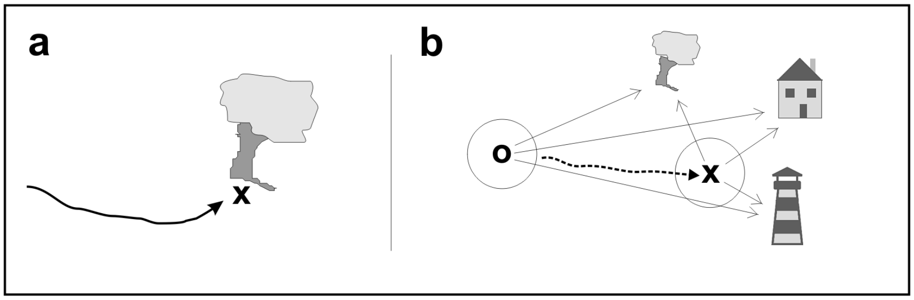


Abbildung 3: (**a**) (*aiming*): Einfache Zielführung mit Hilfe eines einzelnen Zieles (**b**) (*guidance*): Ein Zielpunkt kann durch die Konstellation mehrerer Landmarken zueinander definiert sein. Um das Ziel wiederzufinden, wird zu jedem Zeitpunkt die Bewegungsrichtung (gestrichelte Linie) ausgewählt, in der die Ähnlichkeit des aktuellen Bildes (Position **o**) verglichen mit der (zuvor gespeicherten) Bildinformation am Zielort (**x**) maximiert wird.

durch die Konstellation mehrerer Landmarken zueinander definiert sein (O'Keefe and Nadel (1978)). Dazu muss am Zielort zuvor die Position der umliegenden Landmarken gelernt werden. Ein verbreitetes Modell zur Speicherung der visuelle Information schlägt vor, diese in einem sogenannten Schnappschuss zu speichern (Cartwright and Collett (1982)). Zum Wiederfinden des Zielortes bewegt sich ein Individuum nun so, dass die aktuelle visuelle Information mit dem gespeicherten Schnappschuss in Übereinstimmung kommt. Dazu wird kontinuierlich der Schnappschuss des Zielortes mit dem gegenwärtigen Bild verglichen und die Richtung eingeschlagen, in der sich der Unterschied zwischen den beiden Bildern minimiert. Dieser Mechanismus wird von Trullier et al. (1997) *guidance* genannt. Basiert das Verfahren rein auf Bildinformationen, wird oft auch der Begriff *visual homing* (visuelles Heimfinden) verwendet (s. Abb. 3b).

In einem Experiment, das sich in der Raumkognition inzwischen als Standardparadigma etabliert hat, konnte ein solcher Mechanismus bei Ratten gezeigt werden. Dazu wurde ein Becken mit einer trüben Flüssigkeit gefüllt und an einer Stelle knapp unter der Wasseroberfläche eine Plattform positioniert, ein sogn. „*Morris water maze*". Wird eine Ratte an einer beliebigen Stelle ins Wasser gesetzt, lernt diese nach kurzer Zeit die nicht sichtbare Plattform anzuschwimmen, um sich darauf zu stellen. Dabei orientiert sich das Tier an sichtbaren Landmarken ausserhalb oder an den Wänden des Beckens und bestimmt die Eigenposition relativ zu diesen Landmarken (Morris (1981)). Werden die Landmarken um das Becken rotiert, sucht nun die Ratte die Plattform ensprechend an der Position, von der aus der aktuelle Blick wieder mit der gelernten Perspektive übereinstimmt.

Auch bei Bienen und Ameisen konnte dieser Mechanismus beobachtet werden. Werden in einem Versuch zuvor gelernte Landmarken verstellt, verschieben sich auch die Suchmuster der Tiere hin zu dem Ort, an dem nun die größte Übereinstimmung von aktueller und gelernter Ansicht auftritt (Cartwright and Collett (1982); Wehner and Räber (1979)).

Die einfache Zielführung zu einem einzelnen Ort wird erweitert, wenn an dem erreichten Zielort zusätzlich eine Bewegungsanweisung gespeichert wird, die zu einem weiteren Ort führt. Beim Erreichen des ersten Ortes (*place*) wird dann das Ausführen der Bewegungsanweisung (*response*) ausgelöst, wie beispielsweise 'am großen Baum links weiter'. Dieser Reiz-Reaktions-Mechanismus (*place recognition-triggered response*) besteht aus drei Elementen: 1) Erkennen des Ortes, 2) sich an diesem Ort orientieren bzw. ausrichten und 3) die entsprechende Bewegungsrichtung für den nächsten Zielort auswählen (Trullier et al. (1997)).

Durch die Verkettung mehrerer solcher Zielorte kann eine längere Strecke überwunden werden. Allerdings ist diese Strategie unflexibel, insofern keine Repräsentation über die Umgebung vorliegt, sondern nur einer Kette von Reiz-Reaktionen gefolgt wird. Das Individuum hat bei dem Reiz-Reaktionsmechanismus keine Kenntnis darüber, wo es sich genau befindet und welche Lage die einzelnen Orte zueinander haben. Wenn also ein Ort entlang dieser Kette nicht erkannt werden kann, dann ist damit die Route unterbrochen und das Ziel kann nicht mehr ohne Weiteres erreicht werden.

Routen und kognitive Karten

Mehrere Orte (*places*) können durch jeweils assoziierte Bewegungsanweisungen (*responses*) miteinander verbunden werden. Die Verknüpfen mehreren solcher *place recognition-triggered responses* bildet somit eine lineare Route (Kuipers (1978)).

Werden nun zwei oder mehrere Routen, die sich kreuzen, miteinander verbunden entsteht ein Netzwerk, d.h. eine kognitive Karte der Umgebung. Diese Karte kann intuitive räumliche Beziehungen wie Nachbarschaften und relative Positionen von Orten zueinander beschreiben (topologische Karte). Alternativ dazu kann ein Netzwerk von Orten zusätzlich mit genauen Positionsinformationen versehen sein und Angaben zu Entfernungen zwischen einzelnen Orten enthalten (metrische Karte). Im Folgenden werden die Konzepte zu Routen und kognitiven Karten im Einzelnen vorgestellt.

Routennavigation

Die Routennavigation stellt eine Erweiterung der einfachen Orientierungsmechanismen dar. Durch die Verknüpfung verschiedener Orte mittels jeweils dazu assoziierter Bewegungsanweisung lassen sich einzelne Orte zu einer Route verbinden. Eine Route kann also (wie oben beschrieben) als Kette von Reiz-Reaktionen betrachtet werden (Siegel and White (1975)).

Die zu einem Ort gespeicherte Bewegungsanweisung kann beispielsweise in folgender Form vorliegen: "Von hier aus Richtung Norden weiterlaufen", oder: "Von hier aus rechts weiterlaufen". Diese Beispiele für Bewegungsanweisungen sollen zeigen, dass sich eine Bewegungsanweisung an unterschiedlichen Referenzsystemen ausrichten kann. Im ersten Fall handelt es sich um einen globalen (allozentrischen) Referenzrahmen ("nach Norden"). Im zweiten Fall bezieht sich die Anweisung auf einen körperbezogenen, d. h. egozentrischen Referenzrahmen (“rechts“).

Der globale Referenzrahmen hat den Vorteil, dass unabhängig von der Richtung, aus der ein Ort erreicht wird, die weitere Anweisung immer eindeutig ist. Aus der egozentrischen Perspektive spielt die Annäherungsrichtung eine entscheidende Rolle, da die Anweisung “nach rechts“ in unterschiedliche Richtungen weiterführt, je nachdem, aus welcher Richtung ein Ort erreicht wurde. Allerdings wird dafür kein weiteres Wissen über die Umgebung und deren globale Ausrichtung vorausgesetzt.

Die meisten Navigationsleistungen von Tieren und Menschen lassen sich am besten durch die Kombination einer egozentrischen Repräsentation eines Ortes mit grobem geometrischem (allozentrischem) Umgebungswissen erklären (Wang and Spelke (2002)).

Kognitive Karte

Der Begriff der kognitiven Karte wurde von Tolman (1948) geprägt und beschreibt die mentale Repräsentation räumlichen Umgebungswissens.

Einzelne Orte und Routen bilden durch Verknüpfung ein Netzwerk, d.h. eine kognitive Karte. An einen Ort können dabei mehrere Bewegungsanweisungen assoziiert sein. Durch das Vorhandensein einer Auswahl an Bewegungsanweisungen wird eine flexible Navigation ermöglicht (Siegel and White (1975); Wang and Spelke (2002)). Zudem kann die Lage einzelner Orte nicht nur innerhalb einer Route, sondern innerhalb der (benachbarten) Umgebung bekannt sein. Mittels einer solchen Karte können zwischen verschiedenen Orten nicht nur bekannte Routen genutzt werden, sondern auch völlig neue Routen generiert werden (O’Keefe and Nadel (1978)). Die Inhalte und die Form einer kognitiven Karte sind aber nicht zu vergleichen mit einer ’gewöhnlichen’ Landkarte, sondern die Karte ist unterschiedlich detailiert, verzerrt und lückenhaft (Tversky (1981)). Die Eigenschaften und die Nutzung einer kognitiven Karte wird, wie im Weiteren dargestellt, von Trullier et al. (1997) unterteilt in die topologische Karte bzw. Navigation und die metrische Karte bzw. Navigation.

Topologische Navigation. Eine topologische Karte entsteht durch das Zusammenführen einzelner Routen und Orte. Dabei können einzelne Orte ursprünglich in verschiedene Routen integriert gewesen sein. Diese Orte bilden damit Punkte, an denen Routen miteinander verknüpft werden können (siehe Abb. 4b, Ort ’C’). Diesen Orten sind dann mehrere zielabhängige Bewegungsanweisungen assoziiert. Die topologische Navigation erlaubt nun an einem Ort mehrere Handlungsoptionen. Somit können neue Routen generiert werden, wobei jeder Ort in der topologischen Karte als Start- oder Zielpunkt dienen kann. Dennoch sind auf Grund der fehlenden Metrik keine Abkürzungen oder neue Verbindungen zwischen bisher nicht direkt verbundenen Orten möglich (Trullier et al. (1997)).

Metrische Navigation. Durch die topologische Vernetzung ist es möglich, aus bekannten Routen neue Routen zu kombinieren. Allerdings sind keine absoluten Positionen der Orte bekannt. Eine metrische Einbettung enthält Informationen zu Di-

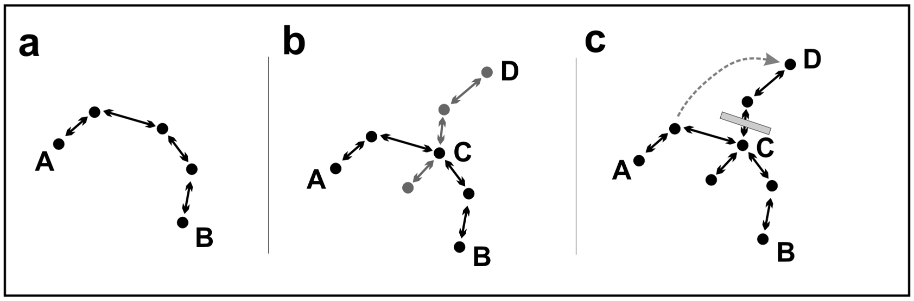


Abbildung 4: (**a**) Einfache Route durch lineare Verknüpfung von Orten. Die Verbindung zwischen zwei Orten kann dabei nur in eine Richtung verlaufen oder bidirektional sein. (**b**) Die Kombination mehrerer Routen erlaubt eine topologische Navigation (s. Text). (**c**) In einer metrischen Repräsentation der Umgebung kann auf Hindernisse (grauer Balken) flexibel reagiert werden und es können neue Wege geplant werden.

stanzen und Richtungen zwischen den Orten (Trullier et al. (1997)). Dies erleichtert es, auf Veränderungen in der Umgebung zu reagieren und Abkürzungen sowie völlig neue Wege zu nutzen (s. Abb. 4c). Die Metrik wird durch Integration verschiedener Informationen aus Wegintegration und Landmarkenwissen in die kognitive Karte eingefügt (Gallistel and Cramer (1996)). Dabei ist die metrische Information nicht immer exakt, sondern kann auch fehlerbehaftet sein. Die metrische kognitive Karte kann daher auch räumliche Verzerrungen aufweisen. Bei der Wahl von zwei Wegen kann dadurch auch der vermeintlich kürzere Weg gewählt werden, obwohl dieser in Wirklichkeit deutlich länger ist (vgl. Stevens and Coupe (1978); Tversky (1981)).

Kommunikation und Raumorientierung

Die Kommunikation mittels visueller Signale oder akustischer Laute ist im Tierreich weit verbreitet. Allerdings gibt es kaum Hinweise darauf, dass bei der Kommunikation Informationen über räumliche Eigenschaften der Umwelt oder Raumbeschreibungen vermittelt werden. Komplexe und flexible Raumbeschreibungen, die sprachlich oder durch Abbildungen weitergegeben werden, scheinen eine explizit menschliche Fähigkeit zu sein. Das Übermitteln von Rauminformation zählt damit zur komplexesten Stufe der beschriebenene Navigations- und Orientierungsstrategien. Im Folgenden wird dies näher erläutert.

Sprache

Die Fähigkeit, über räumliche Strukturen zu kommunizieren, findet sich im Tierreich kaum. Eine Ausnahme ist der sogenannte Schwänzeltanz bei Honigbienen (*Apis mellifera*), entdeckt von Karl von Frisch (1965). Der Tanz ist durch zwei Phasen charakterisiert: (1) einen geradlinigen Abschnitt, bei dem die Biene mit dem Hinterleib wackelt und (2) einem gebogenen Abschnitt ohne Hinterleibsbewegung. Der zweite Teil des Tanzes schließt an den geradlinigen Teil wechselweise links oder rechs herum

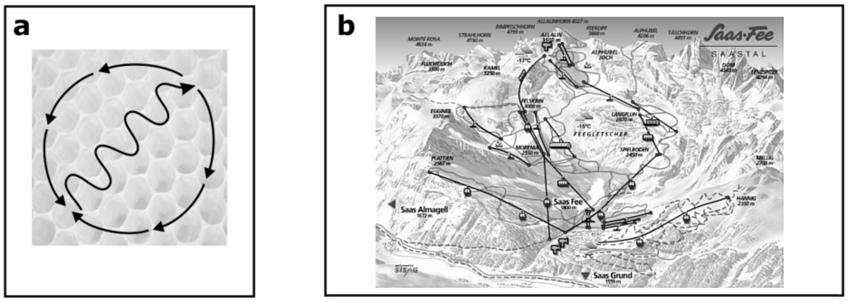


Abbildung 5: (**a**) Verlauf des Bienentanzes zur Rekrutierung anderer Bienen; (**b**) 3D Bergkarte eines Wintersportgebietes mit Zusatzinformationen wie Pisten, Lifte und Beschneiungsanlagen (Abb. mit freundlicher Genehmigung von Saas-Fee/Saastal Tourismus, www.saas-fee.ch)

an und führt in einem Bogen zum Ausgangspunkt des Tanzes zurück. Diese Bewegung wird mehrfach wiederholt (Abb. 5 a). Der Tanz wird im Bienenstock auf den vertikal hängenden Waben ausgeführt. Der Winkel des geradlinigen Stücks relativ zur Schwerkraft korreliert hierbei mit der Richtung der Nahrungsquelle relativ zum Horizontalwinkel (Azimuth) der Sonne. Ein Schwänzeltanz mit dem geraden Stück senkrecht nach oben laufend bedeutet, dass die Futterquelle in Richtung Sonne zu finden ist, ein Tanz auf der Wabe nach unten entspricht der entgegengesetzten Richtung. Entsprechendes gilt für Abweichungen nach links und rechts. Die Distanz zur Nahrungsquelle wird dabei durch die Dauer der ausgeführten Tanzabschnitte angezeigt (Marco and Menzel (2005)).

Eine andere Form von Kommunikation über räumliche Information konnte bisher im Tierreich nicht gefunden werden. Der Mensch hingegen ist fähig, komplizierte räumliche Umgebungen zu beschreiben (Übersicht siehe Levinson (2003)). Dabei kann die Rauminformation mündlich weitergegeben werden oder auch in schriftlicher Textform vorliegen.

Landkarten und Pläne

Straßen- und Landkarten bieten dem Menschen zusätzlich zur Sprache ein Mittel, räumliche Zusammenhänge zu externalisieren und anderen Personen zugänglich zu machen.

Dadurch ist eine Person in der Lage, sich ohne weitere Hilfe in einer ihr völlig unbekannten Umgebung zu orientieren. Neben der Topographie einer Umgebung kann eine Karte noch eine Vielzahl an Zusatzinformationen enthalten, wie z.B. thermographische Daten oder Informationen über die Vegetation. Der Informationsgehalt, der in Karten gespeichert werden kann, ist somit nahezu endlos und an die Bedürfnisse der Benutzer anpassbar (Abb. 5, b).

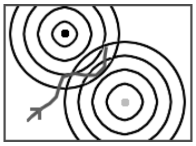
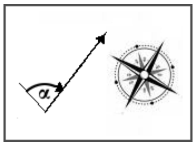
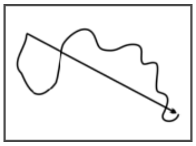
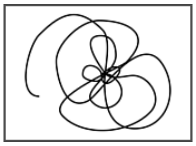
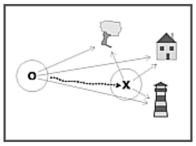
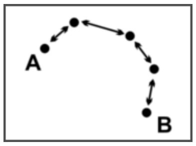
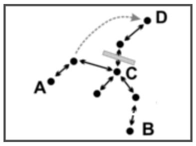
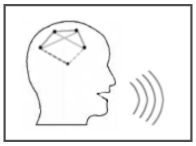
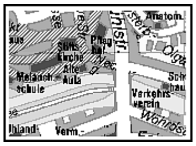
Lokale Navigationsstrategien		
Taxis	Folgen eines Reizgradienten z.B. Phototaxis, Chemotaxis,...	
Richtungsfolgen	Ausrichten entlang eines Kurses (z.B. Magnetkompass)	
Wegintegration	egozentrische Integration der Eigenbewegung	
Suche	Finden bzw. Wiederfinden eines (bekannten) Ortes	
Zielführung	Nutzen einer LM (*aiming*) o. der Konstellation mehrerer LM zueinander (*guidance*)	
Globale Navigationsstrategien		
place recognition-triggered response	lineare Verknüpfung einzelner Orte zu Routen	
Kognitive Karte	Zusammenführung von Routen zu einer topologischen oder metrischen Karte	
Raum und Kommunikation		
Kommunikation über Raum	Weg- und Raumbeschreibungen, aber auch Bienentanz	
Straßen-/Landkarten	externalisiertes Umgebungswissen	

Tabelle 2: Die Navigationshierarchie in der Übersicht.

Hierarchie der Navigationsstrategien

Zur Strukturierung der Navigationsstrategien teilten Trullier et al. (1997) die Strategien in eine Hierarchie mit zunehmenden Komplexitätsstufen ein. Die Hierarchie umfasst vier Stufen: i) Zielführung (*guidance*), ii) ein Reiz-Reaktions-Mechanismus (*place recognition-triggered response*), d.h. einfache Verknüpfung von Plätzen zu Routen, iii) topologische Navigation und iv) metrische Navigation.
Eine andere und erweiterte Form einer Navigationshierarchie haben Franz und Mallot (2000) vorgeschlagen. Unterschieden wurden hierbei lokale Strategien und Strategien für die Überwindung grösserer Distanzen. Für die lokale Navigation wurden vier Strategien definiert: (i) Richtungsfolgen (dies umfasste auch die Wegintegration), (ii) *aiming*, (iii) *guidance* und (iv) die Suche. Während die lokalen Strategien ein Individuum nur zu einem einzigen Ziel führen, erlauben die folgenden drei Strategien das Finden von Wegen in größeren Umgebungen: (i) *place recognition-triggered response*, (ii) die topologische und (iii) die metrische Navigation (von Franz and Mallot (2000) mit Übersichtsnavigation bezeichnet).

Am Ende dieser Einleitung möchte ich die vorgestellten Navigationshierarchien noch durch den Bereich der Kommunikation erweitern, d. h. der Verwendung von Sprache (z.B. Wegbeschreibungen) und Kartenabbildungen (z.B. Straßenkarten). Zusätzlich wird auch die Taxis mit aufgenommen, die, wie das Beispiel der Pheromonspuren bei Ameisen zeigt, ein wichtiger Mechanismus zur Unterstützung der Raumorientierung sein kann. Die Zusammenfassung der vorgestellten Strategien zeigt abschließend die erweiterte Navigationshierarchie in Tabelle 2.

Wissenschaftliche Fragestellung

Die einzelnen Stufen der Navigationshierarchie (Tab. 2) beschreiben die grundlegenden Strategien aus der Literatur zur Raumkognition. Innerhalb einer Stufe können dabei unterschiedliche Mechanismen vorliegen. Beispielsweise kann bei der Wegintegration sowohl der optische Fluss, als auch die Propriozeption zur Distanzbestimmung genutzt werden. Wie und ob einzelne Tiere und Menschen eine dieser Strategien nutzen (können), ist daher vorwiegend abhängig von der jeweils zur Verfügung stehenden Sensorik und den Verarbeitungsmöglichkeiten.

Der Fokus der vorliegenden Arbeit richtet sich auf die Erkennung und die kognitive Repräsentation von Orten im Gedächtnis.
Welche Merkmale spielen eine Rolle bei der Ortserkennung? Wie wird ein Ort in eine Route oder kognitive Karte integriert und repräsentiert? Und welche Bedeutung hat die Repräsentation des Ortsgedächtnisses beim Abruf aus dem Langzeitgedächtnis? Die Arbeit folgt bei der Beantwortung dieser Fragen der steigenden Komplexität der Navigationshierarchie.
Im ersten Schritt wird untersucht, wie mittels eines einfachen Systems (d.h. ein Modell einer Wüstenameise) Orte in einer natürlichen Umgebung erkannt und voneinander unterschieden werden können. Zwar liegen bei der Wüstenameise zahlreiche Verhaltensexperimente vor, die ihre Fähigkeiten zur Ortserkennung charakterisieren, dennoch ist noch nicht eindeutig geklärt, welche Umgebungsmerkmale dazu von den Tieren genutzt werden.
Aufbauend auf die einfache Ortserkennung wurde in einem Verhaltensexperiment mit Menschen untersucht, wie einzelne Orte in Routen- und Kartenwissen eingebunden werden. Werden dabei Orte erst zu Routen verknüpft bevor diese dann nach und nach zu einer kognitiven Karte vernetzt werden, oder laufen beide Prozesse parallel, d.h. besteht eine gleichzeitige Integration von Ortswissen in Routen- und Kartenwissen?
Um räumliches Wissen nutzen zu können, muss dieses in einer sinnvollen Struktur vorliegen. Welche Verknüpfungen zwischen verschiedenen Orten werden gespeichert und wie werden sie in einer kognitiven Karte repräsentiert? Wie kann wichtige Zusatzinformation (wie z.B. 'hier gibt es Nahrung') in die Repräsentation eingebunden werden?
Zu diesen Fragen wird in dieser Arbeit ein Konzept zur Repräsentation räumlichen Wissens vorgestellt, in dem nicht nur ökologische Anforderungen und neuronale Mechanismen diskutiert werden, sondern auch eine mögliche Implementierung in künstlichen Systemen berücksichtigt wurde.
Im letzten Teil der Arbeit wird der Frage nachgegangen, in welcher Form Ortswissen beim Abruf aus dem Langzeitgedächtnis bereit gestellt wird. Im Gegensatz zu vielen Laborexperimenten, in denen Räume oder einzelne Orte oft in kurzer Zeit gelernt werden müssen, wurde hier Ortswissen von stark vertrauten, natürlichen Orten abgefragt.

Zusammengefasst gibt die vorliegende Arbeit neue Einblicke in die Raumkognition. Den Schwerpunkt der Arbeit bilden der Erwerb, die Verarbeitung, die Speicherung und der Abruf von Ortswissen. In den folgenden Kapiteln werden dazu fünf Artikel mit qualitativen, theoretischen und empirischen Projekten zum Themenschwerpunkt Ortswissen vorgestellt.

Literaturverzeichnis

Aczel, A. D. (2005). *Der Kompass*. Rowohlt.

Bayer, K. (1993). *Gaius Plinius Secundus d.Ä.: Naturkunde (Historia naturalis)*, chapter Periplus Hannonis, pages 346–353. V. Artemis, Zürich, München.

Braitenberg, V. (1984). *Vehicles: Experiments in Synthetic Psycology*. MA: MIT Press, Cambridge.

Bremmer, F. and Lappe, M. (1999). The use of optical velocities for distance discrimination and reproduction during visually simulated self motion. *Experimental Brain Research*, 127:33–42. Exp Brain Res.

Bures, J., Fenton, A. A., Kaminsky, Y., Wesierska, M., and Zahalka, A. (1998). Rodent navigation after dissociation of the allocentric and idiothetic representations of space. *Neuropharmacology*, 37:689–699.

Cartwright, B. A. and Collett, T. S. (1982). How honey bees use landmarks to guide their return to a food source. *Nature*, 295:560–564.

Cattet, J. and Etienne, A. S. (2004). Blindfolded dogs relocate a target through path integration. *Animal Behaviour*, 68:203–212.

Collett, T. S. and Collett, M. (2000). Path integration in insects. *Current Opinion in Neurobiology*, 10:757–762.

Durgin, F. H., Akagi, M., Gallistel, C. R., and Haiken, W. (2009). The precision of locomotor odometry in humans. *Experimental Brain Research*, 193(3):429–436.

Durier, V. and Rivault, C. (1999). Path integration in cockroach larvae, *Blattella germanica* (l.) (insect: Dictyoptera): Direction and distance estimation. *Animal Learning & Behavior*, 27(1):108–118.

Esch, H. E. and Burns, J. (1995). Honeybees use optic flow to measure the distance of a food source. *Naturwissenschaften*, 82(1):38–40.

Etienne, A. S. and Jeffery, K. J. (2004). Path integration in mammals. *Hippocampus*, 4:180–192.

Etienne, A. S., Maurer, R., and Séguinot, V. (1996). Path integration in mammals and its interaction with visual landmarks. *Journal of Experimental Biology*, 199:201–209.

Franz, M. O. and Mallot, H. A. (2000). Biomimetic robot navigation. *Robotics and Autonomous Systems*, 38:133–153.

Gallistel, C. R. and Cramer, A. E. (1996). Computations on metric maps in mammals: getting oriented and choosing a multi-destination route. *Journal of Experimental Biology*, 199:211–217.

Hoffmann, G. (1983). The search behavior of the desert isopod *Hemilepistus reaumuri* as compared with a systematic search. *Behavioral Ecology and Sociobiology*, 13:93–106.

Homer (ca. 8.Jhd.v.Chr./ Ausgabe 1995). *Odyssee*. Reclam. Übersetzung: Roland Hampe, S. 82.

Horev, G., Benjamini, Y., Sakov, A., and Golani, I. (2007). Estimating wall guidance and attraction in mouse free locomotor behavior. *Genes, Brain & Behavior*, 6(1):30–41.

Jackson, D. E. and Chaline, N. (2007). Modulation of pheromone trail strength with food quality in Pharaoh's ant, *Monomorium pharaonis*. *Animal Behaviour*, 74(3):463–470.

Jackson, D. E., Holcombe, M., and Ratnieks, F. L. W. (2004). Trail geometry gives polarity to ant foraging networks. *Nature*, 432:907–909.

Johnson, D. S. and Nurminen, J. (2008). *Die große Geschichte der Seefahrt*. National Geographic.

Kearns, M., Warren, W. H., Duchon, A. P., and Tarr, M. J. (2002). Path integration from optic flow and body senses in a homing task. *Perception*, 31:349–374.

Klatzky, R. L., Beal, A. C., Loomis, J. M., Golledge, R. G., and Philbeck, J. W. (1999). Human navigation ability: Tests of the encoding-error model of path integration. *Spatial Cognition and Computation*, 1:31–65.

Kuipers, B. (1978). Modeling spatial knowledge. *Cognitive Science*, 2:129–153.

Levinson, S. C. (2003). *Space in Language and Cognition*. Cambridge University Press.

Levitt, T. S. and Lawton, D. T. (1990). Qualitative navigation for mobile robots. *Artificial Intelligence*, 44(3):305–360.

Loomis, J. M., Klatzky, R. L., and Golledge, R. G. (2001). Navigating without vision: Basic and applied research. *Optometry and Vision Science*, 78(5):282–289.

Loomis, J. M., Klatzky, R. L., Golledge, R. G., Cicinelli, J. G., Pellegrino, J. W., and Fry, P. A. (1993). Nonvisual navigation by blind and sighted: assessment of path integration ability. *Journal of Experimental Psychology: General*, 122:73–91.

Marco, R. D. and Menzel, R. (2005). Encoding spatial information in the waggle dance. *Journal of Experimental Biology*, 208:3885–3894.

May, M. and Klatzky, R. L. (2000). Path integration while ignoring irrelevant movement. *Learning, Memory, and Cognition*, 26(1):169–186.

Mittelstaedt, M.-L. and Mittelstaedt, H. (1980). Homing by path integration in a mammal. *Naturwissenschaften*, 67(11):566–567.

Moller, P. and Görner, P. (1994). Homing by path integration in the spider *Agelena labyrinthica* clerck. *Journal of Comparative Physiology A*, 174(2):221–229.

Morris, R. G. M. (1981). Spatial localization does not require the presence of local cues. *Learning and Motivation*, 12:239–260.

Müller, M. and Wehner, R. (1994). The hidden spiral: systematic search and path integration in desert ants, *Cataglyphis fortis*. *Journal of Comperative Physiologie A*, 175:525–530.

Munter, W. (2003). *Drei mal drei (3x3) Lawinen. Risikomanagement im Wintersport*. 4. Aufl. Bergverlag Rother.

Murlis, J., Elkinton, J. S., and Cardé, R. T. (1992). Odor plumes and how insects use them. *Annual Review of Entomology*, 37:505–532.

Nørgaard, T., Henschel, J. R., and Wehner, R. (2003). Long-distance navigation in the wandering desert spider *Leucorchestris arenicola*: can the slope of the dune surface provide a compass cue? *Journal of Comperative Physiology A*, 189(11):801–809.

O'Keefe, J. and Nadel, L. (1978). *The Hippocampus as a cognitive map*. Clarendon Press-Oxford.

Paola, V. D., Marijuán, P. C., and Lahoz-Beltra, R. (2004). Learning and evolution in bacterial taxis: an operational amplifier circuit modeling the computational dynamics of the prokaryotic 'two component system' protein network. *Biosystems*, 74:29–49.

Péruch, P., Borel, L., Magnan, J., and Lacour, M. (2005). Direction and distance deficits in path integration after unilateral vestibular loss depend on task complexity. *Cognitive Brain Research*, 25(3):862–872.

Philbeck, J. W. and O'Leary, S. (2005). Remembered landmarks enhance the precision of path integration. *Psicologica*, 26:724.

Riecke, B. E., van Veen, H. A. H. C., and Bülthoff, H. H. (2002). Visual homing is possible without landmarks - a path integration study in virtual reality. *Presence: Teleoperators and Virtual Environments*, 11(5):443–473.

Rieser, J. J., Ashmead, D. H., Talor, C. R., and Youngquist, G. A. (1990). Visual perception and the guidance of locomotion without vision to previously seen targets. *Perception*, 19:675–689.

Robinson, E. J. H., Jackson, D. E., Holcombe, M., and Ratnieks, F. L. W. (2005). 'No entry' signal in ant foraging. *Nature*, 438:442.

Seguinot, V. V., Cattet, J., and Benhamou, S. (1998). Path integration in dogs. *Animal Behaviour*, 55(4):787–97.

Siegel, A. W. and White, S. H. (1975). The development of spatial representations of large-scale environments. *Advances in child development and behavior*, 10:9–55.

Spennemann, D. H. R. (2005). Traditional and nineteenth century communication patterns in the marshall islands. *Micronesian Journal of the Humanities and Social Sciences*, 4(1):25–52.

Srinivasan, M. V., Zhang, S., Altwein, M., and Tautz, J. (2000). Honeybee navigation: Nature and calibration of the ' odometer'. *Science*, 287:851–853.

Stevens, A. and Coupe, P. (1978). Distortions in judged spatial relation. *Cognitive Psychology*, 10(4):422–437.

Sumpter, D. J. T. and Beekman, M. (2003). From nonlinearity to optimality: pheromone trail foraging by ants. *Animal Behaviour*, 66:273–280.

Tolman, E. C. (1948). Cognitive maps in rats and men. *Psychological Review*, 55(4):189–208.

Trullier, O., Wiener, S. I., Berthoz, A., and Meyer, J.-A. (1997). Biologically based artificial navigation systems: review and prospects. *Progress in Neurobiology*, 51(5):483–544.

Tversky, B. (1981). Distortions in memory for maps. *Cognitive Psychology*, 13:407– 433.

von Frisch, K. (1965). *Tanzsprache und Orientierung der Bienen (engl. version: The Dance Language and Orientation in Bees, 1967).* Springer, Berlin, Heidelberg, New York.

Wang, R. F. and Spelke, E. S. (2002). Human spatial representation: insights from animals. *Trends in Cognitive Sciences*, 376(9):376–382.

Wehner, R. (1992). *Animal Homing*, chapter 3: Arthropods, pages 45–144. Chapman and Hall.

Wehner, R. (2003). Desert ant navigation: how miniature brains solve complex tasks. *Journal of Comperative Physiology A*, 189:579–588.

Wehner, R. and Räber, F. (1979). Visual spatial memory in desert ants, *Cataglyphis bicolor* (Hymenoptera: Formicidae). *Experientia*, 35:1569–1571.

Wiener, J. M. and Mallot, H. A. (2006). Path complexity does not impair visual path integration. *Spatial Cognition and Computation*, 6(4):333–346.

Wilson, E. O. (1963). Pheromones. *Scientific American*, 208:100–114.

Wiltschko, W. and Wiltschko, R. (2005). Magnetic orientation and magnetoreception in birds and other animals. *Journal of Comperative Physiology A*, 191(8):675–693.

Yaski, O., Portugali, J., and Eilam, D. (2009). The dynamic process of cognitive mapping in the absence of visual cues: human data compared with animal studies. *Journal of Experimental Biology*, 212:2619–2626.

Manuskript 1

Simulated visual homing in desert ant natural environments: efficiency of skyline cues

Ziele, Hauptergebnisse und Eigenanteil an der Publikation/dem Manuskript

Staatenbildende Tiere wie Ameisen sind darauf angewiesen, dass einzelne Tiere nach erfolgreicher Nahrungssuche wieder zum gemeinsamen Nest zurückfinden. Während viele Ameisenarten zur Markierung ihrer Laufwege Pheromonspuren nutzen, die eine Verbindung zwischen Nest und Nahrungsquelle herstellen, sind Wüstenameisen auf Grund der hohen Bodentemperaturen, die solche Spuren schnell verdampfen lassen, auf andere Orientierungsmechanismen angewiesen. In Verhaltensexperimenten im natürlichen Habitat der Ameisen konnte beobachtet werden, dass die Tiere, die in der Regel einzeln auf Nahrungssuche gehen, individuellen stereotypen Routen folgen. Obwohl angenommen wird, dass visuelle Informationen die Ameisen auf ihren Routen leiten, ist unklar, wie und welche visuelle Information gefiltert, gespeichert und zur Orientierung verwendet wird.

Der folgende Artikel beschreibt ein Verfahren zur visuellen Navigation, welches sich an Verhaltensexperimenten und sensorischen Eigenschaften der Wüstenameisen orientiert. Auf Basis von Kartierungen[1] des natürlichen Habitats der Ameisen wurde von mir im Computer eine virtuelle Umgebung modelliert. Die Umgebung zeichnete sich dabei durch viele unregelmäßig verteilte Grasbüschel aus. In dieser Umgebung wurde mit Hilfe des simulierten Sensoriums der Tiere die Fähigkeit zum Wiederfinden bekannter Orte analysiert. Dabei wurden für die Ortserkennung zwei visuelle Merkmale getestet: Aus dem panoramischen Sichtfeld wurden zum einen Bild- bzw. Pixelintensitäten ausgelesen und zum anderen die obere Kontur umliegender Objekte gegen den Himmelshintergrund, d.h. die Skyline, extrahiert. Die Ergebnisse der Computersimulation zeigen, dass die Skyline im Sichtfeld ausreichende Informationen liefert, um Orte in diesen natürlichen Umgebungen wieder zu erkennen und von

---

[1] Die Kartographien des Ameisenhabitats wurden freundlicherweisen von Herrn Prof. Rüdiger Wehner und seiner Arbeitsgruppe in Zürich zur Verfügung gestellt.

anderen Orten der Umgebung zu unterscheiden. Hauptvorteile dieses Verfahrens sind Beleuchtungsinvarianz, geringe Speicheranforderungen und biologische Plausibilität.

Die Implementierung der Computersimulation und die vollständige Auswertung der daraus resultierenden Ergebnisse wurden von mir durchgeführt. Der Artikel wurde zusammen mit Herrn Prof. Hanspeter A. Mallot verfasst.

Das Manuskript wurde zur Veröffentlichung in der Fachzeitschrift *Biological Cybernetics* akzeptiert.

Referenz:
K. Basten, H.A. Mallot (accepted), *Simulated visual homing in desert ant natural environments: efficiency of skyline cues*, Biological Cybernetics

Noname manuscript No.
(will be inserted by the editor)

Simulated visual homing in desert ant natural environments: efficiency of skyline cues

Kai Basten,
Hanspeter A. Mallot

the date of receipt and acceptance should be inserted later

Abstract Desert ants, foraging in cluttered semi-arid environments, are thought to be visually guided along individual, habitual routes. While other navigational mechanisms (e.g., path integration) are well studied, the question how ants extract reliable visual features from a complex visual scene is still largely open. This paper explores the assumption that the upper outline of ground objects formed against the sky, i.e. the skyline, provides sufficient information for visual navigation. We constructed a virtual model of the ant's environment. In the virtual environment panoramic images were recorded and adapted to the resolution of the desert ant's complex eye. From these images either a skyline code or a pixel-based intensity code were extracted. Further, two homing algorithms were implemented, a modified version of the average landmark vector (ALV) model (Lambrinos et al. 2000) and a gradient ascent method. Results show less spatial aliasing for skyline coding and best homing performance for ALV homing based on skyline codes. This supports the assumption of skyline coding in visual homing of desert ants and allows novel approaches to technical outdoor navigation.

Keywords visual homing · place memory · navigation · UV vision · landmark cue · local position information

1 Introduction

Many animals face the problem of returning to previously visited places subserving functions such as nesting, hiding, mating, feeding, or spawning. For approaching and recognising such places, visual landmarks in the vicinity of the goal can be used. The ability to find places from visual cues, also called visual homing, has been demonstrated in many animal groups, but has been most intensively studied in insects (Tinbergen and Kruyt 1938; Wehner and Räber 1979; Cartwright and Collett 1983). Animal performance in place recognition has attracted the attention of theoretical and robot studies, searching for computationally cheap solutions for visual navigation problems (e.g. Anderson 1977; Cartwright and Collett 1983; Franz et al. 1998b; Möller and Vardy 2006; Stürzl and Zeil 2007).The basic idea, which has been derived from homing experiments with honeybees, assumes a remembered view or 'snapshot' as a memory of the place at which the view was taken. The approach to this place is then achieved by moving such as to increase some sort of image similarity measure between the current view and the snapshot.

While the homing algorithms have been subject of intensive research (see below), the question of the visual features used in the image comparison process is still largely open. In this paper, we analyse the information content included in images from the natural environment of the desert ant *Melophorus bagoti*. Predicted performance depends on the homing algorithm used and is best for the combination of skyline cues and average landmark vector homing.

Homing procedures. Möller and Vardy (2006) present a breakdown of homing algorithms based on the particular image comparison strategy used. In 'correspondence methods', corresponding regions are identified in the memorised snapshot and the agent's current view. Corresponding features may include contrast borders or edges (Cartwright and Collett 1983), colour regions (Gourichon et al. 2002), or blocks of pixels, intensities, or gradients matched with optical flow techniques (Vardy and Möller 2005). In contrast, 'holistic models' use the views as a whole. The image difference can be determined for example on the basis of a circular array of sectors occupied by landmarks (Anderson 1977), pixel intensities in a horizontal panoramic view of one pixel height (Franz et al. 1998b) or extended to 2D panoramic images (Möller 2009), or the root mean square pixel differences between the current and a stored view (Zeil et al. 2003).

While most holistic models store complete images, the average landmark vector (ALV) approach presented

---

Inst. Neurobiology, University of Tübingen, Auf der Morgenstelle 28, D-72076 Tübingen, Germany
E-mail: kai.basten@uni-tuebingen.de, hanspeter.mallot@uni-tuebingen.de, phone: +49 7071 2974600, fax: +49 7071 292891

by Lambrinos et al. (2000) reduces the whole visual information to just one two-dimensional vector. The ALV is computed by averaging unit vectors pointing to previously extracted landmark features in the panorama. The difference between the current ALV and the memorized goal ALV guides an agent to the goal position. The ALV approach is reminiscent of O'Keefe's (1991) centroid algorithm, where the agent determines its position based on the centroid of the visible landmark cues and the main axis (the "slope") passing through those cues. Like the ALV algorithm, O'Keefe's method requires an allocentric frame of reference. In addition, it assumes that the distances to landmarks cues are known.

Feature selection. All sketched algorithms require that enough information is present in the images to identify a goal and home back to it. With respect to the image features used in the homing process, the following requirements may be formulated: (i) *Relevance:* The rate of change of the feature configuration upon observer movements should occur at a medium pace. If change is too quick (as in densely cluttered environments) no initial similarity between the current view and the snapshot can be detected and catchment areas will be negligible. If change is too slow (as in large voids), the goal cannot be pinned down based on image change, leading to large catchment areas but low homing precision. (ii) *Uniqueness:* Feature configurations should be unique. If similar configurations occur at different places (as might be the case in regular environments), the agent may be trapped in false minima ("visual aliasing"). (iii) *Reliability:* Feature configurations should be stable over time. Temporal change may be due to illumination changes (particularly in outdoor environments, see Zeil et al. 2003), in which case robustness may be increased by choosing as features the distances to landmark objects (Stürzl and Zeil 2007). Displacements of landmark objects may also play a role if urban environments (parked cars) or seasonal changes (deciduous forests) are considered. For a more general discussion of criteria for landmark selection, see Gillner et al. (2008).

Homing performance in realistic environments. In order to judge the advantages and shortcomings of different feature types and homing algorithms, behavioural tests in realistic environments are needed. So far, many computational approaches developed for visual homing were successfully tested in laboratory environments without deep depth structure (Franz et al. 1998b; Lehrer and Bianco 2000; Gourichon et al. 2002; Vardy and Möller 2005; Möller and Vardy 2006) or with agents operating in featureless environments with few prominent landmarks (Lambrinos et al. 2000; Möller 2000; Smith et al. 2007). One exception is the work of Zeil et al. (2003) and Stürzl and Zeil (2007) who recorded space-dependend image manifolds for real outdoor environments using a 3D camera-gantry.

Plan of the paper. In this paper, we address the problem of feature selection by simulating the performance of four algorithms using environmental and behavioural data from the ant species *Melophorus bagoti*. This analysis consists of three parts, (i) a computer graphics model of the ant's natural environment and peripheral visual system, (ii) a feature extraction and coding scheme and (iii) a path planning module deriving a homing direction from stored place code and current view. For the first part, we implemented a computer model of a real world environment mapped in the course of behavioural experiments that were carried out in central Australia by Kohler & Wehner (2005). In this model environment, realistic views or snapshots can be generated at arbitrary positions and used in the simulation of ant homing behaviour. Concerning the coding part of visual homing, i.e. the type of snapshot used to encode the home location, we assume that the ants are able to detect the skyline that horizontally separates ground objects from sky background in the view, and that they use this information in homing. The skyline-approach is compared with another insect-inspired algorithm based on image intensity values (Franz et al. 1998b) and various fields of view. Both coding schemes were combined with each of two path planning mechanisms based on the correspondence approach. The first is a gradient ascent method (Zeil et al. 2003) based on an uncentred correlation function; the second is an extended version of the ALV algorithm (Lambrinos et al. 2000) adapted to continuous snapshot data. Both correspondence methods were applied to both snapshot models containing the skyline information or pixel intensities.

2 Desert ant behaviour

This project was inspired by route learning abilities of desert ants of the species *Melophorus bagoti*. These ants are solitary foragers searching on the semiarid desert floor of Central Australia for insect prey that have died of heat exhaustion. The environment is characterised by distributed clusters of grass tussocks. If trained to shuttle back and forth between their nest and a feeder, each individual ant establishes habitual outbound and inbound routes, which guide it through its environment (Kohler and Wehner 2005). Additionally to the ants' ability to keep track of the direction and distance of their nest by path integration, they are thought to be visually guided along their habitual routes. This assumption is supported by the behaviour of ants picked up close to the nest and released on an arbitrary point along the known shuttle-route. In this case, the ants had no positional information from their path integrator, but were still able to pick up their habitual routes. How ants can do this, despite the fact that many tussocks are small and similar looking, is still unclear (Cheng et al. 2009).

The idea that desert ants can use the elevation of surrounding objects in the visual field (i.e., the skyline) for view-based navigation was initially suggested by Wehner & Räber (1979) in terms of a horizon detector. Further support for this idea comes from the following results:

i) In an open area, surrounded by trees and shrubs, displaced ants of the species *Melophorus bagoti* are visually guided by the panoramic contour towards the direction of the nest (Graham and Cheng 2009b). When upper parts of the visual field were obscured, while the majority (about 60%) of the panoramic contour information remained available, ants still headed nestwards. On the contrary, when the lower part of the panorama including most of the panoramic countour information was obscured, ants headed for random directions. Additionally, ants can be guided by an artificial skyline, created using black plastic sheeting, that matches the previously learned natural panorama (Graham and Cheng 2009a).

ii) In outdoor scenes, Saharan desert ants of the genus *Cataglyphis* use the elevation of borders in the visual field to centre between adjacent low walls (Heusser and Wehner 2002). Equivalent behaviour could be observed in indoor experiments with ants of the European species *Leptothorax albipennis* (Pratt et al. 2001).

iii) Wood ants (*Formica japonica*), previously shuttling back and forth between their nest and a feeding site, were displaced from their nest to locations in the vicinity of the feeding side. Consequently they had no positional information from their path integrator. When the ants were deprived of their familiar skyline panorama, they started a systematic search programme instead of direct homing performance as observed in runs without visual deprivation. Hence, ants use the skyline, in this experiment represented by the upper edges of a line of trees, for orientation (Fukushi 2001).

iv) The spectral properties of the visual receptors of the desert ant *Cataglyphis bicolor* show one receptor type with a sensitivity maximum around 500nm (green) and a second one maximally sensitive to UV light around 350nm (Mote and Wehner 1980). Using UV and green light sensors, Möller (2002) could show that two classes, natural objects and sky, can be separated in the UV-green colour space by a simple linear threshold classifier.

v) Neural correlates supporting the hypothesis of skyline extraction were identified in another insect, the fly *Drosophila*. Visual neurons in the fan-shaped body of the fly brain exhibit a memory trace for the parameter ‘elevation’ in the visual pattern memory (Liu et al. 2006).

These findings support the idea that the elevation of objects in the visual field can be extracted and used for navigation. However, they do not answer the reverse question, i.e.: does the elevation of local environmental structures provide enough information for visual localisation and homing in complex natural scenes? We addressed this question in a computer simulation comprising a model of the habitat of the ants, and of the ants’ visual system, as illustrated in the following section.

3 Methods

3.1 Virtual model of the environment

Based on mappings of the ants’ natural environment (about 9 × 12m patch) near Alice Springs (Australia) carried out by Kohler & Wehner (2005), a virtual computer model was constructed (Fig. 1). The model was generated with the realtime 3D content software *Multi-Gen Creator*. It consists of a list of polygons with associated textures. Every mapped grass tussock of the original environment was transfered into a prismatic object of standard height (25cm). All objects were textured with a green grass pattern comprising transparent regions (α-channel=0) between individual grass blades in the upper part. The ground was textured with a blurred sand pattern. The environment was surrounded by a semispherical dome, covered from within with a texture showing a slightly clouded blue sky. The scene was illuminated by ambient light and a directional light source located 20 degree off the zenith in infinite distance. This source generated slight color differences on grass textures attached to polygons with different 3D orientation and thus patches appear to be 3D objects.

3.2 Image rendering

For a given position and heading direction of the agent the 3D rendering toolkit *OpenGL Performer* computes an image. From five virtual cameras covering the four

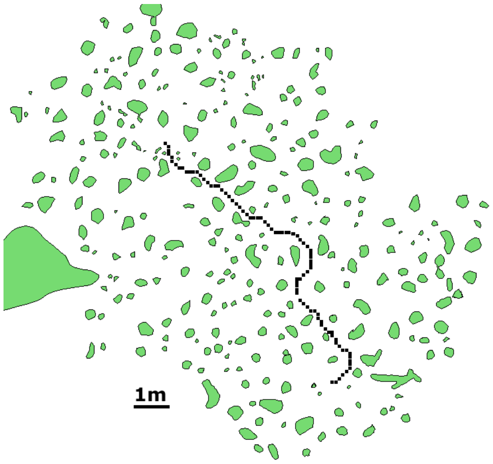


Fig. 1: Map of the modelled environment. Grass patches are indicated by green obstacles. The black markers indicate 86 test locations along a real ant trail used in the results section.

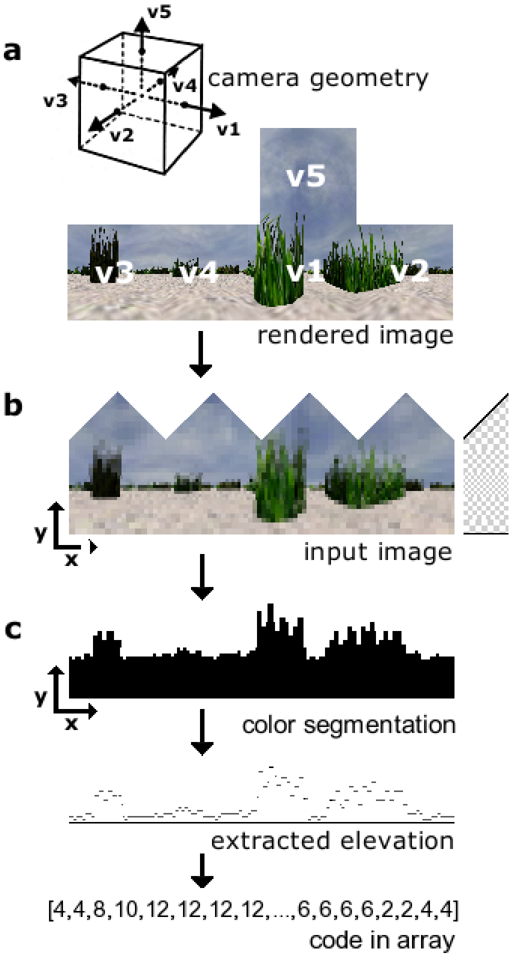


Fig. 2: Image processing. (**a**) In the virtual model, five cameras were defined, which cover all directions in space except the ground. (**b**) The camera images were mapped by unfolding the 'view cube' to one panoramic image. In this step the resolution was adapted to the characteristics of the ant's eye. The inset on the right illustrates the configuration of applied box filters of three different sizes. (**c**) In case of skyline-coding, the skyline was extracted by colour segmentation and the elevation above the horizon was coded in an array of integer values.

horizontal and the upwards direction, one panoramic view was composed (Fig. 2a). The vertical field of view ranged from −45° below to 90° (=zenith) above the horizon. Each camera produced a square image with 60 × 60 pixels; therefore, a central pixel would cover a visual angle of 1.5 × 1.5 degrees. The virtual cameras were located 2 cm above the ground. The top face of the view cube was divided into four isosceles triangles, the bases of which were continuously connected to the according sides of the cube (see Fig. 2b).

The amount of data of the panoramic image was reduced by adapting the resolution of the image to the resolution of a desert ant's complex eye. Since no measurements are available for *Melophorus bagoti*, data from the North African desert ant *Cataglyphis fortis* were used as an approximation (Zollikofer et al. 1995 and cf. Graham and Cheng 2009b). Each complex eye consists of numerous ommatidia, each of which covers about 3 degrees of the visual angle. The angle between the optical axes of the omatidia determines the visual resolution. It is roughly constant along horizontal stripes and varies between 3.0 degrees at the horizon and some 7.0 degrees towards the upper and lower periphery.

Space-variant resolution was modelled by box filters of three sizes, two pixels squared (3°), three pixels (4.5°), and four pixels (6°). The filters were applied at discrete positions, covering the visual field and modelling the viewing directions of the individual ommatidia (see inset on right side of Fig. 2b). Within the square domain of each filter kernel, intensity values were averaged and the average value was writting to each pixel within the box. An example of the resulting image appears in Figure 2b. We will refer to this image as the *input image* for the visual homing procedures. The pixels of the input image will be numbered by $x = 1, ..., x_{max}$ in the horizontal direction ($x_{max} = 240$), and by $y = 1, ..., y_{max}$ in the vertical direction. Due to the cuboid imaging geometry, the maximal possible y-value y_{max} will vary from 60 to 90, depending on x.

3.3 Image Database

For further analysis, 6572 panoramic colour images were recorded in the virtual environment and stored in a reference database. The images were collected at a regular grid with gridsize 0.1 m, excluding the viewpoints within grass tussocks. Thus, the reference database represents the complete view manifold of the environment. All images were taken at a constant elevation above ground (2 cm) with the viewing direction in the horizontal plane and the same global orientation.

3.4 Feature extraction

The visual features encoded from the input image and stored in a snaphot for later retrieval will be referred to as place code. We used and compared two different place codes: (1) the elevation of objects in the visual field, that is, the skyline, and (2) the intensity values of pixels in the image (cf. Franz et al. 1998b). Both place codes require only low-level processing and thus are biologically reasonable.

Skyline

As suggested for insect navigation, in real outdoor environments sky and ground objects can be classified in UV-green colour space, using a linear threshold separator (Möller 2002). In our computer simulation, we used a blue-green colour contrast corresponding to the texturing of sky and tussocks. The separation of foreground and sky pixels obtained from the simulation is shown in (Fig. 3). The skyline elevation η was extracted from the panoramic input image (Fig. 2b) by comparison of the intensities of the blue and green channels according to the following scheme:

$$\eta_x = \sum_{y=y_o}^{y_{max}} h(G_{x,y} - B_{x,y}) \tag{1}$$

where G and B are the green and blue channels of the input image, x numbers the pixels in the horizontal direction and y numbers the pixels in the vertical direction. y_o denotes the y-position of the horizon, which is constant for all images. By h, we denote the Heaviside step function,

$$h(x) := \begin{cases} 1 \text{ if } x \geq 0 \\ 0 \text{ else} \end{cases} . \tag{2}$$

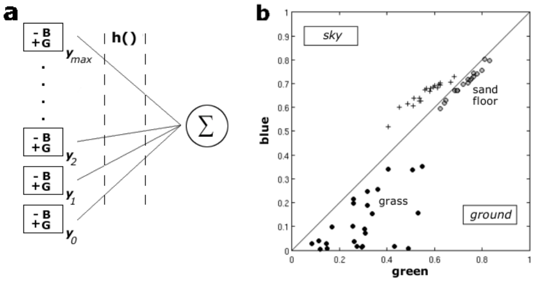


Fig. 3: (**a**) Calculation of elevation η in a column of pixels by counting the number of pixels in which green intensity exceeds blue intensity. (**b**) Labeled object samples obtained from panoramic images of the computer simulation are separated using a simple linear classifier based on the blue (B) and green (G) colour channels.

The sum in Eq. 1 simply counts the number of pixels where green intensity exceeds blue intensity. This proceedure is also illustrated in Figure 3a, together with a classification of simulated sky and ground in Figure 3b. Note that in cases of vertically disconnected objects (such as overhanging branches of a tree) Equ. 1 will count the branch pixels as foreground. This is no problem as long as static environments are considered.

Finally, the input image is encoded as a one-dimensional integer array $S = (\eta_1, \ldots, \eta_n)$, $n = x_{max}$ covering the complete panorama in steps of $\Delta\phi = 2\pi/n$ (Fig. 2c).

Intensity values

The second place code used in our experiments, pixel intensity values, was extracted from greyscale versions of the panoramic input images (cf. Franz et al. 1998a). This place code is usually used with panoramic images reduced in the vertical direction to small rings about the horizon. This seems to work well in environments with only a few, prominent landmarks in front of a featureless background (e.g. Möller et al. 1998; Franz et al. 1998a; Möller 2000; Hübner and Mallot 2007; Smith et al. 2007) or in office environments with few occlusions (Hafner 2001; Stürzl and Mallot 2006). One reason for this may lie in the fact that along the horizontal ring, image features may move along that ring, but will never leave it. In the absence of occlusion, therefore, observer position can easily be inferred from landmark configurations along the horizontal ring. With numerous landmarks and cluttered environments, as are present in the ants' environment, occlusion will prevail, and the horizontal ring will probably not suffice to localise the observer. In the sequel, we will therefore consider vertically extended images of different vertical fields of view (FOV).

If the ants were to rely on intensity-based image codes, the field of view should include image regions

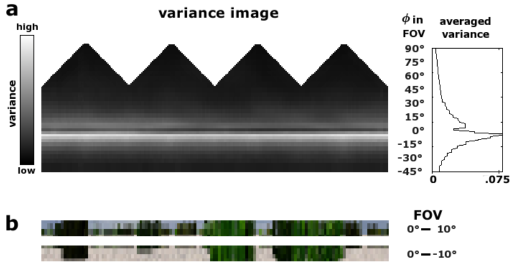


Fig. 4: Local variance of the intensity values (with range $0-1.0$) in panoramic images of the virtual environment (mean variance, $n = 6572$), all images oriented in fixed viewing direction. (**b**) Two (enlarged) examples of input images with vertical FOV of $0° - 10°$ and $0° - -10°$ providing an impression of the image areas with the highest variance.

with high variance, or dissimilarity, across snapshots. In contrast, homogeneous areas such as the sky or the ground are not useful for image based homing. In order to judge the potential relevance of image regions, we analysed the pixel-wise variances of intensity values across all images in our database. The result shows a high variance above and below the horizon (Fig. 4), but not at the horizon itself. In section 4.1 we will analyse image similarity in horizontal image rings of various elevational position and width.

In the following, intensity values I were coded in a one-dimensional integer array $S = (I_1, \dots, I_n)$ ($n = x_{max} \times y_{max}$) by putting the image rows one after the other.

3.5 Spatial specificity of place codes

Correspondence of snapshots was determined for both place codes, elevation and intensity which are now both represented as one-dimensional arrays of numbers. Two views or snapshots S and T were correlated calculating the uncentred correlation coefficient as:

$$\Phi(S,T) = \frac{\sum_{i=1}^{n} S_i T_i}{\sqrt{\sum_{i=1}^{n} S_i^2}\sqrt{\sum_{i=1}^{n} T_i^2}} \qquad (3)$$

The uncentred correlation function results in values Φ between 0.0 and 1.0. A value of 1.0 corresponds to a complete match, that is two identical images, whereas Φ decreases with decreasing similarity of the snapshots. Thus, places represented by snapshots comprising low correlation values can be distinguished with less ambiguity. Generally, correlation values will drop gradually when moving away from a snapshots position (cf. Fig. 5).

All tested images were recorded with a constant orientation. We therefore reduced computational effort by omitting the search for rotational correspondences. For a simple computational approach to simplify rotational correlation search, see Stürzl and Mallot (2006). Ants and many other hymenopteran insects may also align snapshots with help of their sun compass (review: Wehner 1997). Since the compass might be inaccurate or may even be missing in some animals or artificial systems, we will discuss the influence of rotation or compass noise on image ambiguity in Section 4.2.

3.6 Visual Landmark Navigation

The suitability of the two place codes, described by the criteria of relevance, uniqueness and reliability discussed above, was tested by combining each code with two homing models: (1) gradient ascent and (2) average landmark vector (ALV) model. The homing efficiency was characterised by the size of catchment areas. The catchment area is defined as the set of all locations from which an agent can be guided visually to the goal position.

Gradient Ascent

In the simplest case, visual homing can be achieved by continuously moving such as to increase the correlation (Eq. 3) between the stored snapshot and the current place code. Thus, the homing vector is computed as the gradient of Φ

$$\mathbf{v}(r,s) = \nabla\Phi(S(r,s),T) \qquad (4)$$

where r, s, are spatial coordinates with respect to which the gradient is taken, $S(r,s)$ is the current place code and T is the target place code. To each goal location, a catchment area exists within which this procedure will work, but this catchment area may be small. In general, however, the agent may be trapped in local maxima (Fig. 6a and cf. Franz et al. 1998b) which may appear at any distance from the target position. Figure 5 shows a contour plot of the correlation values in a 3 by 3 m patch of the environment with respect to a target location T at the marked coordinates. Multiple local maxima occur with absolute correlation values well below 1. Figure 6a shows the corresponding homing vectors of the gradient ascent algorithm together with the catchment area. For the calculation of the catchment areas, homing paths were simulated starting from all positions in the reference data base. At each step, the agent would proceed to the neighboring grid point with the highest correlation value.

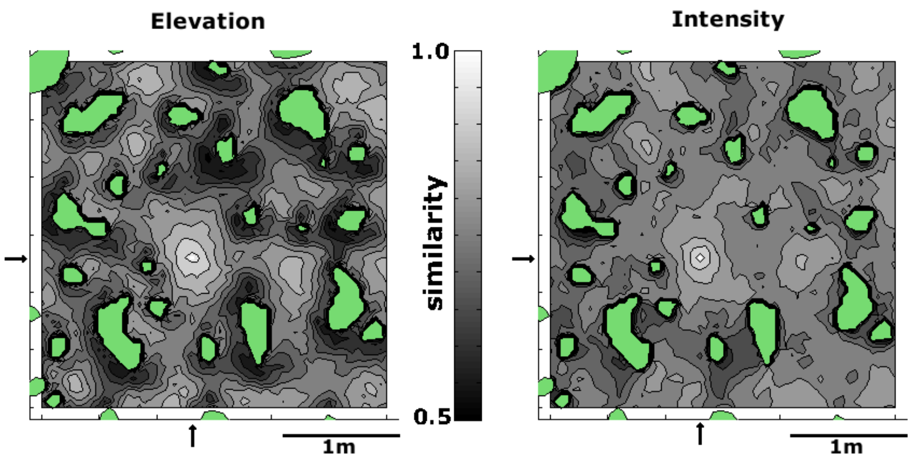


Fig. 5 Contour plots of image correlations $\Phi(S(r,s),T)$. The reference location at which place code T was obtained is marked by arrows at the s- and r-axes. Numerous local maxima are visible. Grass patches are indicated by regions with heavy outlines.

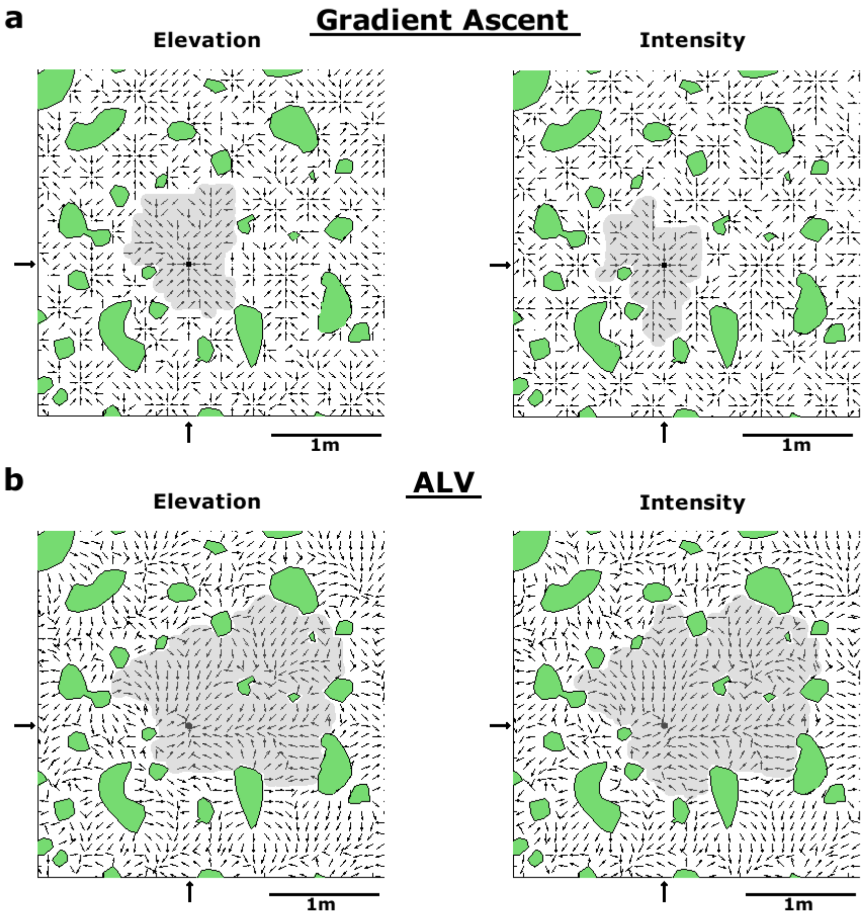


Fig. 6 Homing vector fields for both place codes, elevation and intensities, determined with (**a**) the gradient ascent method and (**b**) the ALV algorithm. Catchment areas are depicted in light grey, and the home or goal position is marked by a black circle and the heavy arrows. Grass patches are indicated by regions with black outlines.

ALV Model

In contrast to the gradient ascent mechanism, the ALV model reduces the full image or place codes S into a single two-component vector. Clearly, this implies enormous reduction of memory requirements. Originally, the ALV algorithm presupposes prior extraction of distinct landmarks or cues. Vectors of unit length pointing to each cue are processed into a single two dimensional vector which essentially points at the mean of all cue-bearings (Lambrinos et al. 2000). We adapted this procedure to pixel-based place codes by first calculating a landmark vector $\mathbf{v}_i$ for each pixel or component of the image code, S_i. The landmark vector points in the (horizontal) direction of the respective image pixel while the vector length reflects the value of S_i, i.e. elevation or intensity, respectively:

$$\mathbf{v}_i = S_i(\cos(i\varDelta\phi), \sin(i\varDelta\phi)). \tag{5}$$

Here, $\varDelta\phi = 2\pi/x_{max} = 1.5\text{deg}$ is the angular resolution along the horizon. The multiple image rows concatenated in the S-code in the case of intensity-coding will contribute to the appropriate directions since $\cos(i + x_{max})\varDelta\phi = \cos(i\varDelta\phi)$ etc. The "Average Landmark Vector" is then obtained by summing all individual landmark vectors ($\mathbf{v}_i$):

$$\mathbf{a} = \frac{1}{n}\sum_{i=1}^{n}\mathbf{v}_i; \tag{6}$$

the procedure is illustrated in Figure 7. For the intensity-based image code, the sketched procedure is equivalent to first vertically averaging the intensity values for all visual directions and then calculating the landmark vectors for these averages.

The ALV-model will be used for both the elevation and the intensity codes. For reasons explained below, only a central horizontal band of the intensity codes will be used in the simulations (0° to 10°).

The ALV model calculates a homing vector $\mathbf{h}$ by simply subtracting the ALV stored at the target position (t) from the ALV at the current position (c):

$$\mathbf{h}(c, t) = \mathbf{a}(c) - \mathbf{a}(t) \tag{7}$$

(Lambrinos et al. 2000). Note that the original algorithm does not take into account landmark distance, since one feature point and vector is considered for each landmark. In our modification, changing the distance to a landmark will affect the visual angles subtended both in the vertical and horizontal directions and will directly affect the elevation code. In this way, distance information is implicitly represented. Figure 6b shows two examples for catchment areas calculated with the ALV model for both tested place codes, elevation and pixel intensity. For the calculation of the catchment areas, homing paths were simulated starting at all grid points in the reference data base. At each time step, the ALV was computed and the agent proceeded to the neighboring grid point closest to the ALV direction.

In our continuous ALV version the number of vectors associated with a given landmark depends on the distance of the landmark (i.e. apparent size in the image) to the current position. Möller (2000) demonstrated that this feature may lead to performance loss since homing directions may form a saddle point around a home location, such that some trajectories will actualy miss the home and lead astray. This problem did not occur in our simulations. One reason for this might be the fact that in our algorithm the total number of vectors is fixed while Möller (2000) assigns vectors only to identified landmark regions in the image; the number of vectors therefore changes with the visual angle subtended by landmarks.

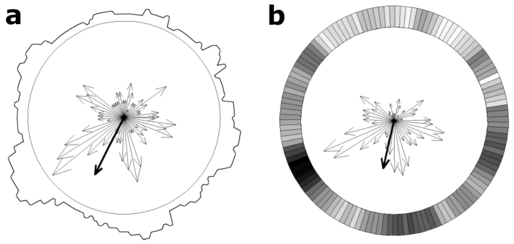


Fig. 7: Average landmark vector (ALV) model applied to (**a**) elevation code (skyline indicated by heavy black line) and (**b**) intensities code (indicated by pixel greylevels). The *small arrows* indicate the individual landmark vectors $\mathbf{v}_i$, scaled by the elevation or intensity in the according direction, the *heavy arrows* indicate the resulting average landmark vector $\mathbf{a}$.

4 Results

4.1 Spatial aliasing and uniqueness of place codes

In this section, we address the question which of the discussed place codes is most useful to the navigating ant. In general, usefulness requires a low degree of ambiguity since otherwise, places will be confused. We therefore analyse the images occurring along a path travelled by a real ant (Kohler and Wehner 2005) and analyse the ambiguities occurring when using the two types of place coding.

86 places were defined in the virtual environment, along a real ant trail (see Fig. 1). From these places, input images and place codes were generated. For the

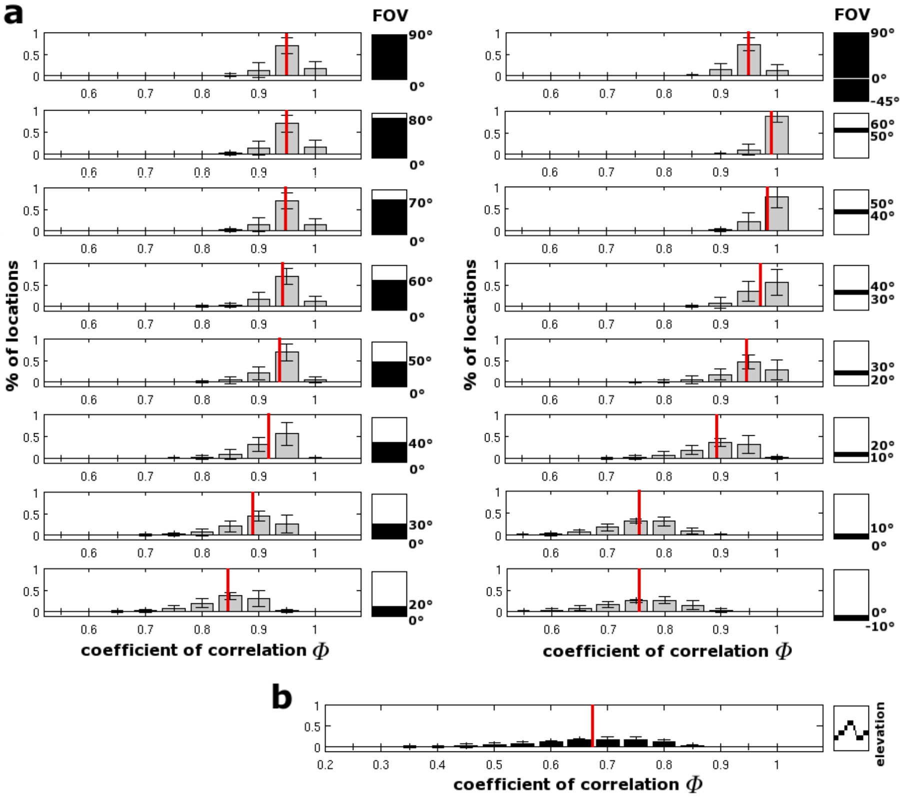


Fig. 8 Histograms of code correlation in the VR environment, for details see text. The red bar indicates the overall mean correlation, $\overline{\Phi}$. (**a**) Intensity code. The different panels show correlations obtained with various vertical fields of view (FOV), as indicated by the insets to the right of each panel. (**b**) Elevation or skyline code. Elevation is always calculated from the full field of view.

intensity code, a dependence of pixel variance on vertical position has been demonstrated in Figure 4. We therefore tested different vertical fields of views (FOV), comprising panoramic image strips of various vertical position and extension. In contrast, the elevation code was always extracted from the full panoramic view.

For each of the 86 test positions and each field of view, we calculated the code correlations with all 6572 snapshots from our reference database according to Eq. 3. The resulting Φ-values were transfered into histograms which were then averaged bin-wise across the 86 test positions. Figure 8 shows the resulting mean histograms of the correlation values. The error bars for each bin indicate the standard deviation across the 86 test positions; the red bars show the overall average correlation value $\overline{\Phi}$. If a place code generates low ambiguity, its histogram will be shifted leftwards in Figure 8.

Using pixel intensities (Fig. 8a), the best results, i.e. lowest correlation values, were achieved using a vertical FOV from 0° to +10° ($\overline{\Phi} = 0.757$) and a vertical FOV 0° to −10° ($\overline{\Phi} = 0.756$). These are the areas in the image previously identified as the areas with the highest variability (Fig. 4). In the following, unless otherwise noted, the intensity code will be used with a vertical FOV of 0° to +10°. The elevation code (Fig. 8b) generates even smaller correlation values ($\overline{\Phi} = 0.672$) than indicated for intensity code with the best vertical field of view. Therefore using the elevation or skyline code will provide the ant with the least ambiguous landmark information.

4.2 Correlation of misaligned images

For matching the current view and a stored snapshot is is important to properly align the two views with respect to some global orientation, or to each other.

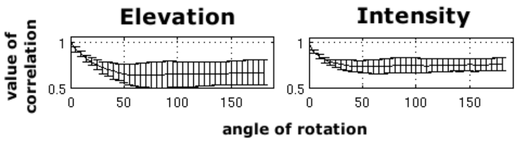


Fig. 9: Effect of rotational misalignment on code correlation. Rotated snapshots were compared to their unrotated originals (mean and standard deviation over 86 test images).

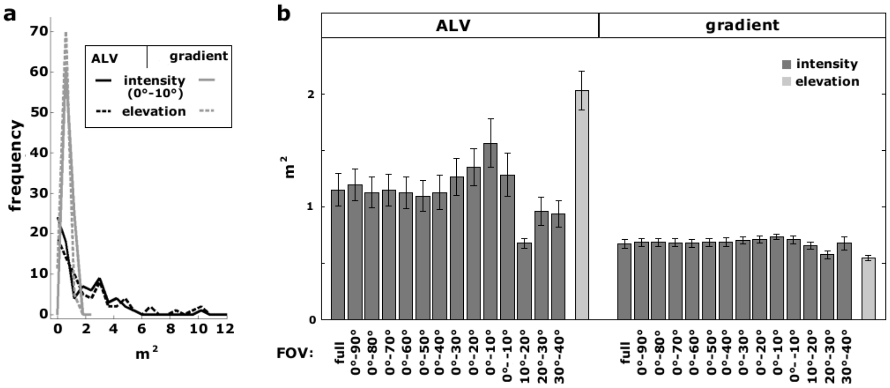


Fig. 10 (**a**) Distribution of catchment area sizes for four combinations of coding and homing schemes. Intensity code data are shown for the FOV of $0° - 10°$, i.e. the FOV leading to the largest catchment areas. (**b**) Average catchment area sizes and standard errors for all goal images of the test set (n=86). Various FOVs for intensity codes were tested (cf. Fig. 8).

To determine the effect of heading rotations (caused e.g. by an imprecise compass) on the ability of recognising a snapshot, we rotated the snapshots of our test set and compared them with their unrotated originals (Fig. 9).

On average, the value of code correlation decreases primarily in the range of rotations up to 40 degrees. For larger rotations, correlation drops to 0.75 for intensity coding (FOV=0°-10°) and to 0.66 for elevation coding. These values correspond to the average correlation values $\overline{\Phi}$ found already in Figure 8. Thus, the ability of place recognition decreases with increasing misalignment of current and stored views. However, small rotations ($< 5°$) caused e.g. by noise of the compass system have minor effects. No clear preference for the intensity or elevation coding scheme can be derived from this result.

4.3 Homing performance and the relevance of place codes

As a third criterion addressing the quality of a place coding and homing scheme, we consider the size of the resulting catchment areas (see Section 3). Figure 6 shows examples for the two place codes, elevation and intensity (with various FOV), and the two homing algorithms, gradient ascent and ALV. Catchment areas were calculated taking all images of our test set ($n = 86$) as goal locations and all other images in the database ($n = 6572$) as starting points. In the resulting homing fields (cf Fig. 6), catchment areas were defined as sets of points from which the goal can be reached (i.e. up to a residual homing error of 7 cm) by following the local homing vectors. Finally, the catchment area sizes were calculated.

The distributions of catchment area size obtained with the various procedures is depicted in Figure 10a. Most notably, the sizes of the catchment areas show greater differences between the homing algorithms than between the place codes. While the gradient ascent algorithm mostly generates smaller catchment areas, the ALV model shows a substantial portion of large and very large catchment areas, as indicated by the tail of the distribution in Figure 10a. On average, catchment areas are considerably larger for the ALV algorithm than for gradient ascent (Fig. 10b). Within each homing algorithm we found differences in the catchment area size between the place codes and between the different fields of view used with the intensity code (FOVs $40° - 50°$ and $50° - 60°$ exhibit insufficient catchment areas (cf. correlation histogramms Fig. 8) and are not shown here). Whereas the elevation code allows for slightly larger catchment areas in the ALV homing scheme, the intensity codes lead to somewhat larger catchment areas when homing with the gradient ascent algorithm. FOV size clearly affects the ALV scheme, while the influence on the gradient scheme is rather small. Within the intensity codes best performance (i.e. largest catchment areas) is obtained with a FOV of $0° - 10°$ that might have been expected from its low correlation values reported in section 4.1. The difference between ALV and gradient ascent is also appearant in the examples in Fig. 6. The ALV field is characterised by extended fronts but few fix points while fix points (local maxima) are frequent in the gradient ascent field.

Figure 11a shows examples of homing trajectories using the elevation or skyline code ALV model. An agent was placed at several positions around the goal position. In each step, the agent extracts the skyline from the panoramic view, calculates the ALV, and continuously estimates the next heading direction by the difference of the current ALV and the stored target ALV. Whereas trajectories starting in the catchment

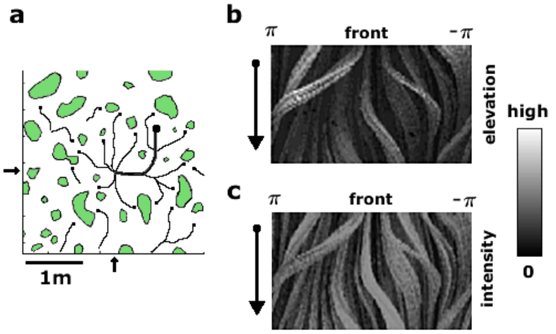


Fig. 11: (**a**) Sample homing trails generated from the ALV model using the elevation code. Starting points are marked by circular dots. The point where several trails converge is the goal location (marked by outer arrows). Agents starting outside the catchment were trapped in false targets (blindly ending trails). **b.c** Development of the place codes along one route (*heavy line*) for the elevation code (**b**) and the intensity code (**c**). Each horizontal line shows one panoramic snapshot. The centre of the snapshot ("front") corresponds to the downward direction in Fig. (a).

area of the target snapshot end at the correct goal position, the agent was trapped at false targets (locations with indistinguishable ALV) otherwise. Developing place codes along a trail are depicted in Figure 11b,c. Both place codes show a smooth continuum. Note the higher overall contrast in the elevation code, which reflects the result from Figure 8, i.e. the higher discriminability of the elevation code.

Figure 12 shows the inverse normalised average landmark vectors ($-\mathbf{a}^o$) for a large portion of the environment. This field is independent of a particular goal location, since only the ALVs, not the homing vectors are shown. Note that all vectors point away from the obstacles, forming fronts along the major passage ways in the environment. Thus, the ALV scheme exhibits an inherent obstacle-avoidance property. This is also visible from Figure 11a, where trajectories do not take the shortest path, but avoid obstacle by wide margins.

5 Discussion

Skyline coding. In this paper we analyse a new insect-inspired approach to visual homing in outdoor environments, which was derived from desert ant's navigational behaviour and perceptual characteristics. On the experimental side the idea of skyline-coding is supported by a number of studies, including Pratt et al. 2001; Fukushi 2001; Heusser and Wehner 2002; Graham and Cheng 2009a,b; here, we provide a computational test for the validity of skyline coding in complex, real-world environments. In summary, skyline coding has the following advantages: First of all, skyline information is easily extracted from the UV and green colour channels, both in ants and in technical applications. Secondly, skyline codes show low similarity or ambiguity (Fig. 8), and comprise catchment areas comparable to intensity coding when combined with both, gradient ascent or ALV homing (Fig. 10). In contrast to intensity coding, which depends strongly on the stability of illumination (Zeil et al. 2003; Stürzl and Zeil 2007), skyline coding is rather robust with respect to both the direction of illumination (sun position) and weather conditions (cloudy skys), see Möller 2002; Kollmeier et al. 2007.

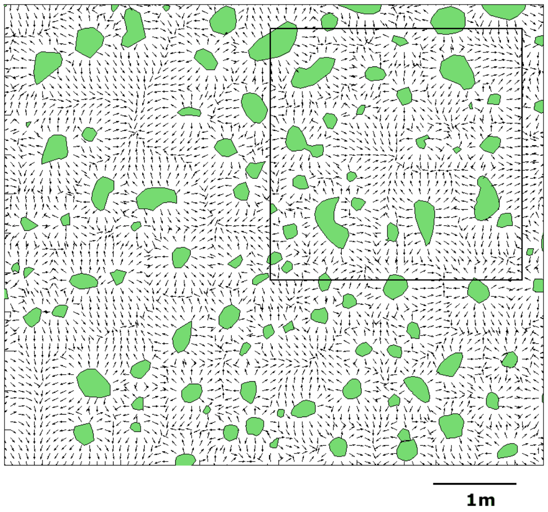


Fig. 12: Vector field made up of inverted and normalised ALVs (calculated from elevation codes). The ALV method exhibits an inherent obstacle-avoidance mechanism, when following the inverted ALVs. The grey frame marks the figure detail of Fig. 5,6.

UV vision. In our simulation, we used blue-green rather than UV-green illumination and contrasts, which clearly does not make any difference in a virtual environment. In real outdoor environments, the use of UV-light for visual homing is important, since best landmark-sky separation in outdoor scenes is obtained by UV light (single channel classifier), or with UV-IR contrast in dual-channel classification. UV-green contrast, as given by the spectral properties of the ant's visual receptors, yields a somewhat lower quality of separation but still performs better than all single channel classifiers and may be sufficient for skyline separation in the real world (Möller 2002; Kollmeier et al. 2007).

Insect colour vision systems are generally organised in terms of UV-blue-green-trichromacy (Briscoe and Chittka 2001). The UV channel may help guiding flower-visitors to their targets by floral colour patterns. Further it has been shown that UV sensitivity can advance detection of celestial polarisation under clouds and canopies (Barta and Horvath 2004). Our results suggest that the UV channel might play also another role, i.e. skyline-based landmark perception for illumination independent visual navigation in outdoor environments.

Skyline cues in technical applications. The usefulness of skyline codes has been noticed also in technical place recognition systems. Cozman et al. (2000) explore the skyline information for visual position estimation in planetary robots. The skyline is preselected by the intensity gradient between sky and ground, and determined by the user in a semi-automatic procedure. The system allows coarse position estimation in natural (planetary) open environments. Skyline extraction was also used in urban position estimation (Johns and Dudek 2006). In this study, it was assumed that the sky region is generally much brighter than foreground objects. The skyline was then extracted by selecting and smoothing the uppermost edge in the output image of an edge-detector. In a performance test, 15 out of 17 test images of buildings could be classified correctly by matching contour segments. It should be noted, however, that intensity values alone do not provide stable features for skyline extraction, since it strongly depends on illumination, wheather conditions, and the surface properties of the ground object. Recently, an algorithm was presented that makes use of image pixel information to determine the position and orientation of the horizon in natural scences. Monitoring the horizon, the attitude of an unmanned aircraft was inferred (Thurrowgood et al. 2009). The sky/ground contrast was enhanced by transforming RGB values to a one-dimensional colour-brightness value. By this value ground and sky could be separated with a small classes overlap of 8.6%.

Homing and image correlation. Catchment areas for the combination of intensity coding and ALV homing were largest for the 10° of visual field above the horizon which also showed the lowest amount of correlation across various images. Also for the gradient ascent algorithm, this field of view yielded the best results, if only by a small margin. In principle, there seems to be no simple theory to relate image correlation to catchment area size. Predictions of this size need to take into account the spatial distribution of image correlation. If we denote by $I(x, y)$ the image taken at location (x, y) and treat it as a pixel-vector, it seems promising to consider the autocorrelation function of the image manifold,

$$\Psi(x', y') := \int\int (I(x, y) \cdot I(x + x', y + y'))dxdy \qquad (8)$$

where $(\cdot)$ denotes the dot product. We hypothesize that for good homing performance, this autocorrelation function should be large for small values of x' and y' and small for larger arguments. In our analysis so far, we have only considered the average value of Ψ for all x' and y'. It should also be noted that catchment area size depends not only on the homing scheme but as well on the structure of the environment (e.g. number and distribution of landmarks).

Further, misalignment of images might cause misinterpretation of the current position estimated by snapshot comparison. An agent can accomplish this by internal rotation of the views or by aligning the views through body rotations (e.g. Franz et al. 1998b; Lambrinos et al. 2000; Stürzl and Mallot 2006; Möller 2009). Experiments with ants with alternately occluded eyes show that memorised views from one eye can not be recognised if seen with the other eye (Wehner and Müller 1985). Thus, an internal rotation of views seems to be unlikely. Since ants can extract a compass direction using their sun compass, it may be possible for the animals to align the body axis with the axis of a memorised view. Indeed, some species of *Cataglyphis* show body rotations on excursions (Wehner 1994). Alternatively, ants could store a set of snapshots with different orientations activated selectively by the current compass direction (Cartwright and Collett 1983) or by the best matching view.

Image resolution. Kohler and Wehner (2005) have shown in their experiments that ants are able to recognise their habitual routes in a complex environment. As a mechanism for this performance, Smith et al. (2007) suggested that ants could use low spatial frequency information (i.e. from large distant landmarks) in combination with local landmark identification. Our results suggest that the ants might be able to identify their position on the basis of a single resolution channel, however with a space-variant distribution of local scales. Indeed, image resolution in the ant's complex eye is higher in the horizontal direction and decreases towards the ground and the zenith where the skyline cues are likely to occur. Thus, the morphology of the eye itself might already serve as a space-variant filter, which is dependent on the elevation in the view.

6 Conclusion

In summary, skyline information used for visual homing has a number of attractive properties. The full panoramic visual input is used and encoded to a ring-like place code representing the skyline with low memory requirements. Further, the skyline can be extracted from colour contrasts with low computational effort, which is plausible with respect to the layout of the ant's eye and photoreceptors. More work is needed to understand the statistics of skyline codes in natural environments of varying object density and height, and the usefulness for such codes in determining directions and locations. Furthermore, the inclusion of UV light modules into visual navigation systems may prove useful to achieve robust outdoor navigation.

Acknowledgements We are grateful to Rüdiger Wehner and his group for providing ant data and for helpful comments on this project. Support by the 'Landesgraduiertenförderung Baden-Württemberg' to KB is gratefully acknowledged.

References

Anderson AM (1977) A model for landmark learning in the honey-bee. J Comp Physiol A 114 (3):335–355

Barta A, Horvath G (2004) Why is it advantageous for animals to detect celestial polarization in the ultraviolet? Skylight polarization under clouds and canopies is strongest in the UV. J Theor Biol 226: 429–437

Briscoe AD, Chittka L (2001) The evolution of color vision in insects. Annu Rev Entomol 46:471–510

Cartwright BA, Collett TS (1983) Landmark learning in bees. J Comp Physiol A 151 (4):521–543

Cheng K, Narendra A, Sommer S, Wehner R (2009) Traveling in clutter: Navigation in the Central Australian desert ant *Melophorus bagoti*. Behav Process 80 (3):261–268

Cozman F, Krotov E, Guestrin C (2000) Outdoor visual position estimation for planetary rovers. Auton Robot 9:135–150

Franz MO, Schölkopf B, Mallot HA, Bülthoff HH (1998a) Learning view graphs for robot navigation. Auton Robot 5:111–125

Franz MO, Schölkopf B, Mallot HA, Bülthoff HH, 1998b) Where did I take that snapshot? Scene-based homing by image matching. Biol Cybern 79:191–202

Fukushi T (2001) Homing in wood ants, *Formica japonica*: use of the skyline panorama. J Exp Biol 204: 2063–2072

Gillner S, Weiß AM, Mallot HA (2008) Visual place recognition and homing in the absence of feature-based landmark information. Cogn 109:105–122

Gourichon S, Meyer JA, Pirim P (2002) Using coloured snapshots for short-range guidance in mobile robots. Int J Robot Autom 17:154–162

Graham P, Cheng K (2009a) Ants use the panoramic skyline as a visual cue during navigation. Curr Biol 19 (20):R935–R937

Graham P, Cheng K (2009b) Which portion of the natural panorama is used for view-based navigation in the Australian desert ant? J Comp Physiol A 195:681–689

Hafner VV (2001) Adaptive homing - robotic exploration tours. Adapt Behav 9 (3-4):131–141

Heusser D, Wehner R (2002) The visual centring response in desert ants - *Cataglyphis fortis*. J Exp Biol 205:585–590

Hübner W, Mallot HA (2007) Metric embedding of view-graphs: A vision and odometry-based approach to cognitive mapping. Auton Robot 23:183–196

Johns D, Dudek G (2006) Urban position estimation from one dimensional visual cues. In: Proceedings of the the 3rd Canadian Conference on Computer and Robot Vision (CRV'06), pp 22

Kohler M, Wehner R (2005) Idiosyncratic route-based memories in desert ants, *Melophorus bagoti*: How do they interact with path-integration vectors? Neurobiol Learn Mem 83:1–12

Kollmeier T, Röben F, Schenck W, Möller R (2007) Spectral contrasts for landmark navigation. J Opt Soc Am A 24:1–10

Lambrinos D, Möller R, Labhart T, Pfeifer R, Wehner R (2000) A mobile robot employing insect strategies for navigation. Robot Auton Syst 30:39–64

Lehrer M, Bianco G (2000) The turn-back-and-look behaviour: bee versus robot. Biol Cybern 83 (3):211–229

Liu G, Seiler H, Wen A, Zars T, Ito K, Wolf R, Heisenberg M, Liu L (2006) Distinct memory traces for two visual features in the *Drosophila* brain. Nature 439:551–556

Möller R (2000) Insect visual homing strategies in a robot with analog processing. Biol Cybern 83 (3):231–243

Möller R (2002) Insects could exploit UV-green contrast for landmark navigation. J Theor Biol 214 (4):619–631

Möller R (2009) Local visual homing by warping of two-dimensional images. Robot Auton Syst 57 (1):87–101

Möller R, Lambrinos D, Pfeifer R, Labhart T, Wehner R (1998. Modeling ant navigation with autonomous agents. In: Proc. 5th Int. Conf. Simulation of Adaptive Behavior, pp 185–195

Möller R, Vardy A (2006) Local visual homing by matched-filter descent in image distances. Biol Cybern 95 (5):413–430

Mote M, Wehner R (1980) Functional characteristics of photoreceptors in the compound eye and ocellus of the desert ant, *Cataglyphis bicolor*. J Comp Physiol 137:63–71

O'Keefe J (1991) Brain and Space. Oxford University Press, Ch. The hippocampal cognitive map and navigational strategies, pp 273–295

Pratt SC, Brooks SE, Franks NR (2001) The use of edges in visual navigation by the ant *Leptothorax albipennis*. Ethol 107:1125–1136

Smith L, Philippides A, Graham P, Baddeley B, Husbands P (2007) Linked local navigation for visual route guidance. Adapt Behav 15 (3):257–271

Stürzl W, Mallot HA (2006) Efficient visual homing based on fourier transformed panoramic images. Robot Auton Syst 54:300–313

Stürzl W, Zeil J (2007) Depth, contrast and view-based homing in outdoor scenes. Biol Cybern 96:519–531

Thurrowgood S, Soccol D, Moore RJD, Bland D, Srinivasan MV (2009) A vision based system for attitude estimation of UAVs. In: The 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp 5725–5730

Tinbergen N, Kruyt W (1938) Über die Orientierung des Bienenwolfes III. Die Bevorzugung bestimmter Wegmarken. Z vergl Physiol 25:292–334

Vardy A, Möller R (2005) Biologically plausible visual homing methods based on optical flow techniques. Connect Sci 17 (1-2):47–89

Wehner R (1994) Neural Basis of Adaptive Behaviour. G. Fischer, Stuttgart, Ch. The polarization-vision project: championing organismic biology, pp 103–143

Wehner R (1997) Orientation and Communication in Arthropods. Birkenhäuser Verlag Basel / Zürich, Ch. The ants celestial compass system: spectral and polarization channels, pp 145–185

Wehner R, Müller M (1985) Does interocular transfer occur in visual navigation by ants. Nature 315:228–229

Wehner R, Räber F (1979) Visual spatial memory in desert ants, *Cataglyphis bicolor* (Hymenoptera: Formicidae). Experientia 35:1569–1571

Zeil J, Hofmann MI, Chahl JS (2003) Catchment areas of panoramic snapshots in outdoor scenes. J Opt Soc Am A 20 (3):450–469

Zollikofer CPE, Wehner R, Fukushi T (1995) Optical scaling in conspecific *Cataglyphis* ants. J Exp Biol 198:1637–1646

Manuskript 2

Route learning strategies in a virtual cluttered environment

Ziele, Hauptergebnisse und Eigenanteil an der Publikation/dem Manuskript

Alltäglich bewegen wir uns von zu Hause zum Arbeiten, zum Einkaufen oder besuchen Freunde und Bekannte. In vertrauten Umgebungen werden die zielführenden Wege dabei oft unterbewußt geplant. Jedoch ist noch nicht vollständig geklärt, welche Navigationsstrategien verwendet werden, um neue Routen in unbekannten Umgebungen zu lernen und zu planen. Um dieser Frage nachzugehen, entwickelte ich zusammen mit Herrn Dr. Jan M. Wiener[2] ein Experiment zum Routenlernen bei Menschen. Für das Experiment wurde eine landmarkenreiche Umgebung, ähnlich der Umgebung, die in der Ameisensimulation verwendet wurde (s. Manuskript 1), in einem Computermodell in virtueller Realität (VR) implementiert. Sie bestand aus einer Vielzahl aus prismatischen Säulen (lokale Landmarken), die sich in Höhe und Textur glichen, aber in der Form ihrer Umrisse unterschieden. Zusätzlich wurden in weiter Entfernung vier große globale Landmarken (Säulen) postiert, die in vier entgegengesetzten Richtungen aufgestellt den Probanden als eine Art Kompass dienen konnten. Die Versuchspersonen mußten nun am Computer mit Hilfe eines Joysticks den Weg zwischen zwei Zielpunkten finden, lernen und ihm in mehrfachen Versuchsdurchgängen wiederholt folgen.

Dabei zeigte sich, dass einige Versuchspersonen eine einzige Route lernten und dieser immer wieder folgten, während andere vielmehr die grobe Richtung zum Ziel wählten und dabei die verschiedenen Hindernissen mal auf der einen, mal auf der anderen Seite passierten und damit mehrere variable Routen nutzten. Die Ergebnisse zeigen, dass sich die Versuchspersonen nicht in zwei Lager aufteilten, d.h. solche, die nur einer konstanten Route folgten und solche, die ganz variable Routen nutzten, sondern vielmehr ein Kontinuum in der Nutzung von sehr konstanten hin zu sehr variablen Routen besteht.

In weiteren Versuchen wurde die Rolle der großen globalen Landmarken und der lokalen Hindernisse untersucht. Es zeigte sich, dass mit steigender Variabilität der Routen die Versuchspersonen vermehrt die globalen Landmarken zur Orientierung

---

[2] (seit 2009) *School of Design, Engineering and Computing*, Bournemouth University, UK

genutzt hatten. Hingegen waren für die Personen, die zu konstanten Routen neigten, die lokalen Landmarken wichtiger für die Navigation.

Das Projekt und die damit verbundene Programmierung wurde gemeinsam mit Herrn Dr. Jan M. Wiener geplant und vorbereitet. Die Versuche und Auswertungen wurden zusammen mit Frau Rebecca Hurlebaus durchgeführt, die im Rahmen dieses Projekts auch ihre Diplomarbeit angefertigt hat. Für die Auswertung der Routenähnlichkeit wurde ein Teil des Verfahrens zur Trajektorienanalyse aus Manuskript 3 verwendet. Die Veröffentlichung wurde mit den Autoren, insbesondere zusammen mit Herrn Dr. Jan M. Wiener, umgesetzt.

Die Publikation ist im Rahmen der *International Conference Spatial Cognition 2008* erschienen in "Spatial Cognition VI, Lecture Notes in Artificial Intelligence 5248, Spinger" (Peer-Reviewed). Der Artikel wurde auf der Konferenz in die Liste der Nominierungen für den "Best Paper Award" aufgenommen.

Referenz:
R. Hurlebaus, K. Basten, H.A. Mallot, and J.M. Wiener (2008), *Route learning strategies in a virtual cluttered environment.* In: C. Freksa et al., (Ed.), Spatial Cognition (SC) 2008, number 5248 in Lecture Notes in Artificial Intelligence, 104-120.

Route Learning Strategies in a Virtual Cluttered Environment

Rebecca Hurlebaus[1], Kai Basten[1],
Hanspeter A. Mallot[1], and Jan M. Wiener[2]

[1] Cognitive Neurosccience, University of Tübingen, Auf der Morgenstelle 28, D-72076 Tübingen, Germany
[2] Center for Cognitive Science, University of Freiburg, Friedrichstr. 50, D-79089 Freiburg, Germany

Abstract. Here we present an experiment investigating human route learning behavior. Specific interest concerned the learning strategies as well as the underlying spatial knowledge. In the experiment naive participants were asked to learn a path between two locations in a complex, cluttered virtual environment that featured local and global landmark information. Participants were trained for several days until they solved the wayfinding task fastly and efficiently. The analysis of individual navigation behavior demontrates strong interindividual differences suggesting different route learning strategies: while some participants were very conservative in their route choices, always selecting the same route, other participants showed a high variability in their route choices. In the subsequent test phase we systematically varied the availability of local and global landmark information to gain first insights into the spatial knowledge underlying these different behavior. Participants showing high variability in route choices strongly depended on global landmark information. Moreover, participants who were conservative in their route choices were able to reproduce the basic form of the learned routes even without any local landmark information, suggesting that their route memory contained metric information. The results of this study suggest two alternative strategies for solving route learning and wayfinding tasks that are reflected in the spatial knowledge aquired during learning.

Keywords: spatial cognition, route learning, navigation.

1 Introduction

Finding the way between two locations is an essential and frequent wayfinding task for both animals and humans. Typical examples include the way from the nest to a feeding site or the route between your home and the office. While several navigation studies, both in real and virtual environments, investigated the form and content of route knowledge (e.g., [1,2,3]), empirical studies investigating the route learning process itself are rather limited (but see [4,5]).

A very influential theoretical framework of spatial knowledge acquisition proposes three stages when learning a novel environment [6]. First, landmark knowledge, i.e., knowledge about objects or views allowing to identify places, is acquired.

C. Freksa et al. (Eds.): SC 2008, LNAI 5248, pp. 104–120, 2008.

In the second stage, landmarks are combined to form route knowledge. With increasing experience in the environment, survey knowledge (i.e. knowledge about distances and direction between landmarks) emerges. According to this model, the mental representation of a route can be conceived as a chain of landmarks or places with associated movement directives (e.g. turn right at red house, turn left at the street lights). This *landmark* to *route* to *survey knowledge* theory of spatial learning has not remained unchallenged: Recent findings, for example, demonstrate that repeated exposures to a route not necessarily resulted in improving metric knowledge between landmarks encountered on the route [5]. Most participants either had accurate knowledge from the first exposure or they never acquired it. Furthermore, results from route learning experiments in virtual reality suggest two spatial learning processes that act in parallel rather than subsequently [4]: (1) a visually dominated strategy for the recognition of routes (i.e., chains of places with associated movement directives) and (2) a spatially dominated strategy integrating places into a survey map. The latter strategy requires no prior connection of places to routes. Support for parallel rather than subsequent learning processes also comes from experiments with rats: depending on the exact training and reinforcement procedure, rats can be trained to approach positions that are defined by the configuration of extramaze cues (c.f. spatially dominated strategy), to follow local visual beacons (c.f. visually dominated strategy), or to execute motor responses (e.g., turn right at intersection; [7,8]). Evidence for a functional distinction of spatial memories also comes from experiments demonstrating that participants who learned a route by navigation performed better on route perspective tasks, while participants who learned a route from a map performed better on tasks analysing survey knowledge [9].

In any case, route knowledge is usually described as a chain of stimulus-response pairs [10,11], in which the recognition of a place stimulates or triggers a response (i.e., a direction of motion). Places along a route can be recognized by objects but also by views or scenes [12]. Evidence for this concept of route memory mostly comes from experiments in mazes, buildings, or urban environments, in which decision points were well defined (e.g. [1,13,3]). Furthermore, distinct objects (i.e., unique landmarks) are usually presented at decision points. Route learning in open environments, in contrast, has received little attention in humans, but has been convincingly demonstrated in ants [2]. The desert ant *Melophorus bagoti* is a singly foraging ant and its environment is characterized by cluttered distributed small grass tussocks. The ants establish idiosyncratic routes while shuttling back and forth between a feeder and their nest. Each individual ant follows a constant route for inbound runs (feeder to nest) and outbound runs (nest to feeder). Usually both routes differ from each other and show a high directionality [14]. In contrast, wood ants can learn bi-directional routes when briefly reversing direction and tracing their path for a short distance [15]. For both ant species view-dependend learning is essential for route learning in open cluttered environments [16]. View dependent representations [17] and view dependent recognition of places has also been demonstrated in humans and has been shown to be relevant for navigation [12].

Most studies investigating route knowledge in humans were conducted in urban environments in which the number of route alternatives between locations as well as possible movement decisions at street junctions are rather limited. How do humans behave when faced with a route learning task in open environments, lacking road networks, predefined places, and unique objects or landmarks? Are they able to learn their way between two locations in such environments? And if so, what are the underlying route learning processes?

1.1 Synopsis and Predictions

In the following we present a route learning experiment in an open cluttered environment characterized by prismatic objects differing in contour but neither in hight nor in texture. The environment did not contain any predefined places, road networks, or unique landmarks. Distal cues were present in form of four large colored columns and background texture. Participants task was to explore the environment and to shuttle between two target locations repeatedly. We monitored participants navigation and route learning behavior during an extensive training phase. Subsequently, we tested the impact and influence of proximal and distal spatial information on participants navigational ability.

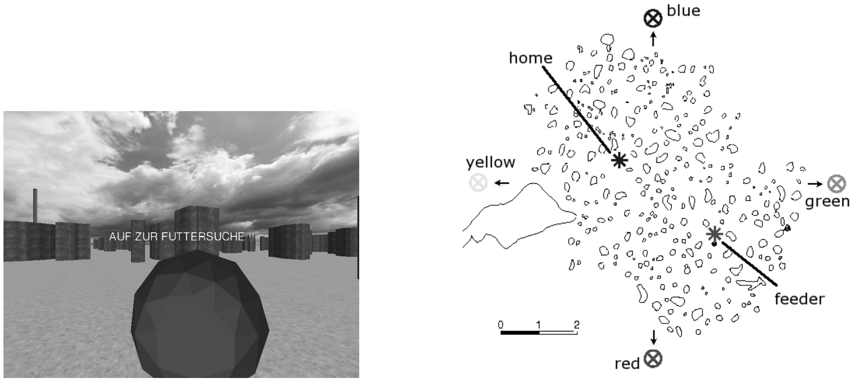
We expected that over an extended period of training, participants were able to solve the general experimental task (i.e., to fast and efficiently navigate between the home and the feeder position). It was, however, an open question whether participants established fixed routes (as ants do when faced with such a task in a similar environment [2]) or whether they learned global directions and distances between the relevant locations. The latter alternative would allow solving the task without explicit route knowledge but requires spatial knowledge that is best described as survey knowledge. In contrast to route knowledge, survey knowledge allows for more flexible navigation behavior when shuttling between distant locations. Consequently, one might expect a higher variability of (similarly efficient) route choices between navigations. Moreover, it is possible that different participants adopted or weighted these alternative spatial learning strategies differently.

In the test phase, we systematically varied the availability of local and global cues to study which spatial information was relevant for solving the wayfinding task. If participants established fixed routes, they were expected to strongly depend on local (i.e, proximal) spatial information to guide their movements. Hence, if that information was removed in a no-local-objects test, their navigation performance was expected to decrease dramatically. If, on the other hand, participants relied on global directions and distal information to solve the task, we expected their navigation performance to drop when such information was removed by adding fog to the environment.

2 Material and Methods

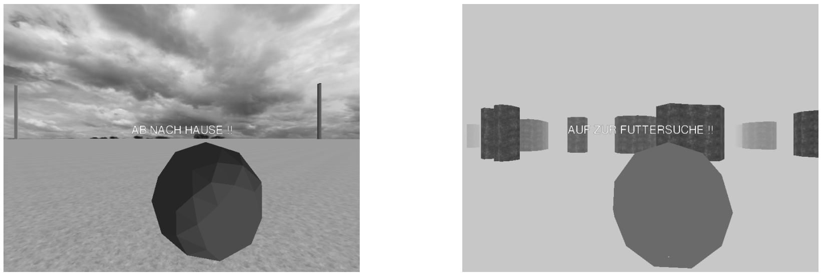
2.1 Participants

Twenty-one students of the University of Tübingen participated in this study (10 females). The average age was 24 years (range 19-28). No participant had



(a) Participants view

(b) Map



(c) Condition: no-local-objects

(d) Condition: fog

Fig. 1. **(a)** The virtual environment from the perspective of the participant at the home position. The sphere was visible only in close proximity. The text specified the current task (here:"Search for the feeder!"). A distal landmark (large small column) is visible in the background; **(b)** A map of the environment: the positions of home and feeder are marked by asterisks; the crossed circles indicate the positions of the colored columns (for illustration columns were plotted closer to the center of the environment); **(c)** The *no-local-objects* condition; **(d)** The fog condition.

prior knowledge of the virtual environment or the experimental hypotheses at the time of testing. They were paid 8€ per hour. One participant (female) had to be excluded, because of motion sickness during the first experimental trials.

2.2 Virtual Environment

The virtual environment was generated using Virtual Environments Library (VeLib)[1]. It consisted of a ground plane cluttered with objects of equal height and texture, which differed only in the shape of their groundplate. The background texture consisted of a cloudy sky and separate flat hills. To provide distinct global landmark information, four large columns of different colour (red, blue,

---

[1] http://velib.kyb.mpg.de/ (March 2008)

green, yellow) were positioned on four sides of the environment, at a distance of 80 units from the center. The directions of these globale landmarks are shown in the environment map in Figure 1b, but are plotted closer to the obstacles for the sake of clarity. Two additional objects, a red sphere and a blue sphere, that marked the relevant location (referred to as home and feeder in the following) were placed in the environment with a distance of $\tilde{5}.5$ units between them. These were so-called pop-up objects that were visible only at close proximity (< 0.4 units). An experimental session always started at the blue sphere that was referred to as the *home* location. In analogy to experiments with ants (see Introduction), the red sphere was referred to as the *feeder*. Figure 1a displays the participants' view within the virtual environment.

2.3 Experimental Setup

The virtual environment was presented on a standard 19" computer monitor. Participants were sitting in front of the monitor on an office chair in a distance of approximately 80cm. Using a standard joypad (Logitech RumblePad 2) they steered through the virtual environment. Translation and rotation velocity could be adjusted separately by the two analog controls. Maximum translation velocity was 0.4 units per second; maximum rotation velocity was 26° per second. All participants were instructed how to use the joypad and had chance to familiarize themselves with the setup.

2.4 Procedure

General Experimental Task and Procedure. The general experimental task was to repeatedly navigate between two target locations, the home (blue sphere) and the feeder (red sphere). During navigation, the target (home or feeder) for the current run was indicated by a text message (e.g., "Go home!"). As soon as the participant moved over the current target (e.g., blue sphere indicating home location), the respective text message changed (e.g., "Search for the feeder!"). Runs from *home* to *feeder* are referred to as outbound runs, runs from the *feeder* to *home* are referred to as inbound runs. Experimental session always started at the home position. As participants were naive with respect to the environment, the experiment had an extensive training-phase prior to the test-phase.

Training-Phase. The training-phase consisted of several sessions, during which participants were instructed to repeatedly navigate between *home* and *feeder*. At the beginning of each session participants were positioned at the home location. Pilot experiments demonstrated that the experimental task was very difficult in the first run; participants were therefore provided with course directional information about the direction from home to the other feeder at the beginning of the first session ("seen from home, the feeder is situated between the red and the green distal column"). Participants unterwent two training sessions per day with a 5min break between them. A single training session was completed with the first visit of the home after 20min. In case participants showed a performance of > 2 runs per minute (for inbound and outbound runs) after 5 training session, they advanced to the test phase. The maximal number of training sessions was 9.

Test-Phase. In the test phase participants were confronted with 2 wayfinding tasks that are reported here and that were always conducted in the same order.

1: No-Local-Objects Condition. Directly after the last training session, participants conducted the *no-local-objects condition*. They were positioned at the home location and were asked to navigate to the feeder. Upon reaching it all local objects (i.e., all local landmark information) disappeared, while the global landmarks (i.e., the distal colored columns and baground texture) remained (see Figure 1c). Participants task was to navigate back to the home location. Once they were convinced to have reached the position of their home, they pressed a button on the joypad. This procedure was repeated three times. After the no-local-objects test, participants were given a 5 min. break. If they already went through 2 training session that day, they conducted the fog test (see below) at the next day.

2: Fog Condition. The *fog condition* was identical to a training session but the visibility conditions were altered by introducing fog in the virtual environment: the visibility of the environment decreased with increasing distance. Beyond a distance of 2.0 units only fog but no other information was perceptible. By these means, global landmarks as well as structures such as view axes or corridors arising from obstacle constellations were eliminated from the visual scene. The fog also covered the ground plane such that it provided no optic flow during navigation. In this modified environment participants had to rely only on local objects in their close proximity to find their way between home and feeder. All participants had 20 minutes to complete as many runs as possible. After that time the fog test stopped with the first visit back at home. Unfortunately, data of three participants had to be excluded from the final data set due to technical problems with the software.

2.5 Data Analysis

During the experiment, participants' positions were recorded with a sampling frequency of 5Hz, allowing to reproduce detailed trajectories. In the following we describe the different dependent variables extracted from these trajectories:

1. **Performance.** Navigation performance is measured in runs per minute. A single run is defined as a navigation from home to the feeder or vice versa. For each experimental session we calculated the average number of runs per minute.
2. **Route-similarity.** The route similarity measure describes how conservative or variable participants were with respect to their route choices. High values ($\approx$1) demonstrate that participants repeatedly choose the same route; low values correspond to a high variablity in route choices. To calculate the route similarity, we used a two step method: (1) the raw trail data was

reduced to sequences of numbers; (2) the similarity of these sequences was compared. For the first step, the environment was tessellated into triangles: we reduced each local obstacle to its center point and applied a Delaunay-triangulation to this set of points. A unique number was assigned to each resulting triangle. Now every run was expressed as a sequence of numbers corresponding to the triangles crossed. To compare these sequences, we used an algorithm originally developed to compare genetic sequences [18]. In each case two single sequences are compaired. The basic principle is to find the number of matches and relate that to the total length of the sequences (for details [19]). A complete match results in a value of 1.0. For each participant these comparisons were done separately for all outbound and inbound routes. In the following we present the mean similarity of all comparisons of runs performed in one session.

3. **Change in performance.** In order to describe navigation performance during the fog test, participants performance (runs per minute) in the fog condition was divided by their performance in the last training session. Equal performance then results in a value of 1.0, increased navigation performance results in values > 1.0 and decreased performance results in values < 1.0.
4. **No-local-object test.** For the homing task in the *No-Local-Objects* test, the following variables were evaluated:
   (a) *Homing error*: Distance between participant's endpoint and actual home position.
   (b) *Distance error*: Air-line distance from start (feeder location) to participant's chosen endpoint compared with the air-line distance from start (feeder location) to the actual home position
   (c) *Angular error*: Angle between the beeline of participants' homing response and the correct direction towards the home location.

3 Results

3.1 Training Phase

Route Learning Performance. 19 out of the total of 20 participants were able to solve the task: they learned to efficiently navigate between the home and the feeder. One participant was removed from the final data set, as he did not reach the learning criterion. This participant also reported to be clueless about the positions of the home and the feeder. For the remaining participants, the time to reach the learning criterion differed: four participants reached it after 5 training sessions, 6 participants after 6 sessions, 2 after 7 sessions, 2 after 8 sessions, and 5 participants needed 9 training session (6.9 sessions on average). The increase in navigation performance was highly significant for both inbound and outbound runs (see Figure 2, paired t-test first vs. last training session, inbound: $t(18) = 14.26; p < .001$; outbound: $t(18) = 10.76; p < .001$). Figure 3 show examples from two participants how route-knowledge evolved with increasing training sessions for two participants. At the end of the training phase, all remaining 19 participants solved the task of navigating between home and

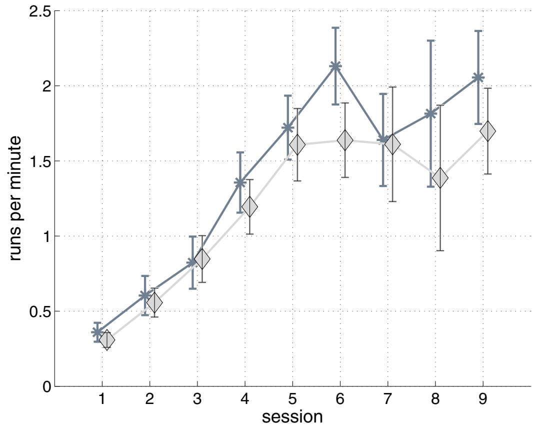


Fig. 2. During the training phase participants' performance increased in number of runs per minute with increasing number of sessions for both, outbound runs (◇), and inbound runs ($*$). Mean values of all participants ± standard error.

the feeder reliably and efficiently. For this result and all other results we did not found any significant gender differences. Since we have small groups (10 female and 11 male) small differences if present are not ascertainable.

Outbound Runs and Inbound Runs. Participants showed better navigation performance (runs/min) on inbound runs as compared to outbound runs (Wilcoxon signed rank test: $p < .01$, see Figure 2). In other words, participants found the way from feeder to home faster than the way from home to feeder. It appears that this difference increases with increasing number of sessions. Note, however, that some particpants reached the learning criterion already after 5 sessions and proceeded to the test phase. In later sessions the number of participants therefore decreases which explains the increasing variations in later sessions and which could account for the saturation effect.

Constant and Variable Routes. Analysing the chosen trajectories in the last training session in detail, reveals remarkable inter-individual differences. While some participants were very conservative in their route choices (see right column in Figure 3), other participants showed a large variability in their choices (see left column in Figure 3). The calculated mean route similarity ranged from .19 in case of very variable to 1.0 for constant routes (mean=.67, std=.24). Figure 4 displays the route similarity values for all participants revealing a continuum rather than distinct groups. Navigation performance (runs/min) in the last training session was significantly correlated with route similarity. Specifically, with higher route similarity the navigation performance increased ($r = .47, p < .05$). Neither navigation speed during the last training session, nor the number of sessions needed to reach the performance criterium to enter the test phase significantly correlated with the route similarity values of the last training session

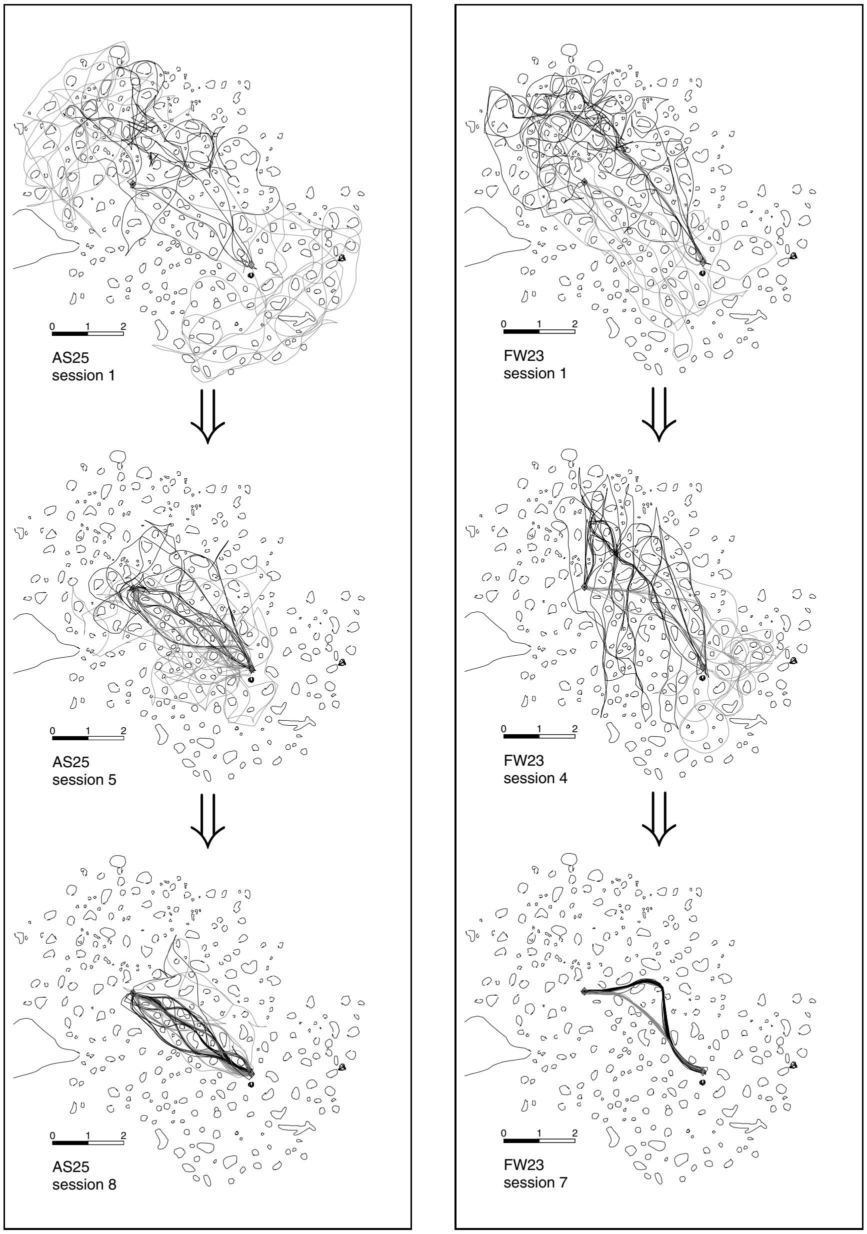


Fig. 3. Evolving route knowledge of two participants. Left column: variable routes with similarity of 0.55 (mean of outbound and inbound runs of the last session, compare to fig. 4 and see text); right column: constant route with similarity 1.0, (mean of outbound and inbound runs of the last session, compare to fig. 4 and see text). *lower left corner* measuring unit, participant, and session number.

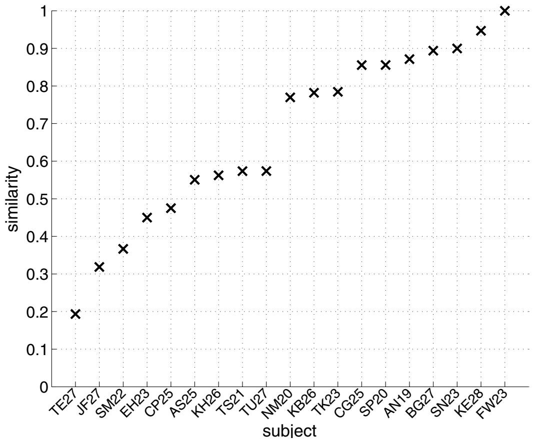


Fig. 4. Route similarity values of all participants of their final session

(correlation navigation speed and route similarity: $r = -.01, p = .97$; correlation number of sessions and route similarity: $r = -.34, p = .16$).

3.2 No-Local-Objects Condition

In the *no-local-objects* condition, all obstacles disappeared after participant reached the feeder. By moving to the estimated position of the home and pressing a button, participants marked the location where they assume the position

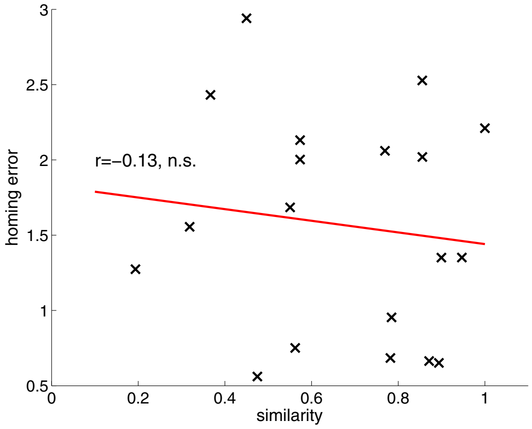


Fig. 5. Navigation without local landmarks: Participants' homing error as a function of their route similarity of the last training session

of their home. On average participants produced a homing error of 1.57 units (std=1.00), an angular error of 16.38 degrees (std=12.17), and a mean distance error of 0.97 units (std=0.76). Together, these results suggest that, in principle, participants could solve the homing task. None of the measures for the homing performance did significantly correlate with the route similarity measure (homing error and route similarity: $r = -.13, p = .61$, see Fig. 5; angular error and route similarity: $r = -.22, p = .36$; distance error and route similarity: $r = -.42, p = .07$). Apparently, participants performance in solving the homing task was independent of whether or not they had established fixed routes during the training phase.

Nevertheless, a closer look at the homing trajectories themself suggests that participants' differed in the strategies they applied to solve the homing task. Figure 6 provides a few examples of homing trajectories. Participants, with low route similarity values (i.e, participants showing a high variability in their choices) show more or less straight inbound routes when homing. Participants with high route similarity values (i.e., participants that established fixed routes during training) generate trajectories that are typically curved, not linear. Moreover, their trajectories were often similar to their habitual routes: the shape of the routes they established during training was roughly reproduced, even if the translational or rotational metric did not fit exactly (see Figure 6). In some cases, established routes were close to the beeline between feeder and home. In such cases it is indistinguishable if the established route is reproduced or if another strategy was used. The same is obviously true for participants showing a high variability in their route choices. While it is not clear how such data could be quantitatively analyzed, Figure 6 demonstrates that some participants with high route similarity values reproduced the form of their habitual routes.

3.3 Fog Condition

In this part of the experiment participants were able to see closely obstacles. Spatial information at larger distances was masked by fog (Fig. 1d). Individual performance (runs per minutes) during the fog test was compared with the performance of the last training session (expressed as change in performance). As expected, most participants show a performance decrease in the fog test (inbound: 14 of 16, outbound: 13 of 16). More interestingly, we found significant correlations between the change in performance and the route similarity in the last training session: participants with low route similarity values show stronger (negative) changes in performance as compared to participants with higher route similarity values (see Figure 7). These correlation were significant for both, inbound runs ($r = .50, < .02$) and outbound runs ($r = .75, p < .001$).

4 Discussion

In this work we presented a navigation experiment investigating human route learning behavior in a complex cluttered virtual environment. In contrast to

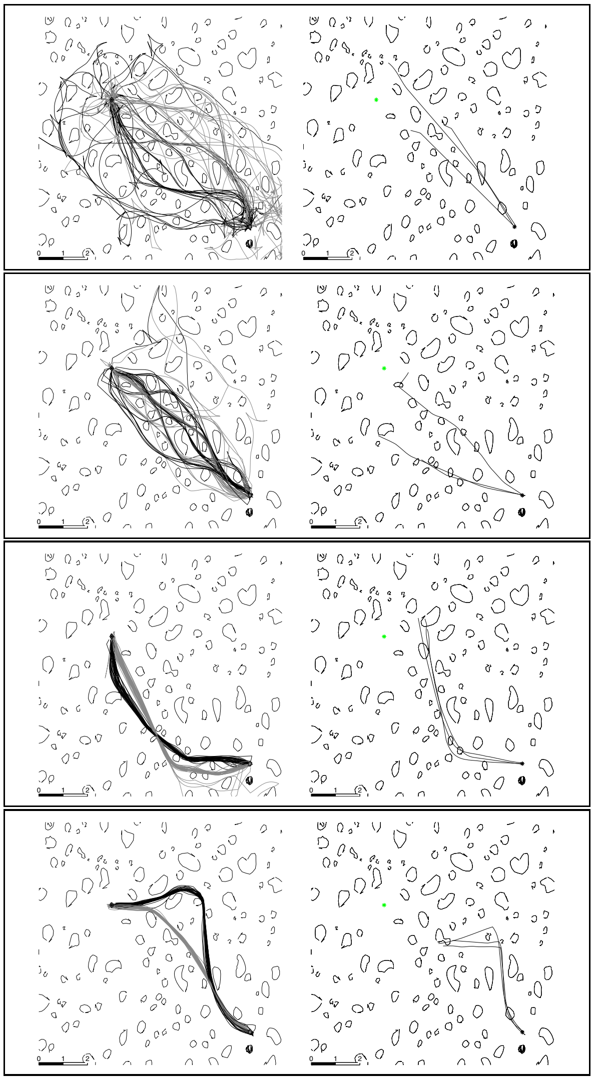


Fig. 6. Four examples of behavior in the last training session (left column) and homing behavior in the *no-local-objects test* (right column). The two top rows show results from a participant with low route similarity values, the lower two rows show examples from a participant with high route similarity values.

most earlier studies on route learning and route knowledge (e.g.,[3,20,21]), the current environment did not feature a road network with predefined places, junctions (decision points), and unique local landmarks. The environment was made up of many similarly shaped objects with identical texture and height that were

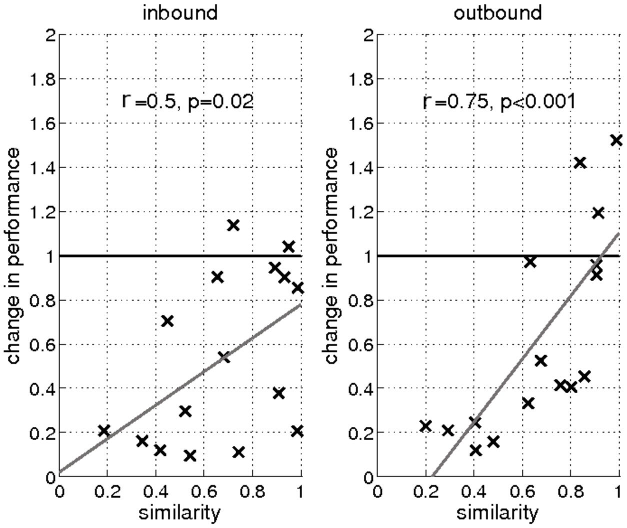


Fig. 7. Participants had to navigate in a foggy VR environment, so only objects in close proximity were visible. Given is the change in performance (runs per minute) of all participants in fog compared to the similarity of routes established in the last training session.

non-uniformly distributed about a large open space. In addition to these local objects, four distal unique landmarks provided global references. Specific interest concerned the question if navigators were able to learn their way between two locations in such an environment. Furthermore, we were interested if all participants used similar or identical route learning strategies (for example: do navigators establish fixed routes or do they rather learn the global layout of the environment and chose between different similarly efficient routes). In the experiment, participants were trained for several session until they were able to efficiently navigate between two locations, the *home* and the *target*. After reaching a learning criterion they entered the test phase, during which the availability of spatial information was systematically varied to investigate which spatial information (local or global) participants used to solve the task.

All but one participant reached the learning criterion after a maximum of 9 training sessions. Navigation performance (measured as runs per minute) clearly increased with the number of training session (see Figure 2). This demonstrates that participants were able to efficiently and reliably navigate in complex cluttered environments lacking predefined places, road networks, and local landmark information that is usually provided by unique objects (e.g., large red house) at decision points or road crossings. Comparisons of navigation performance over the entire training phase revealed differences for outbound runs (home to target) and inbound runs (target to home): specifically, participants found their way faster on inbound runs. This could be explained by the specific significance of the home location, which may result from the fact that each training session started at the home/nest. In central place foragers, like the desert ants, the importance of the nest and its close surrounding is well documented [22]. An alternative explanation for this effect is that the local surrounding of nest and feeder were different (i.e. the spatial distribution of the surrounding obstacles): the nest, for example, was positioned at a larger open space, surrounded by fewer objects, as compared to the feeder. By these means, the nest position might have been recognized from larger distances, hence resulting in an increased performance. Further experiments will have to show whether semantic or spatial (configurational) effects were responsible for the described effect.

The most important result of the training phase is that participants greatly differed with respect to their route choices: using a novel method to compare trajectories (see Section 2.5) we obtained descriptions of the similarity of the traveled paths during the last training session. While some participants were very conservative, selecting the same outbound path and the same inbound path on most runs, others showed a high variability, navigating along many different paths (for examples, see Figure 3). Participants' route similarity values of the last training session were correlated with their navigation performance during that session: participants that established fixed routes during training showed better navigation performance than participants that showed higher variabilities in their route choices. How can these inter-individual differences in route similarity and navigation performance be interpreted? Did different participants employ different navigation or learning strategies, relying on different spatial information?

Results from the test phase in which the availability of different types of spatial information was systematically manipulated allowed for first answers: In the *fog* condition (see Figure 1d) only obstacles in close proximity were visible. By these means, global spatial information was erased (i.e., distal global landmarks and spatial information, emerging by lined-up obstacles such as visual gate-ways or corridors). We observed correlations between participants' route similarity values and their performance in the *fog* condition. Specifically, individuals showing a high variability in route choices showed a clear reduction of navigation performance during the fog condition as compared to the last training session. Individuals with a low variability in route choices, on the other hand, were largely unaffected by the fog. These results suggest that participants with variable route choice behavior strongly relied on distal or global spatial information, while participants exhibiting conservative route choice behavior rather relied on proximal spatial information, as provided by the close-by obstacles or obstacle configurations. A straight forward assumption is that the latter group learned local views (obstacle configurations) and corresponding movement decisions (c.f. [23]) during the training phase that were also available also during the *fog condition*. In other words, route knowledge for these participants would be best described as a sequence of recognition triggered responses [1,3].

If, in fact, participants exhibiting conservative route choice behavior relied on recognition triggered responses, and participants showing variable route choice behavior primarily relied on distal, global spatial information or knowledge, the following behavior had to be predicted for the *no local obstacle condition*: if all local obstacles dissappear after reaching the *feeder* and only the distal global landmarks remained, returning to the home should be impossible for participants relying on recognition triggered responses only. Participants relying on global information, on the other hand, should be able to solve the task. In contrast to these predictions, results demonstrate that all participants were able to solve the task with a certain accuracy (see Figures 5 and 6). Furthermore, virtually no correlation (r=-.13) was found between participants' route similarities in the last training session and their homing performance in the *no local obstacle condition*. This disproves the explanations given above: apparently participants showing conservative route choice behavior did not solely rely on stored views and remembered movement decisions (i.e., recognition triggered responses), but had additional spatial knowledge allowing them to solve the homing task. A detailed inspection of their homing trajectories revealed that some participants reproduced the overall form of their habitual routes from the last training session (see Figure 6). There are two ways of achieving such behavior: (1) participants learned a motor program during training that was replayed during the *no local obstacle condition*, or (2) they possessed a metric representation of the established routes. While this experiment does not allow distinguishing between these alternatives, informal interviews participants after the experiment support the latter explanation.

Taken together, we have shown that participants could learn to efficiently navigate between two locations in a complex cluttered virtual environment, lacking predefined places, decision points, and road networks. In such unstructured environments a route is best described as a sequence places defined by views or object configurations [3], rather than as a sequence of places defined by unique single objects. Analyzing participants' navigation behavior, we could show strong interindividual differences that could be related to different navigation or orientation strategies taking different kind of spatial information into account. Specifically, participants showing a high variability in their route choices depended on distal spatial information, suggesting that they learned global directions and distances between relevant locations. Participants who established fixed routes instead relied on proximal obstacles to guide their movements. However, even if such local spatial information was not available, some were able to reproduce the overall form of their preferred paths. Apparently they learned more than reflex-like recognition triggered responses during training, presumably generating a metric representation of their preferred paths. These results are not in line with the dominant *landmark* to *route* to *survey knowledge* framework of spatial knowledge acquisition [6], stating that survey knowledge emerges not until route knowledge is established. Apparently some participants were able to learn about distances and directions in the environment without first establishing route knowledge (c.f. [5]). The fact that participants' route similarities of their last training session did not fall into two distinct clusters but constituted a continuum, furthermore, suggests that the two learning strategies sketched above are not exclusive but complementary, existing in parallel (c.f. [4]), and that different participants weighted them differently. It is highly likely that these weights are adopted during the course of learning.

Further research is needed to answer questions arising from this exploratory study. For example, what triggers the usage of which strategy? How are the strategies related to each other? And, how is metric information entangled with the strategies applied?

References

1. van Janzen, G., Turennout, M.: Selective neural representation of objects relevant for navigation. Nature Neuroscience 7(6), 572–574 (2004)
2. Kohler, M., Wehner, R.: Idiosyncratic route-based memories in desert ants, Melophorus bagoti: How do they interact with path-integration vectors? Neurobiol. Learn. Mem. 83, 1–12 (2005)
3. Mallot, H., Gillner, S.: Route navigation without place recognition. what is recognized in recognition triggered responses? Perception 29, 43–55 (2000)
4. Aginsky, V., Harris, C., Rensink, R., Beusmans, J.: Two strategies for learning a route in a driving simulator. Journal of Environmental Psychology 17, 317–331 (1997)
5. Ishikawa, T., Montello, D.: Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places. Cognitive Psychology 52, 93–129 (2006)
6. Siegel, A., White, S.: The development of spatial representations of large-scale environments. Advances in child development and behavior 10, 9–55 (1975)
7. Restle, F.: Discrimination cues in mazes: A resolution of the 'place-vs-response' question. Psychological Review 64(4), 217–228 (1957)
8. Leonard, B., McNaughton, B.: Spatial representation in the rat conceptual behavioural and neurophysiological perspectives. In: Kessner, R., Olton, D.S. (eds.) Comparative Cognition and Neuroscience: Neurobiology of Comparative Cognition. Hillsdale, New Jersey (1990)
9. Taylor, H., Naylor, S., Chechile, N.: Goal-specific influences on the representation of spatial perspective. Memory and Cognition 27, 309–319 (1999)
10. Trullier, O., Wiener, S., Berthoz, A., Meyer, J.A.: Biologically based artificial navigation systems: review and prospects. Progress in Neurobiology 51(5), 483–544 (1997)
11. Kuipers, B.: The spatial semantic hierarchy. Artificial Intelligence 119, 191–233 (2000)
12. Gillner, S., Mallot, H.: Navigation and acquisition of spatial knowledge in a virtual maze. Journal of Cognitive Neuroscience 10, 445–463 (1998)
13. Hölscher, C., Meilinger, T., Vrachliotis, G., Brösamle, M., Knauff, M.: Up the down staircase: Wayfinding strategies and multi-level buildings. Journal of Environmental Psychology 26(4), 284–299 (2006)

14. Wehner, R., Boyer, M., Loertscher, F., Sommer, S., Menzi, U.: Ant navigation: One-way routes rather than maps. Current Biology 16, 75–79 (2006)
15. Graham, P., Collett, T.: Bi-directional route learning in wood ants. Journal of Experimental Biology 209, 3677–3684 (2006)
16. Judd, S., Collett, T.S.: Multiple stored views and landmark guidance in ants. Nature 392, 710–714 (1998)
17. Diwadkar, V., McNamara, T.: Viewpoint dependence in scene recognition. Psychological Science 8, 302–307 (1997)
18. Needleman, S., Wunsch, C.: A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. 48, 443–453 (1970)
19. Basten, K., Mallot, H.: Building blocks for trail analysis (2008)
20. Gaunet, F., Vidal, M., Kemeny, A., Berthoz, A.: Active, passive and snapshot exploration in a virtual environment: influence on scene memory, reorientation and path memory. Cognitive Brain Research 11(3), 409–420 (2001)
21. Munzer, S., Zimmer, H., Schwalm, M., Baus, J., Aslan, I.: Computer-assisted navigation and the acquisition of route and survey knowledge. Journal of Environmental Psychology 26(4), 300–308 (2006)
22. Bisch-Knaden, S., Wehner, R.: Landmark memories are more robust when acquired at the nest site than en route: experiments in desert ants. Naturwissenschaften 90, 127–130 (2003)
23. Christou, C.G., Bülthoff, H.H.: View dependence in scene recognition after active learning. Memory and Cognition 27(6), 996–1007 (1999)

Manuskript 3

Trajectory discretisation and sequence analysis for wayfinding experiments in spatial cognition

Ziele, Hauptergebnisse und Eigenanteil an der Publikation/dem Manuskript

Eine gezielte Bewegung im Raum ist das Ergebnis eines oder mehrerer kombinierter Navigationsmechanismen. Diese Navigationmechanismen wurden in zahlreichen Versuchen mit Tieren und Menschen erforscht. Dabei wurden oftmals (Lauf-)Spuren in 2-dimensionalen oder 3-dimensionalen Koordinaten aufgezeichnet. Diese Spuren sind das Ergebnis der Wechselwirkung des einzelnen Individuums mit der Versuchsumgebung. Das bedeutet auch, dass die Struktur der Umgebung Einfluß auf das Orientierungsverhalten hat. Durch den Versuchsaufbau beziehungsweise die Versuchsumgebung können beispielsweise feste Wege (z.B. Labyrinth oder Straßennetz) vorgegeben sein und damit sind oft auch die für die Orientierung wichtigen Entscheidungspunkte (z.B. Kreuzungen) festgelegt. Mit steigender Anzahl möglicher Richtungsentscheidungen an den Entscheidungspunkten (z.B. Kreuzungspunkt von drei, vier oder mehr Wegen) erhöht sich auch die (potenzielle) Variabilität der Laufspuren innerhalb einer Umgebung, da sich die Zahl der möglichen Wege zwischen zwei Orten erhöht.

Ziel des in diesem Kapitel vorgestellten Projekts war die Entwicklung eines Auswerteverfahrens, das es erlaubt, Laufspuren einzelner Individuen, unterschiedlicher Gruppen oder aus unterschiedlichen Versuchsbedingungen statistisch auszuwerten. Dabei sollte die Struktur der Versuchsumgebung berücksichtigt werden, d.h. vor allem die Orte, die Richtungsentscheidungen erlauben.

Das Verfahren ist modular aufgebaut und besteht aus drei Teilen: 1) Reduzierung der Rohdaten zu Sequenzen besuchter Entscheidungspunkte, 2) Bestimmung der Ähnlichkeit dieser Sequenzen zueinander und 3) statistische Analyse der Laufspuren verschiedener Individuen oder aus unterschiedlichen Versuchsbedingungen. Die einzelnen Module können dabei unabhängig voneinander benutzt werden und lassen sich sowohl an den Versuchsaufbau als auch and die Fragestellung des Versuches anpassen.

Die drei Module basieren auf Methoden aus der Geometrie, der DNA Sequenzanalyse und der Statistik, die in diesem Projekt erweitert, angepaßt und neu kombiniert wurden. Das Auswerteverfahren wurde von mir für Tests und zur späteren Verwendung in der Programmierumgebung $Matlab^{TM}$ implementiert. Die in diesem Manuskript verwendeten Beispieldaten wurden im Rahmen des Projektes *Route learning strategies in a virtual cluttered environment* (siehe Manuskript 2) in Zusammenarbeit mit Frau Rebecca Hurlebaus gewonnen. In dem Projekt beschrieben in Manuskript 2 wurde auch der hier vorgestellte Algorithmus in Teilen zur Auswertung der Experimentaldaten verwendet. Das Manuskript wurde zusammen mit Herrn Prof. Hanspeter A. Mallot zur Veröffentlichung vorbereitet.

Referenz:
K. Basten, H.A. Mallot (in prep.), *Trajectory discretisation and sequence analysis for wayfinding experiments in spatial cognition*

Trajectory discretisation and sequence analysis for wayfinding experiments in spatial cognition

Kai Basten and Hanspeter A. Mallot

Cognitive Neuroscience, University of Tübingen, Germany

Abstract

Routes or trajectories recorded with tracking devices are a major type of data in spatial cognition experiments. These data are often interpreted as sequences of discrete decisions taken at possible bifurcation points of the path. Here we present a framework for the analysis of trajectory data in open, channelled, and cluttered environments. In a trajectory discretisation step, continuous position data are transformed into a sequence of places or route sections each associated with an individual movement decision. For the resulting sequences, similarity measures are derived both for pairwise comparisons and for comparisons of sets of trajectories obtained, for example, from different conditions or subjects in one experiment. Statistical significance tests for sequence comparison are also presented. As an example, the overall scheme of analysis is applied to data from a human route learning experiment in a virtual, cluttered environment.

Key words: navigation, position tracking, route analysis, spatial behaviour, spatial cognition, trail analysis, wayfinding

---

1. Introduction

Animals and humans show coordinated and goal directed movements in their environments. Places are recognised and routes are planned. The resulting locomotion pattern is influenced not only by the individual's inner variables, such as spatial memory or motivation, but also by external factors. Such external factors include the path geometry and environmental structures, as well as predator threats and other costs associated with individual path segments. The resulting routes are generally conceptualised as discrete sequences of decisions taking into account both internal and external constraints.

Preprint submitted to Elsevier *March 2, 2010*

In recent years, methods have been developed to describe spatial behaviour in large environments with increasing accuracy. Trajectories are recorded in different ways such as simply following the observed individual, by video analysis, or position tracking in virtual reality (VR) devices. Additionally, modern tracking tools providing exact global data are used such as harmonic radar for insect tracking (e.g. Riley et al., 1996; Menzel et al., 2005; Cant et al., 2005), GPS systems for bird flight monitoring (Lau et al., 2006; Vyssotski et al., 2006; Dell'Ariccia et al., 2008) or primate research (Noser and Byrne, 2007; Byrne et al., 2009). At the same time, mathematical models of trajectory generation have been developed that need to be compared with with other models or experimental data (Warren and Fajen, 2004; Graham et al., 2007).

Since the the original measurements generally provide continuous trajectory data, techniques for deriving and comparing discrete sequences are needed. These techniques will not only depend on the trajectories themselves, but will also take into account the environmental structure from which decision points have to be defined. In this paper, we first outline the characteristics of routes with respect to wayfinding decisions; then we present a modular algorithm for the analysis of trajectory data.

Trajectories will be treated as two- or three-dimensional curves $\mathbf{x}(t)$ specifying a position $\mathbf{x}$ for each measurement time t. Alternatively, pose vectors comprising six degrees of freedom for position and orientation might also be considered. A trace $\mathbf{x}(s)$ is a unit speed curve, parameterised by the path length $s(t) = \int_o^t \|\mathbf{x}'(t)\| dt$ (cf. Korn and Korn, 2000, 17.2-3.b). That is, $\mathbf{x}(s_o)$ is the position of the agent after having travelled for a distance s_o irrespective of the speed profile used during this travel.

Recorded trajectories result from various sensory, physiological and cognitive mechanisms which data analysis seeks to recover. In particular, the following variables have to be considered:

1. Motor variables such as linear and rotational acceleration as well as roll of flying agents. In the simplest case, these variables may be specified as a motor program $(v(t), \omega(t), \theta(t))$ where v is translational forward speed, ω rotational speed about the dorsoventral body axis, and θ torsion, or roll. These local variables uniquely specify the resulting trajectory, according to the 'fundamental theorem of differential geometry' (DoCarmo, 1976).

2. Cognitive variables such as goal representations (both in terms of a home vector or landmarks defining a place), recognition of other places and recognition-triggered responses (cf. Trullier et al., 1997). These are the variables which are most relevant in experiments on spatial longterm-memory.
3. Environmental variables such as configurations of obstacles or free space as well as 'channels' or corridors restricting the agent's route choice. Environmental variables may also be costs associated with particular route sections, for example risks to take or effort to exert.

Trajectories in cluttered environments will often be determined by a goal position or goal direction and one or multiple obstacles that have to be avoided. An important class of trajectory models is built on the idea of attraction by the goal and repulsion from obstacles or repellers (Arbib and House, 1987; Schöner et al., 1995; Fajen et al., 2003; Gerstmayr et al., 2008). For example, in the Fajen et Al. (2003) model, heading Φ is modeled as a dynamical system of the type $\ddot{\Phi} = -b\dot{\Phi} - a(\Phi - \Phi_o, d_o) - \sum_i r(\Phi - \Phi_i, d_i)$. Here, b is a stiffness parameter, a is an attraction function depending on the deviation between the current heading and the goal direction Φ_o and the goal distance d_o, and r is a repellor function depending on the deviation between current heading and the direction and distance of obstacle i, Φ_i and d_i, respectively. This equation can be fitted to natural trajectories around multiple small obstacles.

While the attractor-repellor theory is very elegant for modelling approaches to individual goals, route behaviour is generally considered to be a sequence of multiple such approaches. Cognitive models of route behaviour are usually build on the notion of 'direction' (O'Keefe and Nadel, 1978) or recognition-triggered response (Trullier et al., 1997). In the simplest case, a recognition-triggered response is an association of a recognised place P and an action A to take from there ($< PA >$ association). For honeybees, a model of spatial memory ('the landmark map') has been suggested consisting of a collection of independent associations $< P_i, A_i >$ (Fig. 1 a; Cartwright and Collett, 1987). A discrete set of places P_i is recognised based on the snapshot or local view seen from this place. To each place is associated an action A_i, which in the honeybee example is a vector pointing from P_i to the hive, specified with respect to the stored snapshot or the honeybees's sun compass. A similar approach with $< PA >$ associations is pursued by (Ohashi et al., 2007) in bumblebee traplining. In this paper, the discrete structure of the environment is given by the feeder layout and the trajectory is measured in terms of a discrete sequence of visits of these feeders.

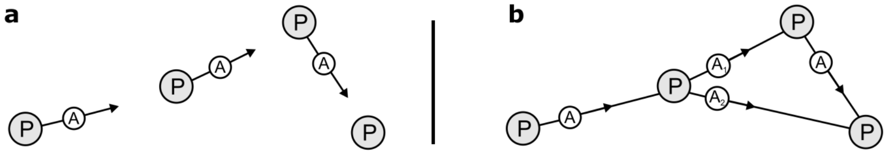


Figure 1: Two models of route knowledge. (**a**) place-recognition triggered response, that is, $< PA >$ association; (**b**) a three-way association $< PAP' >$ implies an expectation about the next place to reach and allows multiple associations. P: place, A: action

Most models of routes and route knowledge assume three-way associations of the type $< PAP' >$ (Fig. 1 b; O'Keefe and Nadel, 1978; Kuipers, 1978, 2000). Thus, if action A is carried out at $< P >$, the agent already has an expectation about the next place to encounter. Although the three-way associations may be stored independently, for example, as elements of a set, they define a graph structure made up of the places P as nodes and the associations $< PAP' >$ as links. This structure allows to store multiple associations for each starting point as in $\{< PA_1P' >, < PA_2, P'' >\}$ since the different actions can be distinguished by the different expected outcomes. The actions themselves may be motor programs ("control laws" according to Kuipers, 2000), view-based homing steps (Franz et al., 1998) or other.

Both two-way and three-way associations will result in sequences of the type $\mathcal{R} = (Start, A_o, P_1, A_1, P_2, A_2, ..., Goal)$ which we suggest as the basic route model. If we consider actions to be implicitly defined by their respective start and goal location, the route model reduces to the simpler sequence $\mathcal{R}^* = (Start, P_1, P_2, ..., Goal)$.

The task of data analysis, then, is to recover this sequence from continuous trajectory data and to establish techniques for quantifying similarities and differences between pairs and sets of routes. We will show in this paper how the geometrical layout of the navigational environment can be used to discretize natural trajectories.

Several methods have been developed to describe and evaluate continuous trajectories without considering the underlying sequences of decision points. Local features describing important properties of a trajectory can be derived from differential geometry, for example velocity, instantaneous turning angles (curvature), or straightness of the routes (Müller and Wehner, 1988; Benhamou et al., 2003; Benhamou, 2004; Bartumeus et al., 2008). These features are independent of the routes position and orientation in space and do not take into account the structure of the environment. As a descriptive method, density plots have been used, which visualize the spatial distribution of trajectories in an environment by accumulating the number and duration of visits of cells in a regular grid (e.g. Collett et al., 1986; Åkesson and Wehner, 2002; Bisch-Knaden and Wehner, 2003; Merkle et al., 2006). A global characteristic of trajectories has been derived from the perpendicular deviation of the trajectory from the beeline (start to end point) at one (Heusser and Wehner, 2002) or more discrete points (Kohler and Wehner, 2005). In this case, a trajectory is reduced to a sequence of local distance measurements $\mathcal{R}' = (d_1, d_2, ...)$ relative to the beeline. Instead of discrete sampling, the difference between two routes can also be described continuously by the area enclosed between the routes (Meade et al., 2005).

In this paper we present a method for the analysis of trails in structured environments using decision points derived from the environmental structure. We describe a method consisting of three consecutive modules that can be used subsequently or for themselves. The method will thus be useful for a wide range of wayfinding experiments. The algorithm combines and adapts established methods and consits of three basic modules: *i*) the discretisation of trajectories to place sequences $\mathcal{R}^*$ by identifying decision points, *ii*) a modified amino acid sequence comparison algorithm to describe similarity of trajectories and *iii*) a Monte Carlo permutation test for statistical analysis.

2. Methods: Three modules for trail analysis

Module 1: discretisation of raw trajectory data

In this first step of analysis, raw data represented as two or three dimensional coordinates, are transformed into sequences of visited decision points. For this purpose we divide the experimental environment into discrete locations. The goal in this step is to define places, where decisions for wayfinding are made. How to do this in different environments is elaborated with examples in the following.

In the simplest case, foraging animals are visiting fixed feeders (e.g. Thomson et al., 1997; Thiele and Winter, 2005; Winter and Stich, 2005). Here, data are already in sequential form $(P_1, P_2, ..., P_n)$ and it suffices to assign a unique number to each feeder to distinguish it them from all others

(Fig.2a). In an open environment, the next target position may be freely chosen without restrictions resulting from corridors or obstacles.

In urban environments (e.g. Aginsky et al., 1997) decision points are given by the junctions and squares of the city map, and are again readily associated with an integer number (Fig. 2b). This also the case for common laboratory mazes such as iterated Y- or T-mazes (Tolman, 1948; Gillner and Mallot, 1998; Wiener and Mallot, 2003; Wolbers and Büchel, 2005), see Fig. 2c. For each trajectory, a sequence of visited decision points can be extracted. Each decision point allows two or more possible decisions according to the local structure.

In irregular, cluttered environments, where no clear junctions can be defined, geometrical tessellations can be used to segment the environment into decision points or areas. Consider a a place surrounded by three obstacles and three passages between them. After entering from one side, a binary decision will lead the agent to continue its path either left or right. Geometrically, this situation is described by three environmental grid points marking the obstacle locations, a triangle connecting these grid points, and a point within the triangle treated as the decision point proper. The local path structure is a three-star or Y-pattern of path segments passing between the obstacles. Larger environments can be partitioned into sets of such triangles, forming a geometrical, triangular tessellation where each triangle contains to one decision point.

As indicated in Fig. 2c, both the numbering of junctions and the introduction of a tessellation of the ground plane yield the same result when applied to regular mazes. In irregular environments, triangular tessellations provide the most natural way to discretize the environment and to generalize the notion of a decision points. Clearly, such irregular environments occur in outdoor experiments, but have also be used in laboratory studies (Kohler and Wehner, 2005; Hurlebaus et al., 2008). For irregular environments with multiple obstacles, we use the standard Delaunay triangulation from computational geometry (cf. de Berg et al., 2000).

The standard way of defining triangular tessellations based on a set of grid points (obstacles) is the well known Delaunay triangulation (cf. de Berg et al., 2000). It is the most natural triangulation of a set of grid points in the sense that (i) the circumcircle of each triangle of the triangular grid does not contain another grid point, and (ii) the minimal angle occurring in the entire triangulation is maximised. This means that on average, the triangles are as compact as possible.

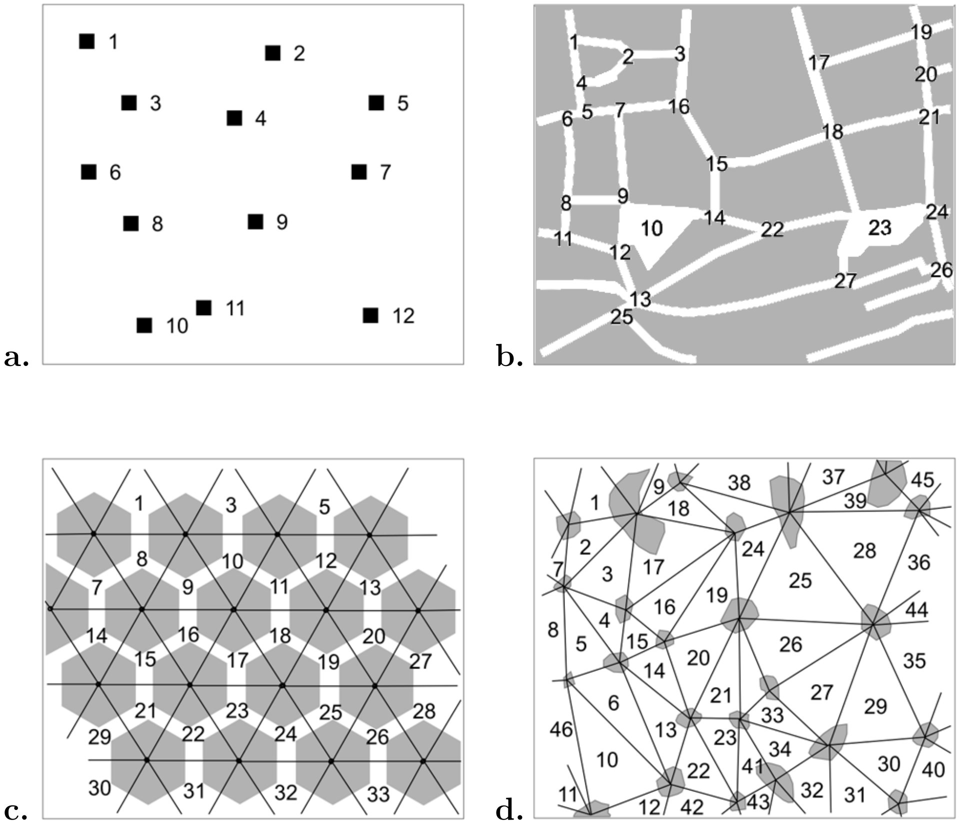


Figure 2: Different setups and environments are structured and numbered, so that trajectories can be discretised into sequences of integer numbers. (*a*) numbered feeder locations, (*b*) in a street raster, binary or multiple decision points are numbered, (*c*) hexagonal environment (iterated Y-maze) with binary decision points. Numbering these points or tessellating the environment with a Delaunay triangulation results in the same structure,(*d*) A cluttered environment without predefined decision points is tessellated using the Delaunay triangulation. A unique number is assigned to each triangle. The discretisation in *a, b* represent locations with multiple possible decisions where to go next, whereas locations in *c, d* lead to binary decisions (leaving a triangle left or right).

The procedure is most obvious for obstacles covering small and convex areas on the ground plane, such as trees in a forest or tussocks in a savanna. Grid points may be generated automatically as the centres of gravity of the obstacle areas, or defined by hand (Fig. 2d). In general, only one point should be assigned to each obstacle. If two or more points are assigned to one obstacle, the triangle sides connecting such points will no longer correspond to a possible passage way. The corresponding triangles will therefore become parts of a corridor-like structure and will no longer contain decision points in the sense of our route model.

As pointed out before, the triangulation guarantees that all decision points are binary, since the surrounding triangle defines a local Y-structure of paths. This may become problematic if four or more grid points fall on one circle, as is the case for an ideal square street crossing. The algorithm will divide the square pattern between the four corners into two triangles by randomly choosing one of the two possible diagonal edges. The crossing will then be represented by two binary decision points corresponding to the two triangles. While this is somewhat arbitrary, it is not a common problem in cluttered environments and will therefore be neglected in further discussion.

Once the triangulation is achieved, a unique ID number is assigned to each triangle (decision area). Thus, even in cluttered environments a trajectory can be reduced to a sequence of visited decision areas satisfying the route description given by $\mathcal{R}*$ (Fig. 2c,d).

In summary, in this first step of our concept for trail analysis we extract decision points from the experimental environment, and massively reduce raw trajectory data into small sequences of integer numbers, that is, decision points $\mathcal{R}^* = (Start, P_1, P_2, ..., Goal)$. Information about speed or time spent at a specific location or in one triangle (i.e. a particular area) is not included in the sequences. An extension regarding this point will be discussed later on in part 4.

Module 2: similarity of sequences

In the first module we showed how to reduce raw trajectory data into small integer sequences. In the second step these sequences are compared in a pairwise fashion. We will present a comparison value for pairs of sequences, ranging between 1 and 0 with 1 meaning 100% similarity or equality .

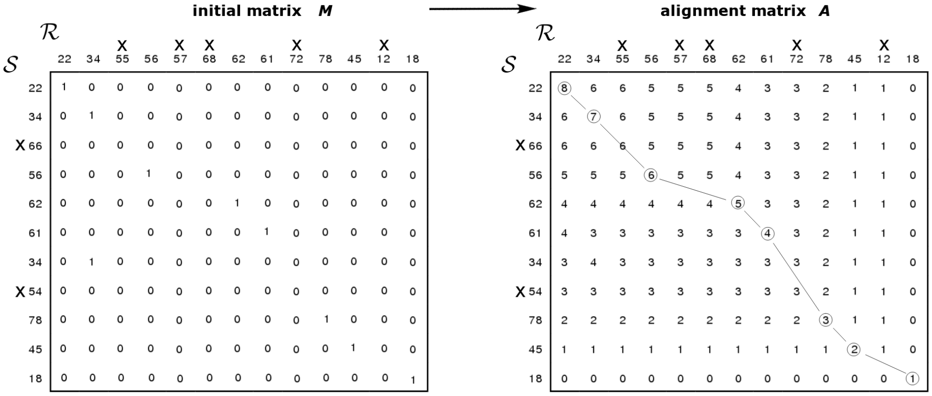
For the analysis of similarity we adapt the method described by Needleman and Wunsch (1970) for the search of similarities in amino acid sequences. The input for this algorithm is given by sequences of integer numbers, in our case, ID numbers of visited feeders, street junctions, or entered triangles from the Delaunay triangulation. Each single test compares two sequences $\mathcal{R} = (r_1, r_2, ..., r_m)$ and $\mathcal{S} = (s_1, s_2, ..., s_n)$ at a time, where r_i is the number of the i-th place visited in route $\mathcal{R}$ and s_j is the number of the j-th place visited in route $\mathcal{S}$. To determine the maximum alignment of these sequences, all possible pairs of numbers are represented in an initial $\mathcal{R} \times \mathcal{S}$ matrix M (Fig. 3, left). All matches of visited decision points in the two sequences are marked by 1, else the value 0 is assigned.

$$M_{i,j} = \begin{cases} 1 \text{ if } r_i = s_j \\ 0 \text{ if } r_i \neq s_j \end{cases} \tag{1}$$

In the next step the largest sequence of matches in matrix M has to be defined without counting matches twice. This is done in an alignment matrix A, which is iteratively filled as follows (cf. Fig 3, right):

```
For   $t = 0, .., \min(r, s) - 1$   Do                                  (2)
        For   $i = 1, ..., r - t$   Do
                $A_{i,s-t} = M_{i,s-t} + \max\{A_{k,l} | k > i, l > s - t\}$
        EndDo
        For   $j = 1, ..., s - t - 1$   Do
                $A_{r-t,j} = M_{r-t,j} + \max\{A_{k,l} | k > r - t, l > j\}$
        EndDo
EndDo
```

That is to say, each cell of matrix A is calculated by adding the value of the corresponding cell of matrix M and the maximum of all cells of matrix A right and down from the current cell. Clearly, the sequence of doing this calculation matters. In the algorithm specified above, calculation starts with $t = 0$ from the last row and column, i.e. with the lower right 'shell' of the matrix. For $t = 1$, the next shell of matrix components to the left and up is considered and so forth until the complete matrix has been processed. The iteration parameter t in eqn. 2 represents the number of the 'shells', i.e. the



initial matrix *M*

$\mathcal{S}$ \ $\mathcal{R}$	22	34	X 55	56	X 57	X 68	62	61	X 72	78	45	X 12	18
22	1	0	0	0	0	0	0	0	0	0	0	0	0
34	0	1	0	0	0	0	0	0	0	0	0	0	0
X 66	0	0	0	0	0	0	0	0	0	0	0	0	0
56	0	0	0	1	0	0	0	0	0	0	0	0	0
62	0	0	0	0	0	0	1	0	0	0	0	0	0
61	0	0	0	0	0	0	0	1	0	0	0	0	0
34	0	1	0	0	0	0	0	0	0	0	0	0	0
X 54	0	0	0	0	0	0	0	0	0	0	0	0	0
78	0	0	0	0	0	0	0	0	0	1	0	0	0
45	0	0	0	0	0	0	0	0	0	0	1	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	1

alignment matrix *A*

$\mathcal{S}$ \ $\mathcal{R}$	22	34	X 55	56	X 57	X 68	62	61	X 72	78	45	X 12	18
22	⑧	6	6	5	5	5	4	3	3	2	1	1	0
34	6	⑦	6	5	5	5	4	3	3	2	1	1	0
X 66	6	6	6	5	5	5	4	3	3	2	1	1	0
56	5	5	5	⑥	5	5	4	3	3	2	1	1	0
62	4	4	4	4	4	4	⑤	3	3	2	1	1	0
61	4	3	3	3	3	3	3	④	3	2	1	1	0
34	3	4	3	3	3	3	3	3	3	2	1	1	0
X 54	3	3	3	3	3	3	3	3	3	2	1	1	0
78	2	2	2	2	2	2	2	2	2	③	1	1	0
45	1	1	1	1	1	1	1	1	1	1	②	1	0
18	0	0	0	0	0	0	0	0	0	0	0	0	①

Figure 3: Comparison of two sequences $\mathcal{R}$ and $\mathcal{S}$ according to equation 2. In the initial matrix M, for all matches of r_i and s_j a 1 is assigned to the according matrix field. In a second step the largest sequence of matches is evaluated in the alignment matrix A. The similarity of these example sequences is $V = 0.615$. All numbers without an equivalent in the other sequence are marked with an X.

shortest distance of the current cell from the right or bottom margin of the matrix. As a result, cell $A_{1,1}$ will hold the maximum number of matches between the two routes. From this, the similarity measure V is calculated as:

$$V = A_{1,1}/\max\{r, s\}. \tag{3}$$

The similarity measure V will produce reasonable results not only for simple routes but also in general cases with various types of loops. This is not true for simpler algorithms such as counting the matches irrespective of their sequence of occurring. To see this, we denote this simpler measure as $\#M$ and define it as $\#M := \sum_{i,j} M_{i,j}$. Since $\#M$ does not take into account the sequence of matches, comparison of inverted routes will yield maximal similarity. Another problem occurs in place repetitions, for example if one route contains an excursion starting and ending in the same place. This place will be counted twice in $\#M$, yielding an overestimate in similarity. The problem is even more apparent in 'double repetitions' if two identical routes loop through a sequence twice ($\mathcal{R} = \mathcal{S} = \{1, 2, 3, 1, 2, 3\}$). In this example, $\#M = 12$, i.e. more than the length of the routes. $A_{1,1}$ will yield 6, resulting in the correct similarity $V = 1$. For further examples of the behaviour of the measures $\#M$, $A_{1,1}$, and V see Table 1.

Table 1: Similarity measures for complex route comparisons.

reference route	$\mathcal{R}$	$= \{1, 2, 3, 4, 5, 6, 7, 8\}$
insertion:	$\mathcal{S}_1$	$= \{1, 2, 3, 4, \mathbf{11}, \mathbf{12}, \mathbf{13}, 4, 5, 6, 7, 8\}$
		$\#M = 9,\ \ A_{1,1} = 8,\ \ V = 0.727$
repetition:	$\mathcal{S}_2$	$= \{1, 2, 3, 4, 5, \mathbf{3}, \mathbf{4}, \mathbf{5}, 6, 7, 8\}$
		$\#M = 11,\ \ A_{1,1} = 8,\ \ V = 0.727$
inversion:	$\mathcal{S}_3$	$= \{1, 2, 3, \mathbf{6}, \mathbf{5}, \mathbf{4}, 7, 8\}$
		$\#M = 8,\ \ A_{1,1} = 6,\ \ V = 0.750$

Module 3: statistical comparison of sets of trajectories

The third part of our modular framework tackles the problem of statistical evaluation. So far, similarity values for pairs of trajectories have been determined. Now, comparison of sets of trajectories obtained from different animals, groups, or under different experimental conditions will be considered. These sets of trajectory data are statistically tested following the question, whether a set of routes $R = (\mathcal{R}_1, \mathcal{R}_2, ..., \mathcal{R}_n)$ differs significantly from a second set $S = (\mathcal{S}_1, \mathcal{S}_2, ..., \mathcal{S}_m)$, regarding the variations within the sets R and S.

The basic idea of the algorithm is to divide the union set $R \cup S$ randomly into two subsets of sizes n and m and compare the set similarities of such random groups to the similarity between the original groups. More formally, similarities between all routes from R and S can be expressed in a matrix, to which a general method of matrix comparison, called Mantel test, will be applied (Mantel, 1967). The Mantel test was originally developed in epidemiology where it is used to investigate spatial and temporal clustering of disease.

For the Mantel test, two matrices are used. First, a distance matrix X is defined containing the pairwise similarity values V (cf. equation 3) for all routes in the sets R, S. It is a $(n+m)^2$-matrix defined by

$$\begin{aligned}
X_{i,j} &= V(R_i, R_j) - \delta_{ij}, && i,j = 1,\dots,n \\
X_{n+i,j} &= V(S_i, R_j), && i = 1,\dots,m; \quad j = 1,\dots,n \\
X_{i,n+j} &= V(R_i, S_j), && i = 1,\dots,n; \quad j = 1,\dots,m \\
X_{n+i,n+j} &= V(S_i, S_j) - \delta_{i,j}, && i,j = 1,\dots,m,
\end{aligned}$$

where n,m are the numbers of routes contained in R, S, respectively. The variable $\delta_{i,j}$ takes the value 1 for $i = j$ and 0 otherwise. It ensures, that trivial comparisons of the form $V(R_k, R_k) = 1$, i.e. comparisons of a route with itself, are excluded from the analysis. Such comparisons might otherwise bias the test (cf. Schnell et al., 1985; Harvey et al., 1988).

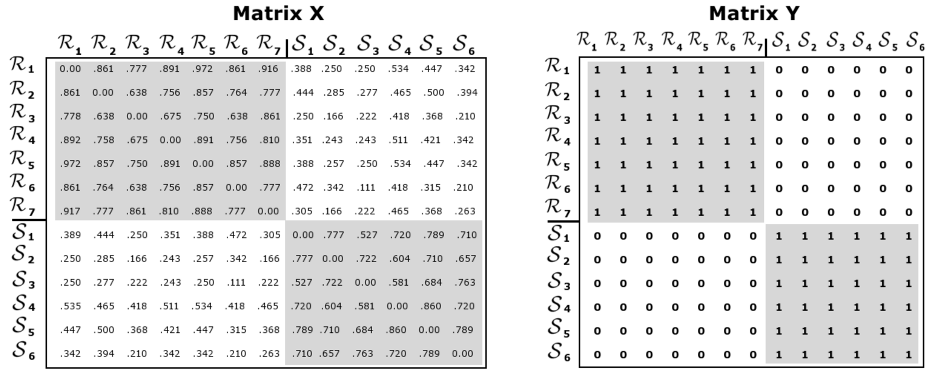
The second matrix Y is a $(n + m)^2$ design matrix denoting whether or not two routes belong to the same set. Thus, it consists of four blocks of components taking value 1 for within-set comparisons and value 0 for across-set comparisons (Fig. 4).

The test has been generalised to cases with two continuous similarity measures, a dependent (i.e., experimentally measured) one such as V and an independent one, D, say, derived from the experimental design. So far, we used $D = 1$ for in-group comparisons and $D = 0$ for between-group comparisons. Examples for continuous D values include difference of training times, genetic relationship, dominance hierarchy, etc. (Schnell et al., 1985). The two matrices are filled by setting the respective V and D values for each pair of individual routes.

Statistical analysis of the resulting matrices is based on a permutation test determining whether the observed difference between the samples is large enough to reject the null hypothesis, which states that both sets of trajectories are drawn from the same distribution. First, the products of corresponding elements of the two matrices X and Y are summed up for all rows i and columns j:

$$U_o = \sum_i \sum_j X_{i,j} Y_{i,j} \tag{4}$$

The resulting value U_o serves as reference value for the following permutation test. For this, the columns j of matrix X are permuted by a scheme π while the rows i are left unchanged:



Matrix X

	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	$\mathcal{R}_4$	$\mathcal{R}_5$	$\mathcal{R}_6$	$\mathcal{R}_7$	$\mathcal{S}_1$	$\mathcal{S}_2$	$\mathcal{S}_3$	$\mathcal{S}_4$	$\mathcal{S}_5$	$\mathcal{S}_6$
$\mathcal{R}_1$	0.00	.861	.777	.891	.972	.861	.916	.388	.250	.250	.534	.447	.342
$\mathcal{R}_2$	.861	0.00	.638	.756	.857	.764	.777	.444	.285	.277	.465	.500	.394
$\mathcal{R}_3$	.778	.638	0.00	.675	.750	.638	.861	.250	.166	.222	.418	.368	.210
$\mathcal{R}_4$	.892	.758	.675	0.00	.891	.756	.810	.351	.243	.243	.511	.421	.342
$\mathcal{R}_5$	.972	.857	.750	.891	0.00	.857	.888	.388	.257	.250	.534	.447	.342
$\mathcal{R}_6$	.861	.764	.638	.756	.857	0.00	.777	.472	.342	.111	.418	.315	.210
$\mathcal{R}_7$	.917	.777	.861	.810	.888	.777	0.00	.305	.166	.222	.465	.368	.263
$\mathcal{S}_1$	.389	.444	.250	.351	.388	.472	.305	0.00	.777	.527	.720	.789	.710
$\mathcal{S}_2$	.250	.285	.166	.243	.257	.342	.166	.777	0.00	.722	.604	.710	.657
$\mathcal{S}_3$	.250	.277	.222	.243	.250	.111	.222	.527	.722	0.00	.581	.684	.763
$\mathcal{S}_4$	.535	.465	.418	.511	.534	.418	.465	.720	.604	.581	0.00	.860	.720
$\mathcal{S}_5$	.447	.500	.368	.421	.447	.315	.368	.789	.710	.684	.860	0.00	.789
$\mathcal{S}_6$	.342	.394	.210	.342	.342	.210	.263	.710	.657	.763	.720	.789	0.00

Matrix Y

	$\mathcal{R}_1$	$\mathcal{R}_2$	$\mathcal{R}_3$	$\mathcal{R}_4$	$\mathcal{R}_5$	$\mathcal{R}_6$	$\mathcal{R}_7$	$\mathcal{S}_1$	$\mathcal{S}_2$	$\mathcal{S}_3$	$\mathcal{S}_4$	$\mathcal{S}_5$	$\mathcal{S}_6$
$\mathcal{R}_1$	1	1	1	1	1	1	1	0	0	0	0	0	0
$\mathcal{R}_2$	1	1	1	1	1	1	1	0	0	0	0	0	0
$\mathcal{R}_3$	1	1	1	1	1	1	1	0	0	0	0	0	0
$\mathcal{R}_4$	1	1	1	1	1	1	1	0	0	0	0	0	0
$\mathcal{R}_5$	1	1	1	1	1	1	1	0	0	0	0	0	0
$\mathcal{R}_6$	1	1	1	1	1	1	1	0	0	0	0	0	0
$\mathcal{R}_7$	1	1	1	1	1	1	1	0	0	0	0	0	0
$\mathcal{S}_1$	0	0	0	0	0	0	0	1	1	1	1	1	1
$\mathcal{S}_2$	0	0	0	0	0	0	0	1	1	1	1	1	1
$\mathcal{S}_3$	0	0	0	0	0	0	0	1	1	1	1	1	1
$\mathcal{S}_4$	0	0	0	0	0	0	0	1	1	1	1	1	1
$\mathcal{S}_5$	0	0	0	0	0	0	0	1	1	1	1	1	1
$\mathcal{S}_6$	0	0	0	0	0	0	0	1	1	1	1	1	1

Figure 4: For statistical analysis of trails measured from two different individuals we built a similarity matrix (left) with all values V from pairwise comparisons. Accordingly we built the design matrix Y indicating within and between-group comparisons by values 1 and 0, respectively.

$$U' = \sum_i \sum_j X_{i,\pi(j)} Y_{i,j}. \tag{5}$$

From the large set of possible permutations ($(n+m)!$), we chose a random sample of permutations and calculate the corresponding sample of U' values. From this 'Monte Carlo' sample, a significance level α is calculated by the relative number of cases where U' is larger than the reference value U_o:

$$\alpha = \frac{1}{n_\pi} |\{{U'}_{n_\pi} > U_o\}|, \tag{6}$$

where n_π is the size of the Monte Carlo sample. Significance will depend on n_π but will approach a limit value for large sample sizes. Therefore, the sample size should be increased to the point where further increase does not substantially affect the resulting estimate of α.

A value α gives the significance level without further calculations. That is to say, for $\alpha < 0.05$ the null hypothesis can be rejected on the 0.05-level. The permutation is non-parametric in the sense that the distribution of the V-values does not matter (Good, 1994).

3. Case Study

The presented method is suitable for the analysis of routes from spatial cognition experiments, such as finding the way between two or more locations. Examples include the way from the nest to a feeding site in animals or the route between home and office in humans. The process of learning routes was investigated in a related study, where naive participants had to learn a way between two target locations in a cluttered virtual environment (Hurlebaus et al., 2008). The virtual environment was presented on a standard 19" computer monitor and participants controlled their translation and rotation using a joypad. The environment contained 291 columnar objects with irregular foot-prints, equally textured surface and equal height (2 meters). Additionally, four large, differently coloured poles were placed outside of the environment, at about 80 meters distance, providing a global frame of reference comparable to a compass. The target locations were marked with small spheres visible in close proximity only. This setup provided no predefined decision points compared to mazes or street layouts.

Finding a way between the two targets was very challenging for the participants. On average they needed around seven learning sessions of 20 minutes each (two a day) to reach the final performance criteria (> 2 runs per minute, 1 run $=$ navigation from one target to the other). The formation of routes was observed over subsequent trials (Fig. 5). Here, we show exemplarily for one participant how to apply our method to this experiment in order to address the following questions:

1. How conservative are participants in their route choices and do they establish habitual routes?
2. Does route formation differ for the outbound (start to goal) and inbound (goal to start) directions?

Figure 5 shows route formation during learning sessions for one subject. Two effects can be seen, route convergence and direction dependence. While initially, search patterns and large variations prevail, the routes get more and more fixated for the later sessions. Finally, two slightly different, fixed routes emerge, one for each direction of navigation.

We applied the route analysis described in the previous sections. The Delaunay triangulation of the environment, which was used for route discretisation, is shown in Fig. 2d. Results of pairwise comparisons (V-values)

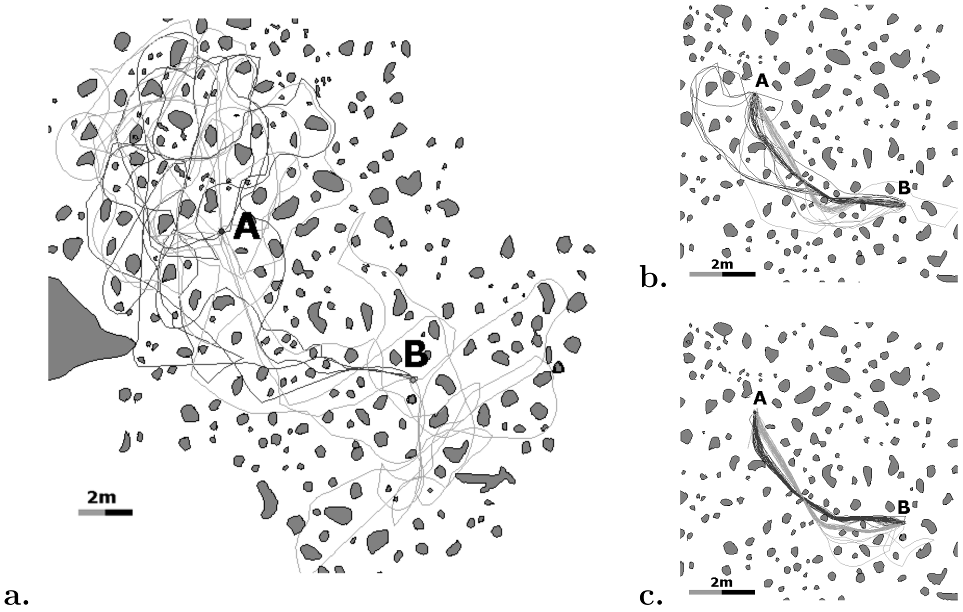


Figure 5: Route formation during training. (a) Session no.1 (b) session no.3 (c) last learning session (no.5), *light gray*: $A \rightarrow B$, *dark gray*: $B \rightarrow A$.

averaged over all pairs of routes taken from each learning session are given in the top two rows of table 2. The increasing number quantifies the mentioned convergence effect for both route directions.

session	1	2	3	4	5
mean similarity of routes $A \rightarrow B$	0.168	0.456	0.745	0.968	0.872
mean similarity of routes $B \rightarrow A$	0.102	0.365	0.701	0.838	0.986
p($A \rightarrow B = B \rightarrow A$)	0.078	< 0.05	< 0.001	< 0.001	< 0.001

Table 2: Evaluation of trajectories shown in figure 5. Upper rows show the mean similarities ($\overline{V}$) of trajectories for each direction and session. Results of the statistical analysis (α-values) are shown in the bottom row.

The analysis of routes showed that this participant developed fixed routes for each direction. In the next step, the difference between the sets of inbound and outbound routes was tested statistically. For this comparison, the place sequence for the inbound routes were inverted such that identical inbound and outbound routes would yield maximal similarity values. The null hypothesis (no difference between inbound and outbound routes) could be rejected in the second and all later sessions of the experiment (table 2 bottom row). These results refer to the data of only one subject. The full analysis of all participants presented in Hurlebaus et Al. (2008) demonstrated strong interindividual differences. While some participants always selected highly conservative routes (as shown here), others showed distributed routes of high dissimilarity. Participants forming conservative routes could be divided further into two groups using either identical or different in- and outbound routes.

The algorithm presented is based on a comparison between in-group and between-group variances. If within-group variance is high, existing between-group variances may not reach significance. An example for this effect is visible in Figure 5a. In this first learning session, inbound and outbound routes differ in their general extend and the focus of the search loops. However, this difference does not reach significance (2) due to large ingroup variability and small sample sizes.

4. Some Considerations and Extensions

Three-dimensional trajectories. Natural navigation happens in the three-dimensional world, even though, crawling or walking creatures will be limited to two-dimensional surfaces. As long as each x-y coordinate value can be associated with a unique elevation z in the environment, trajectories will be essentially two-dimensional. Truly three-dimensional trajectories may arise in "multi-storey" environments such as buildings or trees. In these cases, the discretised route will contain decision points from which the agent may proceed upwards or downwards. For example, wayfinding experiments in multi-storey buildings will result in three-dimensional data, when it is possible to use staircases or lifts to get to other floors (e.g. Hölscher et al., 2006). With respect to trajectory analysis, it suffices, however, to consider discrete decision points, keeping in mind that these may allow movements not only in the plane but also upstairs and downstairs (Fig. 6 a). Locations that allow wayfinding decisions in three dimensions can also be defined for tree climbing animals using branching points or branch segments as decision points (Fig. 6 b). In contrast, open 3D-spaces rarely provide obvious decision points. In such cases, the Delaunay tessellation can again be applied. In three dimen-

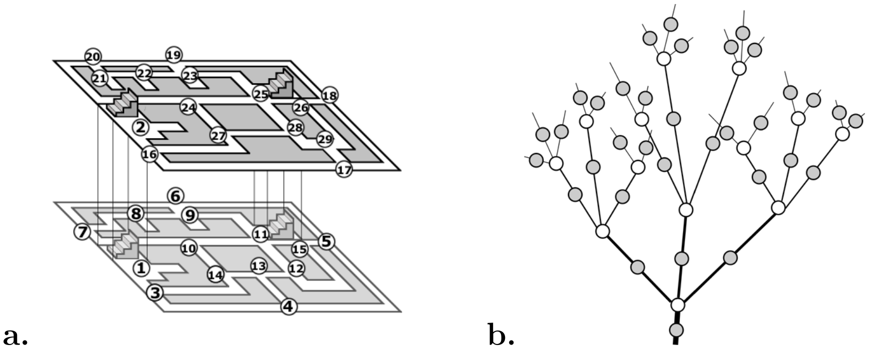


Figure 6: Decision points in 3D-space. (**a**) Numbered decision points in a multi-storey building. Staircases are treated in the same way as junctions in corridors, (**b**) A tree provides decision areas at all branching points (white circles) or optionally at branch segments (grey circles).

sions, triangles are replaced by tetrahedra (three-dimensional simplexes) as elements of the tessellation. Trajectories will then be described as sequences of passed simplexes.

In summary, our analysis can cope with trajectory data of different dimensionality. This is a result of the basic concept of the presented algorithm, that is reduction of trajectory data into sequences of visited decision points P with $P(x, y, z, ...)$. Also, the techniques for defining decision points can be scaled from two to three dimensions.

Space and time. The described method of trajectory analysis does not take into regard locomotion speed or stops along a trail. Thus, two trajectories comprising the same spatial layout may differ with respect to their temporal progress. As an example, consider two routes with identical place sequence but different times spent at each decision point. Let these time intervals be denoted by Δt_i^r for the i-th place of route $\mathcal{R}$, and Δt_j^s for the j-th place of route $\mathcal{S}$. In this situation, the algorithm for pair comparisons should be modified accordingly. So far, matches of visited decision points result in an entry of the value 1 in matrix M (Equ.1). Now, each match is weighted by the quotient of the time intervals spent at the decision point:

$$M_{i,j} = \begin{cases} \min\{\frac{t_i^r}{t_j^s} \ , \ \frac{t_j^s}{t_i^r}\} & \text{if } r_i = s_j \\ 0 & \text{if } r_i \neq s_j \end{cases} \tag{7}$$

Exact matches in space and time will still result in a matrix entry of 1, while the entry will drop to 0.5, say, if the times spend at a given location differ between the two routes by a factor of two.

Clearly, this procedure covers only differences in timing occurring at decision points. This is no problem in cases such as the one depicted in figure 2a, where the transitory sections of trajectories are not recorded anyway since the complete route data consist of the visiting sequence and duration of a set of feeders. In cases where the discretisation is obtained using the Delaunay triangulation, place transitions are "instantaneous", since the entire time spent in each triangle counts for the Δt of this triangle. In the same manner, transition times in mazes may be included in the Δt of the start or goal node of the transition, providing another means for dealing with temporal structure.

5. Conclusion

Routes can be conceived of as sequences of visited places and chosen actions, that are wayfinding decisions made at these places. The central idea of the presented framework for trajectory analysis is to extract decision points in various experimental environments taking into regard the structural characteristics of the environment and the experimental design. Together with the two subsequent steps, calculation of route similarity and statistical analysis, this framework is highly adaptable to various experimental conditions.

Extracted sequences of subsequently visited decision points are relevant for theories of decision making and route memory in spatial cognition and may be used for a subsequent analysis of behavioural and cognitive mechanisms. This analysis will need to take into account also additional information such as possible landmarks or egomotion data used by the agent and should eventually lead to a comprehensive, mechanistic model of wayfinding behaviour.

Acknowledgments

Support by the 'Landesgraduiertenförderung Baden-Württemberg' to KB is gratefully acknowledged. We thank Johannes Thiele for helpful discussions and comments.

References

Aginsky, V., Harris, C., Rensink, R., Beusmans, J., 1997. Two strategies for learning a route in a driving simulator. Journal of Environmental Psychology 17, 317–331.

Åkesson, S., Wehner, R., 2002. Visual navigation in desert ants *Cataglyphis fortis*: are snapshots coupled to a celestial system of reference? Journal of Experimental Biology 205, 1971–1978.

Arbib, M. A., House, D. H., 1987. Vision, brain, and cooperative computation. MIT Press, Ch. Depth and detours: an essay on visually guided behavior, pp. 129–163.

Bartumeus, F., Catalan, J., Viswanathan, G., Raposo, E., da Luz, M., 2008. The influence of turning angles on the success of non-oriented animal searches. Journal of Theoretical Biology 252, 43–55.

Benhamou, S., 2004. How to reliably estimate the tortuosity of an animal's path: straightness, sinuosity, or fractal dimension? Journal of Theoretical Biology 229, 209–220.

Benhamou, S., Bried, J., Bonadonna, F., Jouventin, P., 2003. Homing in pelagic birds: a pilot experiment with white-chinned petrels released in the open sea. Behavioural Processes 61, 95–100.

Bisch-Knaden, S., Wehner, R., 2003. Landmark memories are more robust when acquired at the nest site than en route: experiments in desert ants. Naturwissenschaften (2003) 90:127 130 90, 127–130.

Byrne, R., Noser, R., Bates, L., Jupp, P., 2009. How did they get here from there? detecting changes of direction in terrestrial ranging. Animal Behaviour 77 (3), 619–631.

Cant, E. T., Smith, A. D., Reynolds, D. R., Osborne, J. L., 2005. Tracking butterfly flight paths across the landscape with harmonic. Proceedings of the Royal Society of London B 272 (1565), 785–790.

Cartwright, B. A., Collett, T. S., 1987. Landmark maps for honeybees. Biological Cybernetics 57, 85–93.

Collett, T. S., Cartwright, B. A., Smith, B. A., 1986. Landmark learning and visuo-spatial memories in gerbils. Journal of Comparative Physiology A 158 (6), 835–851.

de Berg, M., van Kreveld, M., Overmars, M., Schwarzkopf, O., February 2000. Computational Geometry, 2nd Edition. Springer.

Dell'Ariccia, G., Dell'Omo, G., Wolfer, D., Lipp, H.-P., 2008. Flock flying improves pigeons' homing: Gps track analysis of individual flyers versus small groups. Animal Behaviour 76, 1165–1172.

DoCarmo, M., 1976. Differential Geometry of Curves and Surfaces. Prentice Hall.

Fajen, B. R., Warren, W. H., Temizer, S., Kaelbling, L. P., 2003. A dynamical model of visually-guided steering, obstacle avoidance, and route selection. International Journal of Computer Vision 54, 13–34.

Franz, M. O., Schölkopf, B., Mallot, H. A., Bülthoff, H. H., 1998. Where did I take that snapshot? Scene-based homing by image matching. Biol Cybern 79, 191–202.

Gerstmayr, L., Mallot, H. A., Wiener, J. M., 2008. A minimalistic model of visually guided path selection and obstacle avoidance behavior. In: Spatial Cognition 2008. LNAI 5248. pp. 87–103.

Gillner, S., Mallot, H. A., 1998. Navigation and acquisition of spatial knowledge in a virtual maze. Journal of Cognitive Neuroscience 10, 445–463.

Good, P., 1994. Permutation tests. Springer Series in Statistics. Springer-Verlag.

Graham, P., Durier, V., Collett, T., 2007. The co-activation of snapshot memories in wood ants. J. Exp. Biol. 210, 2128–2136.

Harvey, L., Davis, F., Gale, N., 1988. The analysis of class dispersion patterns using matrix comparisons. Ecology 69 (2), 537–542.

Heusser, D., Wehner, R., 2002. The visual centring response in desert ants - *Cataglyphis fortis*. J Exp Biol 205, 585–590.

Hölscher, C., Meilinger, T., Vrachliotis, G., Brösamle, M., Knauff, M., 2006. Up the down staircase: Wayfinding strategies and multi-level buildings. Journal of Environmental Psychology 26 (4), 284–299.

Hurlebaus, R., Basten, K., Mallot, H., Wiener, J., 2008. Route learning strategies in a virtual cluttered environment. In: Freksa, C., Newcombe, N., Gärdenfors, P., Wölfl, S. (Eds.), Spatial Cognition VI. No. 5248 in LNAI. Springer, pp. 104–120.

Kohler, M., Wehner, R., 2005. Idiosyncratic route-based memories in desert ants, *Melophorus bagoti*: How do they interact with path-integration vectors? Neurobiol Learn Mem 83, 1–12.

Korn, G., Korn, T., 2000. Mathematical Handbook for scientists and engineers. Dover Publications, Mineolu, New York.

Kuipers, B., 1978. Modeling spatial knowledge. Cognitive Science 2, 129–153.

Kuipers, B., 2000. The spatial semantic hierarchy. Artificial Intelligence 119, 191–233.

Lau, K.-K., Roberts, S., Biro, D., Freeman, R., Meade, J., Guilford, T., 2006. An edge-detection approach to investigating pigeon navigation. Journal of Theoretical Biology 239, 71–78.

Mantel, N., 1967. The detection of disease clustering and a generalized regression approach. Cancer Research 27 (2), 209–20.

Meade, J., Biro, D., Guilford, T., 2005. Homing pigeons develop local route stereotypy. Proceedings of the Royal Society B 272, 17–23.

Menzel, R., Greggers, U., Smith, A., Berger, S., Brandt, R., Brunke, S., Bundrock, G., Hülse, S., Plümpe, T., Schaupp, F., Schüttler, E., Stach, S., Stindt, J., Stollhoff, N., Watzl, S., 2005. Honey bees navigate according to a map-like spatial memory. Proceedings of the National Academy of Sciences 102, 3040–3045.

Merkle, T., Knaden, M., Wehner, R., 2006. Uncertainty about nest position influences systematic search strategies in desert ants. The Journal of Experimental Biology 209, 3545–3549.

Müller, M., Wehner, R., 1988. Path integration in desert ants - *Cataglyphis fortis*. Proceedings of the National Academy of Sciences 85, 5287–5290.

Needleman, S., Wunsch, C., 1970. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J. Mol. Biol. 48, 443–453.

Noser, R., Byrne, R., 2007. Travel routes and planning of visits to out-of-sight resources in wild chacma baboons, *Papio ursinus*. Animal Behaviour 73, 257–266.

Ohashi, K., Thomson, J., D'Souza, D., 2007. Trapline foraging by bumble bees: IV. optimization of route geometry in the absence of competition. Behavioral Ecology 18 (1), 1–11.

O'Keefe, J., Nadel, L., 1978. The Hippocampus as a cognitive map. Clarendon Press-Oxford.

Riley, J. R., Smith, A. D., Reynolds, D. R., Edwards, A. S., 1996. Tracking bees with harmonic radar. Nature 379, 29–30.

Schnell, G., Watt, D., Douglas, M., 1985. Statistical comparison of proximity matrices: applications in animal behaviour. Animal Behaviour 33, 239–253.

Schöner, G., Dose, M., Engels, C., 1995. Dynamics of behavior: theory and applications for autonomous robot architectures. Robotics and Autonomous Systems 16, 213–245.

Thiele, J., Winter, Y., 2005. Hierarchical strategy for relocating food targets in flower bats: spatial memory versus cue-directed search. Animal Behaviour 69, 315–327.

Thomson, J. D., Slatkin, M., Thomson, B. A., 1997. Trapline foraging by bumble bees: II. Definition and detection from sequence data. Behavioral Ecology 8, 199–210.

Tolman, E. C., 1948. Cognitive maps in rats and men. Psychological Review 55 (4), 189–208.

Trullier, O., Wiener, S. I., Berthoz, A., Meyer, J.-A., 1997. Biologically based artificial navigation systems: review and prospects. Progress in Neurobiology 51 (5), 483–544.

Vyssotski, A., Serkov, A., Itskov, P., Dell'Omo, G., Latanov, A., Wolfer, D., Lipp, H.-P., 2006. Miniature neurologgers for flying pigeons: Multichannel eeg and action and field potentials in combination with gps recording. Journal of Neurophysiology 95, 1263–1273.

Warren, W. H., Fajen, B. R., 2004. Behavioral dynamics of human locomotion. Ecological Psychology 16 (1), 61–66.

Wiener, J. M., Mallot, H. A., 2003. "fine-to-coarse" route planning and navigation in regionalized environments. Spatial Cognition and Computation 3 (4), 331–358.

Winter, Y., Stich, K., 2005. Foraging in a complex naturalistic environment: capacity of spatial working memory in flower bats. Journal of Experimental Biology 208, 539–548.

Wolbers, T., Büchel, C., 2005. Dissociable retrosplenial and hippocampal contributions to successful formation of survey representations. The Journal of Neuroscience 25 (13), 3333–3340.

Manuskript 4

Embodied Spatial Cognition: Biological and Artificial Systems

Ziele, Hauptergebnisse und Eigenanteil an der Publikation/dem Manuskript

Tiere und Menschen orientieren sich in ihrer Umwelt vorwiegend, um Orte wiederzufinden, Wege zu planen und auf unvorhergesehene Änderungen in der Umgebung zu reagieren. Entscheidend für das Orientierungsverhalten sind dabei das Sensorium des Individuums, die Informationsverarbeitung und die im Gedächtnis gespeicherte Repräsentation der Umgebung.

Der vorliegende Artikel gibt eine Übersicht über einzelne Orientierungsmechanismen und deren Komplexität und Bedeutung in der Raumkognition. Für die Repräsentation des Umgebungswissens werden dabei unterschiedliche Graphenstrukturen diskutiert. In einem Graph werden Orte durch Knotenpunkte und deren Verbindungen durch Kanten repräsentiert. Mit den Knotenpunkten und Kanten können weitere Informationen assoziiert werden. So können einzelnen Orten sowohl funktionale Eigenschaften (z.B. Informationen über Nahrungsvorkommen), als auch metrische Information (z.B. Entfernung zwischen zwei Orten) zugeordnet werden.
Das vorgestellte Konzept thematisiert nicht nur die natürlichen Navigationsleistungen, sondern kann auch als Grundlage bei der Konstruktion navigierender künstlicher Systeme dienen.

Viele im Artikel vorgestellten Aspekte basieren auf langjähriger Forschungsarbeit am Lehrstuhl *Kognitive Neurowissenschaft* (Prof. H.A. Mallot, Universität Tübingen). Eine vorausgehende ausführliche Übersicht zur Raumkognition wurde im Rahmen eines Berichts (*Deliverable*) des EU-Projektes *Wayfinding* verfasst (H.A. Mallot, K. Basten, D. Berger, and A. Schnee, *Spatial Cognition: From Animals to Humans*). Aus diesen Grundlagen entstand in enger Diskussion mit Prof. Hanspeter A. Mallot der vorliegende Artikel.

Der Artikel wurde in der Fachzeitschrift *Image and Vision Computing* veröffentlicht.

Referenz:
H.A. Mallot, K. Basten (2009), *Embodied spatial cognition: Biological and artificial systems*, *Image and Vision Computing*, 27, 1658-1670

Image and Vision Computing 27 (2009) 1658–1670

Contents lists available at ScienceDirect

Image and Vision Computing

journal homepage: www.elsevier.com/locate/imavis

Embodied spatial cognition: Biological and artificial systems

Hanspeter A. Mallot, Kai Basten *

Department of Cognitive Neurobiology, Faculty of Biology, University of Tübingen, 72076 Tübingen, Germany

ARTICLE INFO

Article history:
Received 3 December 2007
Received in revised form 14 August 2008
Accepted 2 September 2008

Keywords:
Spatial cognition
Spatial memory
Landmark use
Graph representations
Embodiment

ABSTRACT

In this paper, we sketch out a computational theory of spatial cognition motivated by navigational behaviours, ecological requirements, and neural mechanisms as identified in animals and man. Spatial cognition is considered in the context of a cognitive agent built around the action–perception cycle. Besides sensors and effectors, the agent comprises multiple memory structures including a working memory and a longterm memory stage. Spatial longterm memory is modelled along the graph approach, treating recognizable places or poses as nodes and navigational actions as links. Models of working memory and its interaction with reference memory are discussed. The model provides an overall framework of spatial cognition which can be adapted to model different levels of behavioural complexity as well as interactions between working and longterm memory. A number of design questions for building cognitive robots are derived from comparison with biological systems and discussed in the paper.

1. Introduction

1.1. Embodied spatial cognition

Behaviour in animals and man can be described by the perception–action-cycle linking sensory and motor components via central processing within the agent and via a feedback loop established by the environment. The cognitive processing required to initiate and control higher level behaviour will thus depend on the available sensorium, the effectors, the environment, and the behavioural tasks pursued by the agent. A major difference between robotic and biological systems lies in their respective sensors and the types of data these sensors provide. In biological systems, local views of the environment and distance measurements obtained from egomotion are prominent types of sensory data, whereas range and absolute position sensors (GPS) play a large role in technical systems. In biological systems, the association of a recognized view with some motor action (recognition triggered response), i.e. a stereotyped perception–action-pair, is therefore considered a basic element of spatial memory which at the same time allows to generate simple route following behaviour. More abstract forms of spatial memory are required for the planning of novel routes (cognitive maps) indicating that the abstraction level of memory scales with the complexity of the tasks performed.

In this paper, we will develop a framework for spatial cognition based on the action–perception-cycle and a biologically plausible sensorium. The resulting theory of spatial cognition may be called *embodied* in the sense that cognitive processing is shaped by the sensorium, the effectors and the behavioural tasks to be carried out by the agent. Following the discussions presented by Lakoff [49] and Gallagher [25], the following aspects of embodiment will be considered:

1. *Adaptation to sensorium:* Perception and the available types of sensory information influence the computations to be performed by the central processor.
2. *Abstraction:* Memory structures are “close” to the data structures of sensors and motor control, i.e. the transformations from sensors to memory and memory to action require only a small number of processing steps.
3. *Task-specificity:* Central processing is adapted to the behavioural repertoire determined by motor abilities and tasks.
4. *Incrementality:* Learning of spatial representations occurs incremental both in a local sense (learn about current place) and in a tasks-specificity sense (learn how to solve current task).

As a consequence of the task-specificity, the cognitive apparatus of animals and man is not a universal computer, but rather a collection of partially independent adaptations to various problems which are relevant in an animal’s life (see, for example, Hauser [33]). Indeed, the now common delineation of *spatial* cognition from other domains such as visual cognition, tool use, and social cognition (e.g. Spelke & Kinzler) [73] is an example for the task-specific approach. In this paper, we limit ourselves to the spatial domain, which is also the most widespread domain of cognition in the animal kingdom. Possible extensions of the framework to other domains are left to future research.

---

* Corresponding author. Tel.: +49 7071 2978830.
E-mail addresses: hanspeter.mallot@uni-tuebingen.de (H.A. Mallot), kai.basten@uni-tuebingen.de (K. Basten).

0262-8856/$ - see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.imavis.2008.09.001

Table 1
Task hierarchy in spatial cognition

Task	Required information processing and memory
Recognizing places	Memory of local position information characteristic of places
Finding home after excursion	Two mechanisms: landmark guidance and/or path integration
Following a route	Associate places with motor actions (S–R or S–R–S)
Recombining known route segments	Graphs (networks) of S–R–S associations
Route selection and planning	Working memory
Cross-country shortcuts	Metric embedding of places
Communicating about space	Naming places and actions

S, Stimulus; R, Response.

The framework for spatial cognition presented in this paper starts from stimulus–response schemata and simple versions of the perception–action-cycle. The cognitive level is reached by including various types of graph structures in the central memory structures allowing to model topological navigation and declarative memory.[1] The framework is extended to include hierarchical representations, route planning, functionality of places, and metric information. The approach is evolutionary in that the adaptive value of a behavioural competence and the required information processing are scaled up jointly, in small, simultaneous steps.

The main scope of this paper is to describe a general framework for modeling biological navigation systems in various species. The framework also allows comparisons with technical approaches and may be useful in situations such as navigation of large outdoor environments, learning high level spatial maps representing regions and places functionalities, and interfacing navigation systems with human users.

1.2. Tasks in spatial cognition: repertoire and ontologies

As a result of its evolutionary origin, the ontology of spatial knowledge, i.e. the selection of data types represented in spatial memory, reflects the behavioural repertoire for which it evolved. Table 1 gives an overview of tasks or behavioural competences in spatial cognition together with the required representations and information processing abilities, which may be summarized as the spatial ontology. While the main ordering criterion is based on behaviour, not on information processing, the resulting hierarchy is similar to the hierarchies proposed by Kuipers [46,47] and by Trullier et al. [80].

1.2.1. Place recognition and homing

The simplest tasks requiring some sort of spatial memory occur in a life style called *central place foraging*, found in many insects or other animals building some sort of borrow or nest and taking foraging trips from there (e.g. Wehner) [84]. One type of memory needed is a longterm or reference memory of the place itself, i.e. a representation of sensory cues allowing to recognize the place once it is reached. Depending on this recognition mechanism, there will always be a neighbourhood around that place from which homing is possible (catchment area) and a smaller neighbourhood within which no further approach can be made (confusion area), see Franz et al. [23]. In more complex memories, the mechanisms developed for the actual home may be applied to additional places in just the same way.

Another strategy that may be used for homing especially if excursions are longer, is path integration. It is based on continuous updating of a so-called home-vector, i.e. a metric representation of the vector linking home to the agent's current position. This updating requires some kind of working memory since the home-vector changes with egomotion. Vector knowledge may also be stored in longterm memory, for instance in the form of labels attached to remembered places and specifying the distance and direction to the central place (e.g. Cartwright and Collett) [11].

1.2.2. Route following

Extending the "catchment area" of the central place by storing additional places as stepping stones leads on to the next behavioural level which is following a route or sequence of steps. Routes are generally thought to be composed of discrete decision points which have to be recognized from some longterm memory contents, and route segments which may be traveled without longterm memory involvement. Decisions at the decision points are based on stimulus–response (SR) learning achieved by operant conditioning or reinforcement learning. While the chain-idea of route memory is theoretically appealing, direct evidence for these view comes mostly from maze experiments, where decision points are defined by the experimenter (e.g. Mallot and Gillner, Janzen and van Tourennout, Waller and Lippa) [51,42,83]. Of course, routes are also formed in cluttered, irregular environments where no explicit decision points are provided (Hurlebaus et al.) [40]. Route knowledge in open environments has been convincingly demonstrated e.g. in ants (Wehner et al.) [85] and bats (Schaub and Schnitzler) [67]. In these cases, routes seem to be continuous, giving no evidence for the discrete, stepwise character assumed in standard models of route memory. Still, the discrete structure may underlie also continuous routes, if decision points are close and catchment areas are large.

1.2.3. Inferring novel routes

The map level of spatial knowledge is reached if navigators are able to recombine parts of known routes into novel routes (Tolman 1948, O'Keefe and Nadel) [79,62]. Extending the route concept to maps amounts to replacing the chains of decision points and actions by networks or graphs which contain bifurcations and loops. In this situation, actions cannot simply be associated to decision points, since one decision point may indeed allow different actions such as "turn left" or "move uphill". Choosing a particular action depends on the currently pursued goal and requires two novel abilities: (i) a so-called declarative knowledge of the type "choosing action A_i at place P_j will lead to place P_k" (a three-way association: stimulus, response, expected next stimulus [S–R–S′]) and (ii) a planning stage allowing to select a sequence of places eventually leading to the goal. The planning stage is a type of spatial working memory generating executable routes from the map-like long-term memory. It may also allow to consider alternative routes and make high-level navigational decisions. The distinction between routes and maps, originally suggested with theoretical and behavioural arguments, is now well established also on neurophysiological grounds. Hartley et al. [32], for example, have shown that the basal ganglia are involved in stereotyped route following while map navigation activated cortical areas.

The process of generating a navigable route from a cognitive map,[2] i.e. route planning, requires a working memory to store the steps of the route while navigating. Bumble bees, for example, show systematic "trap-line" foraging patterns in spatial arrays of feeders,

---

[1] Declarative memory is a part of longterm memory storing declarations rather than procedures or associations. In spatial memory, map knowledge is considered declarative while route knowledge is not.

[2] The term cognitive map is used for different types of declarative spatial memory. In the simplest case, it is a graph of places and place transitions, but it may also include metric and hierarchical structure. We will use the term only in contexts where the different meanings cannot be confused.

requiring both route memory and some route planning abilities (Ohashi et al.) [60]. In humans, path planning is generally studied in the context of choosing between alternative routes to the same goal (Bailenson et al., Wiener and Mallot, Wiener et al.) [6,88,90]. We will discuss these performances in more detail below.

1.2.4. Metric knowledge

Metric information may be used for quite different tasks in spatial cognition. Motor activity always includes some metric knowledge about the movements performed. Together with egomotion estimates from visual and vestibular senses, this information is used in path integration. In longterm-memory, metric information is important to find shortcuts over novel terrain, i.e. shortcuts which are not just combinations of known route segments. Also, it will help making choices between alternative routes in route planning. Theoretically, three possible types of metric long-term memory may be distinguished, (i) full metric maps, (ii) metric embeddings, and (iii) local metric information.

In the *full metric map* approach, the cognitive map is a continuous chart in which known places or objects are plotted. Each location is represented in the full metric map (with some resolution) and a label provides information whether this location is occupied or not. Alternatively, the graph of places and actions discussed in the last paragraph might be augmented by *local metric information* such as distances associated with graph links or angles associated to adjacent pairs of graph links. Local metric information is relational and does not require the specification of a coordinate system. An intermediate case is *metric embedding* where coordinate values are associated to nodes only. Like the full metric map, metric embedding requires a coordinate system but does not represent points which are not actually occupied by a place node of the graph.

While these theoretical concepts are quite clear, experimental evidence on metric long-term memory is scarce. Metric judgments often lack symmetry, i.e. distance from A to B may be judged differently from the distance from B to A, (McNamara & Diwadkar) [52] or violate the triangular inequality. Foo et al. [20] have adapted the standard triangular completion paradigm used in studies of path integration (a working memory task) for long term memory. If two legs of a triangle are learned independently until each leg is well represented in longterm memory, the pointing performance between the open ends of the two legs is still poor. This is surprising if the places (vertices of the triangle) were represented with coordinate values in some global coordinate frame. Similarly, if subjects are taken to an unknown environment and are given repeated training for 10 subsequent days, their performance in pointing tasks between various points on their daily route is poor and does not improve with training (Ishikawa and Montello) [41]. Again, this result is surprising if metric localization in a global framework is assumed (Gallistel) [26].

1.2.5. Communication about space

In human spatial cognition, navigational aids such as drawn or printed maps, verbal directions, street signs and public transportation systems play an essential role. These navigational aids depend on social interaction and language. The behaviours which they subserve are therefore not purely spatial, but integrate elements from the spatial and social domains.

As a result of the predominance of such navigational aids in human spatial cognition, it may appear that spatial behaviour in general relies on language-based mechanisms. The view taken in this paper is rather opposite. Since wayfinding abilities can be observed throughout the animal kingdom, they must be a forerunner of language based cognition, not a consequence. Indeed, in their review on language evolution, Hauser et al. [34] see spatial cognition as part of a general "concept system" which is a preadaptation to language. This does not preclude the possibility of closer interactions between language and space in humans.

In accordance with this view, navigational behaviour and social communication in animals are rather independent. While navigational competences are widespread, communication abilities are rather rare and often limited to mating and warning. The most sophisticated example for communicating spatial information is certainly the honeybee dance language discovered originally by von Frisch [82]. Here, a honeybee returning to its hive from a profitable food source communicates the distance and direction of the food location to its nestmates. The key feature of this communication, however, is recruitment, i.e. getting fellow worker bees to search in the advertised location. Distance and direction to the food source are encoded in the frequency of "waggle"-movements of the abdomen and dance direction relative to the gravitational vertical, respectively. The function of these spatial codes is purely for communication, they are not the information that the dancer would use to go back to the food source.

1.3. A cognitive agent

Animal behaviour is generally analysed in the perception-action-cycle, where central processing is considered as some kind of controller trying to keep the organism in a favourable state. For the discussion of spatial cognition, it is important to distinguish within this controller different types of memory, including working and longterm (or reference) memory components. In order to structure the discussion, we briefly discuss a cognitive agent for spatial tasks depicted in Fig. 1. Clearly, the involvement of the various memory components will vary for different navigational mechanisms and different animal species. The similarity to the standard von Neuman architecture of a computer is not intended to mean that parallel processing is considered unlikely. Indeed, the working memory stage is a rich, structured system that contains a planning device selecting between different subgoals and navigational strategies. This planing device might well realize a subsumption architecture. Also, it should be stated that the boxes making up this agent are only logical in nature, giving no one-to-one relation to parts of the brain.

In order to address different types of spatial behaviour in robots, animals, and humans, we have divided the environment in two parts. The inner shell in Fig. 1 includes natural environments not intentionally changed by activities of the agent or conspecifics. Here, the standard wayfinding behaviour takes place as discussed, e.g. by Trullier et al. [80]. Humans (and some animals) change their environments by building structures or by providing information to conspecifics. The according behaviours require additions in various parts of the agent as will be discussed below.

1.3.1. Sensory input

In the wayfinding shell, sensory input relevant to spatial behaviour includes two types, landmarks (or local position information) and egomotion. For both types, the visual modality will be of great relevance.

Sensory input relevant to spatial tasks can be described by the local position information, i.e. the total sensor readings obtained from each position. Local position information can be thought of as landmark information in a very broad sense, which may be transformed to more explicit landmark information by applying different amounts of processing, leading (in the visual domain) to snapshots, geometric descriptions, or recognition of landmark objects. A second class of sensory information concerns egomotion which may be measured through different sensory modalities. Together, landmark and egomotion cues comprise "natural" cues whose presence in the environment is incidental and not due to some sort of information posting. In contrast, information is said

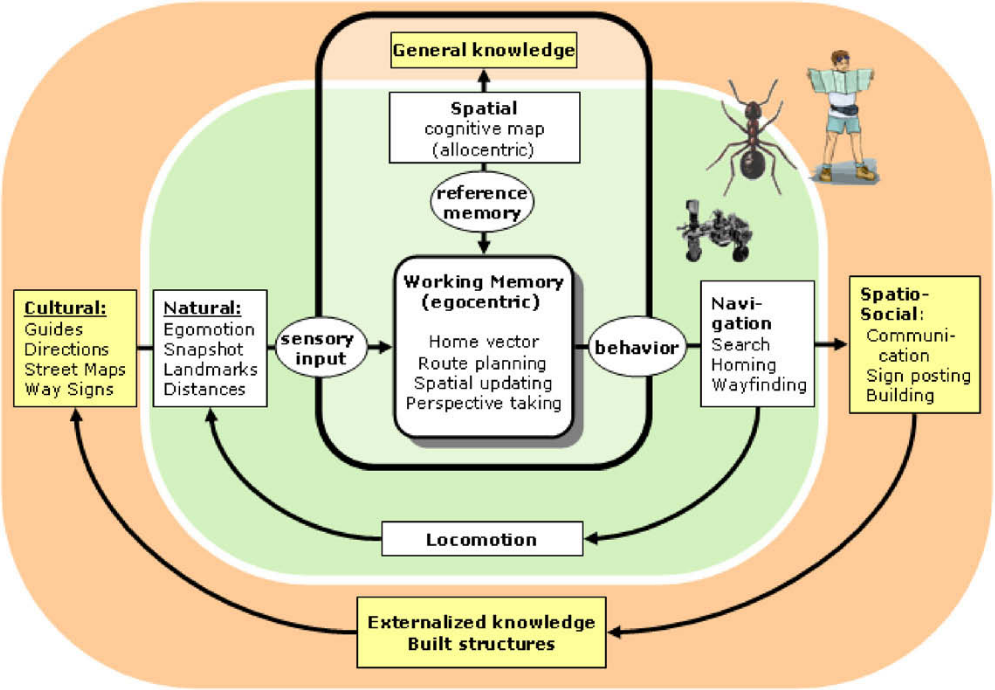


Fig. 1. Cognitive agent for spatial behaviour. Inner box with heavy outline: Agent with input and output. *Inner shell:* wayfinding behaviour, as is found in animals and humans. *Outer shell:* Interactions of social and spatial behaviour restricted mostly to humans. To date, robots generally also operate in the wayfinding domain.

to be "cultural", if available due to according efforts of the navigator or its conspecifics. In animals, examples for provided information include spatial communication such as the honeybee waggle dance, pheromone trails in ants, or olfactory territory markings. In humans, such provided information is abundant, including verbal directions, streets and corridors, way signs, printed maps, etc.

1.3.2. Reference memory

Longterm memory of space has been the subject of research on spatial cognition for decades. The emerging view of the "ontology" of spatial memory, i.e. of the types of structures represented, includes places (represented, e.g. as feature vectors of local position information), actions required to move from one such place to another, local metric information associated to such actions, etc. In many cases, spatial performance will additionally rely on general knowledge not referring to a particular place. Examples range from simple rules like "walk downhill to reach the water" to general knowledge about buildings or the abilities needed to interpret a printed map.

1.3.3. Working memory

Baddeley [5, p. 34] defined *general working memory* as the "temporary storage of information that is being processed in any of a range of cognitive tasks". In this sense, it is the central processing unit integrating sensory data and memory retrieved from the reference memory stage to control behaviour and generate new reference memory items. Working memory tasks in animal spatial cognition include simple path integration based on a home-vector representation, route planning, and object permanence (spatial updating). The spatial framework used in these tasks is ego-centric. Although working memory is depicted as one central stage in Fig. 1, we do not exclude the possibility, that different tasks have separate working memory instances.

1.3.4. Behaviour and action

The "core" spatial behaviour studied extensively in both animals and humans is wayfinding, i.e. getting from one place to another. Indeed, the wayfinding hierarchy sketched out in Section 1.2 is usually considered to cover the entire field of spatial cognition, at least in animals. However, wayfinding is not the only way in which both humans and animals interact with space. Building (nest) structures for individuals or groups, establishing and defending territories, or cooperation in larger groups and swarms all require some spatial abilities.

1.3.5. Environment

In the general view of spatial behaviour outlined here, the environment has to be considered as a part of the entire system. Human navigation rarely takes place in the wild, but deals mostly with finding intersections, reading signs and maps, or following verbal instructions. Descriptions of man-made environments together with their natural origins are sought in spatial information theory and GIS.

1.4. Design questions for spatial cognitive agents

For the spatial domain of cognition, a computational theory is emerging which builds on a number of simple mechanisms of orientation and place recognition which are integrated and scaled to an overall system. By and large, the resulting theory extends Kuipers' [46] spatial semantic hierarchy, which bases longterm, declarative knowledge of space (the cognitive map) on a graph of recognizable places and control rules describing transition between these places. In Section 2, we will review the current state of the graph theory of cognitive maps theory from a biological point of view.

From the comparison of biological and robotic spatial cognition systems, a number of design questions can be identified which are treated differently in different approaches. These design questions include

1. *Landmark representation:* Are landmarks a separate concept or data type included in the cognitive map or are they just properties of places from which they are perceived?
2. *Metric information:* While metric is clearly included in path integration and other motor mechanisms it is not clear how much metric information is actually represented in the cognitive map.
3. *Hierarchies:* The concept of a "place", even if envisaged as a geometric locus, will always include some spatial extension, if only as a result of uncertainty. In general, however, spatial knowledge will also represent regions of various hierarchical levels, which may also overlap.
4. *Functionality:* Spatial knowledge may be represented jointly with non-spatial information. As an example, consider the world graph approach of Arbib and Lieblich [3] where possible behaviours are stored together with each place representation.
5. *Using the cognitive map:* Reading the cognitive map requires a planning stage which buffers the information relevant for the current task and translates it into subgoals and motor actions. We will discuss the representation of actions as labels attached to graph links, and the planning stage as a spatial working memory.

The paper is structured by the graph approach to spatial cognition, which has been briefly discussed already in Section 1.2. In Section 2, we will consider various types and topologies of graphs that have been used for modeling spatial cognition. In Sections 3 and 4, individual components of the graph will be discussed. We will finish with some considerations concerning multiple memories of space, such as reference and working memories.

2. Graph representations of spatial memory and spatial behaviour

2.1. General

As usual, we consider graphs as structures composed of a set of nodes (or vertices) $V = \{v_i, i = 1, \ldots, n\}$ and a set of directional links (or edges) $E = \{e_i = (v_{i1}, v_{i2}, \lambda_i), i = 1, \ldots, m\}$. In our context, nodes are interpreted as spatial entities such as places, poses (place plus viewing direction), regions, etc. They are treated as empty containers that may carry different types of information such as views visible from each node, coordinates, etc. We will call these types of information "labels". The links are treated as triples $e_i = (v_{i1}, v_{i2}, \lambda_i)$ where $v_{i1}, v_{i2} \in V$ are the nodes connected by the link and λ_i is a label or set of labels attached to the link. These labels may contain information needed to proceed from the start node to the end-node of the link.

Graph descriptions of space are used extensively in computational geometry, in particular in the context of architectural planning. The scope of the discussion here is not so much on space itself, but on the mental representation of space. Still, many concepts will be reminiscent of ideas used in computational geometry.

Historically, graphs have been used to model spatial memory and spatial behaviour at least since the work of Tolman [78, Fig. 61, p. 177]. In Tolman's terms, the nodes of the graph are goals, or "ends" and the edges are "means-ends"-associations; the whole network is described as a "means-ends-field". Clearly, this is a very general characterization, intended to describe behaviour and problem solving in general, rather than a pure representation of space. An end, in Tolman's scheme, is a state taken by the agent and may be both spatial (e.g. "I have reached place A") or non-spatial (e.g. "I have obtained object B"). A means is a procedure for achieving state transitions. Thus, the means-ends-field appears quite similar to the idea of a finite state machine and its associated graph of state transitions.

More recent graph models of spatial memory and spatial behaviour have used different specifications of what nodes and links of the graph actually are or represent. Choices include places, views, ways, and regions for the nodes and according transition rules for the links. In this section, we discuss the main types of "space graphs", i.e. graph structures used in modeling spatial cognition; an overview is given in Fig. 2. We do not consider here the recognition of the various nodes or the information used for this task; this question will be postponed to Section 4.1.

2.2. Topologies

One dimension, along which space graphs may vary is topology. In animal behaviour, three major types of topologies are generally considered: chains, stars, and nets.

Chains of places and place transitions are called route memories. It is generally assumed that the recognition of a place (or other type of graph node; see below) triggers some action which then leads to the next node in the chain ("recognition-triggered response" of Trullier et al. [80]; "direction" of O'Keefe and Nadel [62]). Thus, the graph nodes have a label saying what to do when reaching the node (see below, Section 4).

Star topologies are often considered for insect spatial memories, especially for honeybees and ants. The centre of the star is the home or hive, or more generally the central place of a central place forager (Cartwright and Collett, Menzel et al., Dyer, Wehner et al.) [11,55,18,85]. Honeybees are able to learn routes from the hive to various feeders and back, but do not seem to take shortcuts between different feeders with non-overlapping catchment areas. In cases where shortcuts have been observed, conspicuous landmarks may have been used that result in large catchment areas. It should be noted, however, that experiments using radar techniques for tracking whole flight paths of bees, seem to indicate that honeybees might be able to integrate at least two vectors and to choose between at least two locations (Menzel et al.) [56,57].

True network topologies with loops are assumed for space graphs in mammals including man. While mathematically, the step from chains to networks may appear minor in nature, the presence of bifurcations (nodes with three or more links) requires an additional planning or choice mechanism between the possible transitions.

2.3. Place graph

The simplest type of graph representation of space is the place graph where each node represents a place in the world while the links are transversable connections between places. In Fig. 2a, we have assumed that these connections are directed, allowing to associate different travel instructions with each direction. Connections between places, if existing, need not be two-way, for example when it is possible to get from place A to place B immediately, while the reverse direction may require some deviation (Kuipers, Gillner and Mallot, Franz et al.) [46,47,29,22]. In animal and human behaviour, directionality of wayfinding knowledge is a widespread phenomenon. Knowing an outbound path does not necessarily imply that the way back is easily found. If the path is known, it may be quite different from the outbound route; for examples from ant navigation, see Kohler and Wehner [45], and Wehner et al. [85].

The notion of a "place", unanimous as it may seem, is not easily defined. Geometrically, a place may be identified with a point in some mathematical space, but in our context, places are memory items which have to be recognized based on some sort of landmark

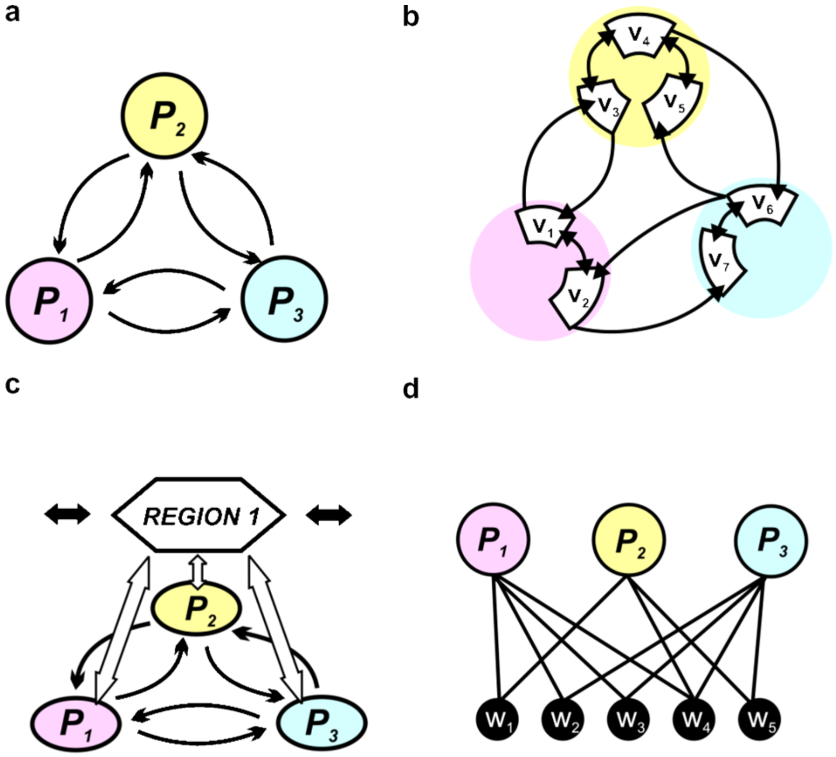


Fig. 2. Graph models of spatial memory. (a) Place graph with nodes (places) P_i. (b) View graph with nodes v_i. The coloured regions indicate viewpoints (places) from which the views were taken. (c) Hierarchical graph with place nodes P_i and an additional regional node. (d) Bipartite Graph of places P_i and ways W_i.

information. Empirically, knowledge of a place will always be described by an extended neighbourhood area of some anchor point, most notably the catchment area and the confusion area of a place. The confusion area is the set of points which cannot be distinguished from the goal location based on the cues represented in memory (see Hübner and Mallot) [39]. The catchment area is frequently studied in animal behaviour; it is a neighbourhood of the goal such that the animal is able to find the goal (or its confusion area) when released anywhere within the catchment area (Zeil et al.) [91]. Catchment areas can be quite large, especially when distant landmarks are used for characterizing a place.

In neural network models of cognitive maps, place recognition is often modeled by "place fields" similar to the hippocampal place-fields known from rodent neurophysiology (O'Keefe et al., McNaughton et al.) [61,54]. In this case, the place is described by the neuron's firing field, i.e. a region of several tens of centimetres in diameter. This firing field is again a neighbourhood of the cell's 'preferred' place, together with a weighting function declining towards the margins of the region. Since the firing fields of different neurons overlap, the place cells constitute a population code of space (e.g. Arleo et al.) [4].

The graph models used in architectural planning are also based on neighbourhoods of places, most notably the isovist or viewshed of a place (Benedikt) [8]. The isovist of a point comprises the set of all points that can be seen from one reference point. Points and their isovists can be connected in a visibility graph. In general, graphs based on catchment area and navigability will not be identical to graphs of isovists and mutual visibility, since navigable place transitions need not be visible from the starting point. Still, the isovist concept seems to capture some aspects of human space perception, such as the appreciation of architectural spaces (Wiener et al.) [87].

One problem shared by graph models form cognitive science and architectural planning is the selection of discrete points around which neighbourhoods are considered. In the animal literature, little is known about the selection of places to store in memory besides the primary goal locations such as nest and feeder; for one example from wood ants, see Durier et al. [17]. It seems likely, that simple optimality criteria apply.

2.4. View graph

While nodes of a place graph represent geometric points together with some neighbourhood, view graphs are based on the pose of the observer, i.e. on a combination of position and heading or body orientation. As in the place graph, this representation is not in terms of a coordinate vector, but rather in terms of a view of the environment associated with the heading direction and position as a vantage point.

At any one place, different views arise from different viewing directions. This is true even if the field of view spans the entire panorama, since the image coordinate system of the view will be centred around the heading direction. Indeed, despite their panoramic fields of view, wasps and other insects seem to use visual landmark information in a direction specific way, recognizing a place or panoramic view only if looked at from the direction in which it was encoded (Zeil et al.) [92]. This might be due to the fact that view information is encoded jointly with geocentric compass information provided by the skylight polarization pattern (Rossel and Wehner, Wehner and Müller) [66,86]. Evidence for directional encoding of positional information has also been presented for view-cells in the hippocampus of non-human primates, which seem to replace the mostly non-directional place-cells found in rodents (Rolls et al.) [65]. In psychology, directional effects of place recognition have been presented by McNamara et al. [53], and Schweitzer et al. [71].

Views might also be associated with place connections, in the sense that a particular view of or from a place is obtained when approaching or leaving the place using a particular connection. As indicated in Fig. 2b, the nodes of the view graph correspond to links between places in a place graph. Schölkopf and Mallot [70] have presented a theory of cognitive map based entirely on such views. Connections between views indicate that during navigation, two views can be encountered in immediate temporal sequence. The view graph is a dual of the place graph in the sense that each place connection corresponds to a view and each complete bipartite subgraph of the view-graph corresponds to a place. Actions corresponding to view transitions can be coded in an egocentric way, such that coordinate transforms in the process of self-localization and route generation can be avoided. Learning a view-graph amounts to accumulating schemata of the form $(\text{view}_i, \text{action}_k, \text{view}_j)$ by adding nodes and links to the current graph. Exploration can be guided by the fact that the view-graph can be decomposed into complete bipartite subgraphs, each corresponding to a place.

One computational problem solved by the view-graph is that of snapshot alignment. Even if panoramic snapshots are taken to represent places, place recognition from unknown directions requires costly comparison operations such as calculating the complete circular cross-correlation function between the stored and current snapshots. Stürzl and Mallot [76] have suggested a coarse-to-fine approach to this problem, calculating the raw view alignment from the low spatial frequency components. The view-alignment problem shows that view-graph representations may be useful even if panoramic vision is present. In less than panoramic vision, it seems even more natural to base spatial memory on views. Indeed, Gaussier et al. [27] have used the view-graph in robot applications.

The notion of the view graph is also akin to the idea of the axial graph used by Hiller et al. [38] for describing road choices by drivers in central London. It should be noted, however, that urban environments are channeled in the sense that movement is restricted to a discrete, low number of directions at any one time. The view graph approach is not limited to this situation but may be used in open environments as well.

2.5. Hierarchies and regions

Evidence for the representation of regions in spatial longterm memory comes from a number of sources including human nearness judgments and path choices made on printed maps (Bailenson et al.) [6] and optimal foraging in insects (Ohasi et al.) [59]. In navigational studies with humans, Wiener and Mallot [88] and Wiener et al. [90] investigated way choices in environments offering simple way alternatives. Subjects were familiarized with an environment comprising a loop of six places, each place with an outward connection ending blindly in another place. Each place was marked with an object taken from one of three semantic groups. Places with objects from each group were grouped together forming semantic regions containing two adjacent places from the loop together with their dead end extensions. When asked to navigate from one place to the place directly opposing on the loop, subjects preferred the route initially staying in the semantic region of the start place and then immediately leading to the goal region. The alternative route would lead from the start region to the third, irrelevant region and only then into the goal region. The results clearly indicate that regional knowledge is represented in longterm memory and used in spatial planning. Wiener and Mallot [88] suggest that the underlying memory structure is a hierarchical graph with additional regional nodes connected to all place nodes comprising the region (Fig. 2c). Thus, the hierarchical graph contains two types or nodes, places and regions, and three types of connections, place-to-place, place-to-region, and region-to-region. Path planning uses the lowest hierarchy level at the starting point but may rely on coarser levels for more distant parts of the journey.

2.6. Ways and places

An alternative approach to space graphs that combines elements of the place and view graph together with aspects of hierarchies has been presented by Kuipers et al. [48], see Fig. 2d. In this "skeleton" approach, two types of nodes exist, one for places and another one for paths or ways. A place node is connected to a way node, if the way passes through the respective place. Overall, place and way nodes thus form a bipartite graph. Routes are not represented by place-to-place connections, but by connections of the type place-to-way and way-to-place. An interesting feature of this approach is that it can explain the emergence of "highways" which are used whenever possible, thus introducing a backbone or skeleton in the cognitive map. As in the space syntax approach mentioned above (Hiller et al.) [38], the skeleton approach applies best to channeled spaces.

2.7. Landmarks as graph nodes

Throughout this section, we have considered nodes as spatial entities which have been occupied by the agent at some time or another. Places, poses, views, ways, and regions all are accessible to the agent and become nodes of the space graph only when the agent decides to store them as its current position. In contrast, landmarks need not be accessible spatial entities, either because they are unreachable or too far away (e.g. stars), or because they do not correspond to objects or places at all (e.g. view axes). We will discuss this notion of landmarks below (Section 4.1). However, some landmarks do correspond to accessible locations and may therefore be treated as nodes of a space graph. Indeed, this is the standard approach in robotics SLAM algorithms, where landmarks are localized in an occupancy grid and subsequently used for global metric embedding via triangulation. It should be noted, however, that this approach limits the use of landmarks, excluding possibly useful cues that do not lend themselves for localization.

3. Links and labels attached to links

Two graph nodes will be connected by a link, if it is known how to navigate from one node to the other. Links may be directional if only one travel direction can be carried out based on the knowledge stored in the graph. In the simplest case, the rule allowing navigation between two nodes may be the same rule used also for recognizing the place. In this case, all the agent needs to know is which neighbors can be found from each node and then start the search or homing procedure. In the logic of snapshot-based memory, this means that a node v_1 is (directionally) connected to a node v_2 if and only if v_1 is contained in the catchment area of v_2 (Franz et al.) [22].

Without further labeling, a link only contains the information that there is a way of navigating between the connected nodes. If this navigation requires special means or additional parameters, these have to be included in memory. One example of this is given by the "local vectors" in ant navigation which an ant can activate from memory to navigate from a recognized point to its nest (Bisch-Knaden and Wehner) [9]. The local vector contains information on distance and direction and is used just as a home-vector obtained from path integration. Unlike this vector, it is no longer subject to error accumulation once it has been correctly associated to that path. More generally, labels associated to links may be motor programs corresponding to the respective path segment. Just as the "local vector", such motor programs contain implicit distance and turning information which can be used in the construction of metric maps as described below. Also, labels attached to links may contain landmark information or guidances, such as in the procedure "follow the brick wall".

Allowing labels attached to links is one possibility of action representation. Alternatively, one might consider to attach these label to the nodes together with other labels discussed below. Indeed, in route memories, i.e. chain-like graphs, both possibilities are equivalent. In true maps (graphs containing nodes with three or more links), however, the following problem arises: Consider a label λ_i attached to a link from node υ_{i1} to node υ_{i2}. If this was to be attached to the start link υ_{i1}, it must take the form $(\lambda_i, \upsilon_{i2})$, since otherwise it would not be clear to which neighbour of υ_{i1} the action specified by λ_i will lead. As a result, there will be two types of labels attached to node υ_{i1}, one depending only on υ_{i1} itself (these are the labels discussed in Section 4 below) and one depending on υ_{i1} and another node from its neighbourhood. While this is possible, we consider it more parsimonious to attach label λ_i to the link itself, since this allows to treat chains and true graphs in the same way. In this scheme, each planning step from one node to the next is divided into two substeps, first the selection of the next node and second the retrieval of the required action (i.e. the label attached to the link to the next node).

4. Labels attached to graph nodes

4.1. Landmarks (local position information)

The most basic task in spatial behaviour is probably recognizing places. Even in simple search behaviour, the goal has to be recognized once it has been reached. In its simplest form, place recognition may be restricted to one or a few special places such as nest entries or feeding sites, but more complex spatial memories will contain larger numbers of known places. Place recognition has to rely on some sort of landmark information or distinct configuration of landmarks, i.e. sensor data characteristic of each place. Thus, the problem of place recognition is largely identical to the problem of landmark recognition. Let us define a landmark as a piece of sensory information, characteristic of a place, that is stored in memory and used in place recognition. Based on this definition, three questions about landmarks can be asked:

1. *Depth of processing:* What codes are generated from the sensory input to be stored in memory?
2. *Landmark selection:* Which parts of the input information are used to form landmark codes?
3. *Landmark usage:* How are landmarks used in spatial behaviour?

For the last question, it formally suffices to consider landmarks as characteristics of places, or, more generally, of the particular type of graph nodes considered in each model. The standard distinction between guidance (piloting) and direction (recognition-triggered response) made by O'Keefe and Nadel [62] and by Trullier et al. [80] can then be presented as two subsequent steps: The first step is pinning down the location of a place by landmark guidance, i.e. by moving such that a memorized view occurs on the navigator's retina. If this place (or state) is not the final destination, the next step will be a direction associated not with the landmark itself but with the graph link from the node recognized by means of the landmark to the next goal node. As mentioned before, attaching the action label to the graph link rather than to the start node of the link allows to treat route and topological map knowledge in just the same way.

4.1.1. Depth of processing

The most general account of landmark information is the notion of "local position information", defined as the sum of all sensory inputs perceivable at a certain location (Trullier et al.) [80]. Although this definition is not restricted to the visual modality, local views or snapshots will often be the most important type of local position information. In an extended environment, the local position information is a vector-valued function of position and pose (orientation of body and sensors) called the view manifold (Franz et al.) [22]; the components of the vector are the sensor readings currently obtained. For the visual modality, the local position information is the image or view locally obtained. These views, parameterized with the agent's position and pose, thus form a manifold containing all visual landmark information available in the environment. More specific kinds of landmarks can be derived by applying various amounts of image processing (see Table 2). For robots, low-frequency Fourier components or the panoramic image can be used to bring snapshots into register for subsequent image comparison (Stürzl and Mallot) [76]. Honeybees have been shown to use raw image information ("snapshots") and find places by matching currently visible and remembered snapshots (Cartwright and Collett) [10]. In addition, Zhang et al. [93] demonstrated that honeybees are also able to match features defined by motion contrast or motion parallax. The usage of raw snapshot information in humans has recently been demonstrated in virtual reality experiments using a feature-less smooth colour cycle as a texture of the wall in a circular room. Subjects presented with the view visible from a given place in the room were able to approach this place from other locations (Gillner et al.) [30]. In addition, Christou and Bülthoff [13] have shown that landmark objects are recognized faster and more accurately if they are presented in a familiar rather than in a novel orientation, indicating that views rather than 3D objects are remembered in the navigation task.

In a mechanism called the geometric module (Cheng) [12], rats determine their position from local depth maps, i.e. the distances to the surrounding walls. This information is also derived from visual input but requires more elaborate processing. In rats, possible cues to determine the distance of walls include motion parallax, the elevation of the lower edge of the wall in the rat's visual image, and maybe some stereopsis. In humans, it was shown that young children (Hermer and Spelke) [36] as well as non-attentive adults (Hermer-Vazquez et al.) [37] also use geometric information to find a hidden object in a room.

In everyday language, a landmark is an identified, nameable object or sight, whose recognition in an image requires a fully-fledged object recognition system. This type of information indeed increases navigation performance (Heft) [35]. The relation between landmark recognition and object recognition in general is not entirely clear (e.g. Fig. 3). While many computational problems are identical, neural processing of the two recognition systems in humans shows an interesting dissociation (Janzen and van Tourennout) [42]: The recognition of attentively perceived objects is associated with activation in the right fusiform cortex, whereas objects at decision points are processed mainly in the parahippocampal region, indicating a difference in the neural processing underlying landmark- and object recognition.

Table 2
Depth of processing in landmark recognition

Input	Image processing	Output
View manifold	Local recording	Raw snapshot at discrete points
	Edge detection	Edge based snapshot
	Depth estimate	Depth signature (geometric module)
	Object recognition	Identified landmarks
	Associating a text	Named landmark

Fig. 3. Places and landmarks. The photograph shows the view from a place called "four lakes view" near the town of Boppard, Germany. The four lakes are actually sections of the meandering river Rhein separated by mountain ridges. Thus, the name does not refer to landmark objects existing in the world (the "lakes"), but rather to the view obtained from the place. (photograph: F. König, http://de.wikipedia.org/wiki/Vierseenblicklift, GNU General Public Licence).

The level with the largest depth of processing in the sketched landmark hierarchy consists in the assignment of names to the recognized landmarks. Nameability of landmarks may be more relevant in spatial language and direction giving, than in actual navigation.

4.1.2. Landmark selection

As compared to the complete view manifold, or its processed versions, landmark information must be limited in two ways. First, only parts of each image should be stored as a memory. Second, landmarks should be recorded only at selected observer locations. Both selections will be guided by the following criteria:

1. *Salience:* Landmarks should be well recognizable. In verbal recall experiments, named objects are usually those standing out from the environment in size, shape, or functionality (Appelyard) [2].
2. *Relevance:* Landmarks should be remembered at places where navigational decisions are required. If adult subjects are asked to choose objects with potential landmark value they tend to name objects in the vicinity of road crossings (Allen et al.) [1] where a decision has to be made. Note that the distinction of landmark and object recognition in (Janzen and van Tourennout) [42] was based on this criterion.
3. *Permanence:* The landmark and its position should be constant over time. In children, navigational errors have been shown to result from choosing as landmarks salient but non-permanent objects such as fancy cars, which may be gone when visiting their original location again (Cornell et al.) [15]. Rat head direction cells follow the more distant of two independently moving landmarks even if they cover the same visual angle. A possible interpretation of this preference is that the more distant object will be larger and therefore less likely to move (Zugaro et al.) [94].

4.2. Coordinates

In the graph approach, different types of metric representation can be considered. (i) In the most complete case, the full metric map, a node is introduced for each point in a sampled plane. These nodes are defined primarily by their coordinates and gain place or landmark information only after according evidence has been collected. A global coordinate system is needed to specify the coordinates of each point. The full metric map is used for example in robotic Simultaneous Localization and Mapping (SLAM) algorithms (Thrun) [77]. (ii) By "local metric information", we denote distance information between pairs of nodes, typically represented as labels of graph edges, and angular information associated to the central one of a triple of nodes (or pairs of edges). Graph nodes are introduced only for places, poses etc, which at some point have been occupied by the agent. Local metric information may be globally inconsistent and does not assume a global coordinate frame. (iii) "Metric embedding" is an extension of the local metric case using the same basic perceptual data, i.e. node recognition, node distances, and bearings. Metric embedding is intermediate in that local metric information is checked for consistency by local triangulation. Inconsistencies are removed by optimizing the local metric data or by assigning optimized coordinate values to each image point. In the first case, the representation remains coordinate-free, while in the second case, a global coordinate frame will be introduced.

Given the available experimental data summarized in Section 1.2, it appears that the human cognitive map uses local metric (type ii) and maybe some metric embedding (type iii). Behavioural evidence for complete metric maps (type i) seems to be missing. Metric perceptions in printed maps are probably closer to image processing than to navigation. An algorithm for recovering a metric embedding from local metric data has been presented by Hübner and Mallot [39].

As a caution remark, it should be noted that abandoning the idea of a global metric framework for the cognitive map makes it hard to interpret data on global orientation in maps. For example, Restat et al. [64] have shown that urban environments on slanted landscapes are learned together with the up–down-axis of the environment. One possibility might be to represent slope locally, i.e. to orient the snapshots of all nodes into the upward direction, say, as suggested by Jeffrey et al. [43] for rodent hippocampal place fields.

4.3. Functionality

In neuropsychology, the separation of a purely spatial component from other types of declarative longterm memory, especially the episodic memory[3], remains a matter of debate (Eichenbaum)

---

[3] In neuropsychology, episodic memory is the memory for events which is considered a part of declarative memory. Since events always have a location, interactions between episodic and spatial memory may be expected.

[19]. One point where spatial and more general memory types may meet is the issue of functionality, i.e. memories of places associated with actions that may be carried out at each place. With the addition of functionality labels, the graph structure discussed so far becomes a much more useful tool for the organization of behaviour, allowing to generate route plans from functional demands (see for example Arbib and Lieblich) [3].

In biology, the simplest function associated with a place seems to be the place's attractivity. A model for the formation of territories based on a pre-existing spatial map and learning of attractivity values has been suggested by Stamps and Krishnan [75]. An extension for simultaneous learning of places and place attractivities has been presented by Schmolke and Mallot [69]. Indeed, the cognitive representation of territories seems to be an evolutionary old trait which may even form the root of individual recognition. Bee and Gerhardt [7], for example, showed that territorial neighbours are tolerated by American bullfrogs as long as they are recognized in their proper territory. If two owners of adjacent territories change their calling location, they are recognized as intruders even so they would be tolerated when calling from their familiar sites. This result indicates that individual recognition may root in territoriality, which requires a memory of place function.

5. Multiple spatial memories

So far, we have considered the reference or longterm memory of space, of which the space graph appears to be a suitable model. Longterm memory is required for the recognition of places, as well as for route following and route planning. In addition to this reference memory, a behaving agent, both robot or animal, will need some sort of working memory as an interface between perception, reference memory, and behaviour. Working memory can be involved in spatial behaviour either in isolation or combined with long-term memory. We will therefore discuss theoretical approaches to working memory structure in this section.

5.1. Ego- and allocentric memories

One important distinction between different types of spatial memory is that between allocentric and egocentric memory frames. The best way of making this distinction seems to be in terms of each memory's behaviour upon movements of the observer. We call a representation of space or objects "egocentric" if it changes in predictable ways as the observer moves, while the observer representation ("ego") remains unchanged. An egocentric representation could be a local map with the observer always at the origin, heading into the angular direction 0 degrees. If the observer ("ego") turns, the representation of outside objects will turn in the opposite direction, thus moving on circles about the origin. If "ego" translates, object representations will be translated in the opposite direction. The simplest type of an egocentric representation is the home-vector assumed in ant path integration (see Section 1.2). This home-vector is continuously updated during motion; it is thus a type of working memory where memory of preceeding steps is lost upon each update. Clearly, representations changing with each observer motion are not suitable for longterm memories which should be rather more stable.

The opposing term, "allocentric" representation, is used for representations of places or objects which do not change as the observer moves. In this case, the position of ego must be represented independently, as it will not stay in the centre of the map. Such representations do not necessarily need a coordinate frame with a well-defined "allo" as its centre. They may also be coordinate-free, e.g. based on distances and angles between objects or places. Still it will satisfy the requirement of stability with respect to observer motion. The term "allocentric" is thus missleading, as no centre needs to be defined. For a discussion of these terms, see also Klatzky [44]. Note that snapshots used as place-descriptors are allocentric in this definition even though they show the place's environment in an egocentric view.

Due to interactions between ego- and allocentric mechanisms and representations, the distinction between both types may not always be clear-cut. One example is the metric embedding of local distance and angular information obtained from ego-centric path integration into an allocentric global map. Interactions between path integration and allocentric metric maps have been studied experimentally. Blindfolded human participants performing a path integration task have been found to show less uncertainty in familiar as compared to novel environment, indicating that prior knowledge of the environment is integrated with the ongoing path integration mechanism (Philbeck and O'Leary 2005) [63].

5.2. Working memory

5.2.1. Path integration

The best studied example for a spatial working memory is path integration, i.e. the ability of many animals to return to their nest in a straight line, even after a long, winding excursion. Unlike long-term memory, path integration does not require a permanent memory of the trail, but only a continuous update of a home-vector pointing from the agent's current position to the "home" position. In insects, this vector is generally assumed to be given in an egocentric reference frame (see Merkle et al. [58] for a review of models of path-integration in insects). In humans, it is generally assumed that solving path integration tasks involves more complex memories, but direct evidence for this assumption is scarce. The "encoding error model" (Fujita et al.) [24], assumes that a sequence of egomotion steps is stored during the entire excursion and path integration is performed from these memories only when the return to home is being attempted. This model predicts that path integration should require longer processing time and be more prone to error for more complex excursions. However, this prediction could not be confirmed in behavioural experiments (Wiener and Mallot) [89], arguing again for a continuous update of a home-vector.

In rodents, the recent finding of grid cells in the entorhinal cortex has fostered an alternative model of path integration, which may be called locally allocentric. Path integration takes place in a local metric map or chart, realized by a sheet of laterally connected neurons. As the agent moves, an activity peak representing the agent's position will be shifted on that map in the appropriate direction. When the margin of the map is reached, the activity peak will leave and reenter the map from the opposing side, i.e. the map is assumed to realize periodic boundary conditions or a toroidal topology. Thus, as the agent moves for long distances, the activity peak will reach the same neuron over and over again, resulting in a regular grid of firing fields of the map neurons (McNaughton et al.) [54]. Indeed, the grid cells found in the rat entorhinal cortex by Hafting et al. [31] show the expected properties.

5.2.2. Other working memory tasks

Other tasks related to path integration have been studied intensively in humans. In *spatial updating*, subjects are asked to predict the egocentric position of objects or places after performing some egomotion. This requires the same operations used for keeping track of the starting point in path integration. Indeed, in ants, it has been suggested that memories of home and feeding site are both build up by two separate vector memories, each being fed and updated with the same, instantaneous egomotion data (Collett and Collett) [14]. Clearly, when trying to keep track of many objects, a local map model such as the locally allocentric grid-cell buffer seems rather more likely.

In *perspective taking* the subject is asked to imagine how a scene would appear when looked at from a different, not currently occupied view-point. This view-point is often marked by a person or an avatar in virtual reality studies. The computation is the same needed also in spatial updating, but it is not performed automatically, triggered by the actual movement of the observer, but under conscious control. Again, a local metric map might be a useful tool for achieving perspective taking performance. Perspective taking is often discussed in relation to social behaviour, considering the ability to infer what somebody else is seeing as a part of the Theory of Mind (Vogeley et al.) [81]. However, direct evidence for this interpretation seems to be missing.

The prediction of views visible from viewpoints not currently occupied also plays a role in *route planning*, where decision points along a future route have to be foreseen. Clearly, route planning requires a tight interaction between reference and working memory. In the hierarchical graph model of Wiener and Mallot [88] a "focal map" is assumed as a working memory stage (see also Fig. 2c). This focal map contains the current observer position and its surrounding region at highest resolution, while other regions (higher levels of the hierarchical graph) are represented only be their regional nodes. As a result, planning will prefer routes passing through the lowest possible number of regions.

Evidence for route planning abilities in animals has been presented e.g. by Cramer and Gallistel [16] who had vervet monkeys choose between two routes leading to various amounts of food items. The monkey optimized their routes for overall food intake, taking into account at least two future steps.

In terms of the ego- vs. allocentric distinction, home-vectors and object representations spatially updated during bodily motion clearly classify as egocentric memories. In perspective taking and route planning, no actual motions are performed. The memories involved may still be called egocentric with respect to ego's imagined position.

5.3. Longterm memory

An important distinction in longterm memory pioneered by O'Keefe and Nadel [62] and Squire [74] is that between declarative ("locale", map-like) and non-declarative ("taxon", route-like) representations. In the graph approach, the distinction lies simply in the topology of the graph itself. If the graph is a chain, the memory will be of the route type, while for non-trivial graphs with bifurcations (i.e. graphs including nodes with more than two edges) the map type is reached. Declarative memory is thus not associated with individual nodes but is a property of the graph structure as a whole. The recall of declarative memory will always involve a planning stage retrieving a route or path through the graph. Indeed, in the graph view, the need for this additional planning stage marks the major distinction between route and map. The planning stage itself, however, is a type of working memory. Neural network implementations that have been proposed for the planning stage are based on standard algorithms for graph search; for examples see Schmajuk and Thieme [68] and Schölkopf and Mallot [70].

The distinction between map and route type memories has been confirmed by functional magnetic resonance imaging (fMRI) studies showing that planing of novel routes from map knowledge recruits hippocampal and other cortical areas while repetition of familiar routes recruits the caudate nucleus in the basal ganglia (Hartley et al.) [32]. Similarly, in rats, stereotyped spatial behaviour is affected by pharmacological inhibition of the basal ganglia (Gengler et al.) [28].

6. Conclusion: hierarchical models of spatial memory and cognition

In this paper, we have presented a theory of spatial memory based on the graph approach, as summarized in Fig. 4. The graph approach in turn was motivated by the cognitive needs of various navigators ranging from ants foraging from a central "home" place all the way to humans operating in complex environments where the spatial and social dimensions are closely intertwined.

The graph approach unifies various hierarchical theories of spatial cognition which are based on the linear sequence of places, routes, and maps. This linear hierarchy of places, routes, and maps has been discussed in a number of different contexts, including (i) logical and computational complexity (Kuipers) [46,47], (ii) the sequence of acquisition during learning and ontogeny (Siegel and White) [72], (iii) the modularity of neural systems (O'Keefe and Nadel) [62], and (iv) the complexity of behaviour in animals (Trullier et al., Mallot) [80,50] and robots (Franz and Mallot) [21]. The existence of a logical hierarchy, however, does not necessarily imply that the same hierarchy should appear in the sequence of the acquisition of spatial knowledge or in the neural implementation. For example, spatial knowledge may be acquired from a printed map and than be transferred to a route plan used for actual navigation. In this case, the acquisition sequence goes "backwards", from map to route. In general, the different types of hierarchies may differ substantially although significant dependencies may be expected. We argue that the graph approach will be useful to understand the relation between the different types of spatial hierarchies listed above.

Another advantage of the graph model lies in its applicability to navigational systems of different animal species and different complexity. For example, the spatial memories of insects like ants and honeybees can be modeled as graphs with star-topology, where the nest or hive is in the center and other known places are linked to the centre but not to each other. In humans, more complex memories have been demonstrated representing true networks of places and allowing the generation of novel routes from segments of known routes. The graph model can be scaled to these various types of spatial memory and therefore allows detailed comparisons.

The construction of cognitive systems in robotics can take advantage of biological theories of cognition in various ways.

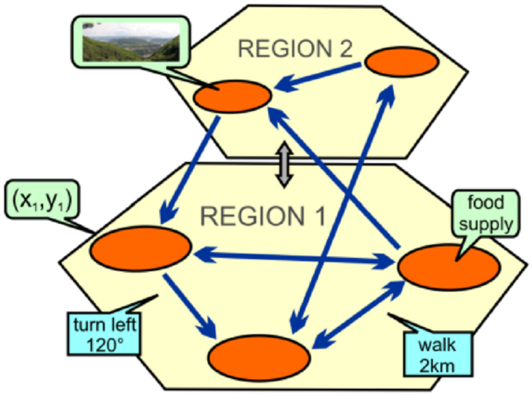


Fig. 4. Summary of space-graph approach to spatial longterm memory. The nodes are assumed to be places, but the model can be generalized to view-graphs etc. Graph links are labeled by action rules providing local (implicit) metric information. For graph nodes, three types of labels are indicated, local position information, global metric coordinates, and functionality of places. Regions are symbolized by the hexagons, which may also be nodes of a hierarchical graph on a higher level.

Besides copying partial solutions, the analysis of biological systems helps structuring the overall problem. In the design questions listed in Section 1.4, we have collected a number of issues where biological approaches seem to differ from robot approaches. The one field where we think interactions might be most promising is the representation of place functionalities, which can be used for joint planning of the spatial and non-spatial components of general action plans.

Acknowledgments

The work described in the paper was supported by the European Commission (FP6-2003-NEST-PATH Project "Wayfinding") and the Deutsche Forschungsgemeinschaft (SFB 550).

References

[1] Allen L. Gary, Kirasic C. Kathleen, Siegel W. Alexander, Herman F. James, Developmental issues in cognitive mapping: the selection and utilization of environmental landmarks, Child Development 50 (4) (1979) 1062–1070.

[2] D. Appleyard, Styles and methods of structuring a city, Environment & Behavior 2 (1970) 100–116.

[3] M.A. Arbib, Lieblich, Motivational learning of spatial behavior, in: J. Metzler (Ed.), Systems Neuroscience, Academic Press, New York, 1977, pp. 221–239.

[4] A. Arleo, F. Smeraldi, W. Gerstner, Cognitive navigation based on nonuniform gabor space sampling, unsupervised growing network, and reinforcement learning, IEEE Transactions of Neural Networks 15 (2004) 639–652.

[5] A. Baddeley, Working Memory, Clarendon Press, Oxford, 1986.

[6] J.N. Bailenson, M.S. Shum, D.H. Uttal, The initial segment strategy: a heuristic for route selection, Memory & Cognition 28 (2000) 306–318.

[7] M.A. Bee, H.C. Gerhardt, Individual voice recognition in a territorial frog (*Rana catesbeiana*), Proceedings of the Royal Society (London) B 269 (2002) 1443–1448.

[8] M.L. Benedikt, To take hold of space: isovists and isovis fields, Environment and Planning B 6 (1979) 47–65.

[9] S. Bisch-Knaden, R. Wehner, Local vectors in desert ants: context-dependent landmark learning during outbound and homebound runs, Journal of Comparative Physiology A 189 (3) (2003) 181–187.

[10] B.A. Cartwright, T.S. Collett, How honey bees use landmarks to guide their return to a food source, Nature 295 (1982) 560–564.

[11] B.A. Cartwright, T.S. Collett, Landmark maps for honeybees, Biological Cybernetics 57 (1987) 85–93.

[12] K. Cheng, A purely geometric module in the rat's spatial representation, Cognition 23 (1986) 149–178.

[13] C. Christou, H.H. Bülthoff, View dependence in scene recognition after active learning, Memory & Cognition 27 (1999) 996–1007.

[14] M. Collett, T. Collett, How do insects use path integration for their navigation?, Biological Cybernetics 83 (2000) 245–259

[15] E.D. Cornell, C.D. Heth, D.M. Alberts, Place recognition and way finding by children and adults, Memory & Cognition 22 (1994) 633–643.

[16] A. Cramer, C. Gallistel, Vervet monkeys as travelling salesmen, Nature 387 (1997) 464.

[17] Virginie Durier, Paul Graham, Thomas S. Collett, Switching destinations: memory change in wood ants, The Journal of Experimental Biology 207 (14) (2004) 2401–2408.

[18] Fred C. Dyer, Spatial memory and navigation by honeybees on the scale of the foraging range, The Journal of Experimental Biology 199 (1996) 147–154.

[19] H. Eichenbaum, Is the rodent hippocampus just for 'place'?, Current Opinion in Neurobiology 6 (1996) 187–195

[20] P. Foo, W.H. Warren, A. Duchon, M.J. Tarr, Do humans integrate routes into a cognitive map? map- versus landmark-based navigation of novel shortcuts, Journal of Experimental psychology: Learning, Memory, and Cognition 31 (2005) 195–215.

[21] M.O. Franz, H.A. Mallot, Biomimetic robot navigation, Robotics and Autonomous Systems 30 (2000) 133–153.

[22] M.O. Franz, B. Schölkopf, H.A. Mallot, H.H. Bülthoff, Learning view graphs for robot navigation, Autonomous Robots 5 (1998) 111–125.

[23] M.O. Franz, B. Schölkopf, H.A. Mallot, H.H. Bülthoff, Where did I take that snapshot? Scene-based homing by image matching, Biological Cybernetics 79 (1998) 191–202.

[24] N. Fujita, R.L. Klatzky, J.M. Loomis, R.G. Golledge, The encoding-error model of path-way completion without vision, Geographical Analysis 25 (1993) 295–314.

[25] Shaun Gallagher, How the Body Shapes the Mind, Oxford University Press, USA, 2005. December.

[26] C.R. Gallistel, The Organization of Learning, The MIT Press, Cambridge, MA, USA, 1990.

[27] P. Gaussier, A. Revel, J.P. Banquet, V. Babeau, From view cells and place cells to cognitive map learning: processing stages of the hippocampal system, Biological Cybernetics 86 (2002) 15–28.

[28] S. Gengler, H.A. Mallot, C. Hölscher, Inactivation of the rat dorsal striatum impairs performance in spatial tasks and alters hippocampal theta in the freely moving rat, Behavioral Brain Research 164 (2005) 73–82.

[29] S. Gillner, H.A. Mallot, Navigation and acquisition of spatial knowledge in a virtual maze, Journal of Cognitive Neuroscience 10 (1998) 445–463.

[30] S. Gillner, A.M. Weiß, and H.A. Mallot, Visual place recognition and homing in the absence of feature-based landmark information, Cognition, in press.

[31] T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, E.I. Moser, Microstructure of a spatial map in the entorhinal cortex, Nature 436 (2005) 801–806.

[32] T. Hartley, E.A. Maguire, H.J. Spiers, N. Burgess, The well-worn route and the path less traveled: distinct neural bases of route following and wayfinding in humans, Neuron 37 (2003) 877–888.

[33] M. Hauser, Wild Minds: What Animals Really Think, Allen Lane The Penguin Press, London, 2000.

[34] M.D. Hauser, N. Chomsky, W.T. Fitch, The faculty of language: what is it, ego has it, and how did it evolve?, Science 298 (2002) 1569–1579

[35] H. Heft, The role of environmental features in route-learning: two exploratory studies of way-finding, Environmental Psychology of Nonverbal Behavior 3 (3) (1981) 172–185.

[36] L. Hermer, E.S. Spelke, A geometric process for spatial reorientation in young children, Nature 370 (1994) 57–59.

[37] L. Hermer-Vazquez, E.S. Spelke, A.S. Katsnelson, Sources of flexibility in human cognition: dual-task studies of space and language, Cognitive psychology 39 (1999) 3–36.

[38] B. Hillier, A. Penn, J. Hanson, T. Grajewski, J. Xu, Natural movement: or, configuration and attraction in urban pedestrian movement, Environment and Planning B 20 (1993) 29–66.

[39] W. Hübner, H.A. Mallot, Metric embedding of view graphs. a vision and odometry-based approach to cognitive mapping, Autonomous Robots 23 (2007) 183–196.

[40] R. Hurlebaus, K. Basten, H.A. Mallot, and J.M. Wiener, Route learning strategies in a virtual cluttered environment, In: C. Freksa et al., (Ed.), Spatial Cognition (SC) 2008, number 5248 in Lecture Notes in Artificial Intelligence, Springer, 2008, pp. 104–120.

[41] T. Ishikawa, D.R. Montello, Spatial knowledge acquisition from direct experience in the environment: Individual differences in the development of metric knowledge and the integration of separately learned places, Cognitive Psychology 52 (2006) 93–129.

[42] G. Janzen, M. van Turennout, Selective neural representation of objects relevant for navigation, Nature Neuroscience 7 (2004) 673–677.

[43] K.J. Jeffery, R.L. Anand, M.I. Anderson, A role for terrain slope in orienting hippocampal place fields, Experimental Brain Research 169 (2006) 218–225.

[44] R.L. Klatzky, Allocentric and egocentric spatial representations: definitions, distinctions, and interconnections, in: C. Freksa, C. Habel, K.F. Wender (Eds.), Spatial Cognition. An Interdisciplinary Approach to Representing and Processing Spatial Knowledge, Number 1404 in Lecture Notes in Artificial Intelligence, Springer, Verlag, Berlin, 1998, pp. 1–17.

[45] M. Kohler, R. Wehner, Idiosyncratic route-based memories in desert ants, *Melophorus bagoti*: how do they interact with path-integration vectors?, Neurobiology of Learning and Memory 83 (2005) 1–12

[46] B. Kuipers, Modeling spatial knowledge, Cognitive Science 2 (1978) 129–153.

[47] B. Kuipers, The spatial semantic hierarchy, Artificial Intelligence 119 (2000) 191–233.

[48] B.J. Kuipers, D.G. Teucci, B.J. Stankiewicz, The skeleton in the cognitive map. A computational and empirical exploration, Environment and Behavior 35 (2003) 81–106.

[49] George Lakoff, Women, Fire, and Dangerous Things, University Of Chicago Press, 1987. April.

[50] H.A. Mallot, Spatial cognition: behavioral competences, neural mechanisms, and evolutionary scaling, Kognitionswissenschaft 8 (1999) 40–48.

[51] H.A. Mallot, S. Gillner, Route navigation without place recognition: what is recognized in recognition–triggered responses?, Perception 29 (2000) 43–55

[52] T.P. McNamara, V.A. Diwadkar, Symmetry and asymmetry of human spatial memory, Cognitive Psychology 34 (1997) 160–190.

[53] T.P. McNamara, R. Ratcliff, G. McKoon, The mental representation of knowledge acquired from maps, Journal of Experimental Psychology: Learning, Memory, and Cognition 10 (1984) 723–732.

[54] B.L. McNaughton, F.P. Battaglia, O. Jensen, E.I. Moser, M.-B. Moser, Path integration and the neural basis of the 'cognitive map', Nature Reviews Neuroscience 7 (2006) 663–678.

[55] R. Menzel, K. Geiger, L. Chittka, J. Joerges, J. Kunze, The knowledge base of bee navigation, The Journal of Experimental Biology 199 (1996) 141–146.

[56] R. Menzel, U. Greggers, A. Smith, S. Berger, R. Brandt, S. Brunke, G. Bundrock, S. Hülse, T. Plümpe, F. Schaupp, E. Schüttler, S. Stach, J. Stindt, N. Stollhoff, S. Watzl, Honey bees navigate according to a map-like spatial memory, Proceedings of the National Academy of Sciences 102 (2005) 3040–3045.

[57] R. Menzel, R.J. De Marc, U. Greggers, Spatial memory, navigation and dance behaviour in *Apis mellifera*, Journal of Comparative Physiology A 192 (2006) 889–903.

[58] T. Merkle, M. Rost, W. Alt, Egocentric path integration models and their application to desert arthropods, Journal of Theoretical Biology 240 (2006) 385–399.

[59] K. Ohashi, J.D. Thomson, D. D'Souza, Trapline foraging by bumble bees: IV. Optimization of route geometry in the absence of competition, Behavioural Ecology (2006), doi:10.1093/beheco/arl053.

[60] K. Ohashi, J.D. Thomson, D. D'Souza, Trapline foraging by bumble bees, Behavioural Ecology 18 (2007) 1–11.

[61] J. O'Keefe, N. Burgess, J.G. Bennett, K.J. Jeffery, E.A. Maguire, Place cells, navigational accuracy, and the human hippocampus, Philosophical Transactions of the Royal Society London B 353 (1998) 1333–1340.

[62] J. O'Keefe, L. Nadel, The Hippocampus as a Cognitive Map, Clarendon, Oxford, England, 1978.

[63] John W. Philbeck, Shannon O'Leary, Remembered landmarks enhance the precision of path integration, Psicologica 26 (2005) 724.

[64] J. Restat, S.D. Steck, H.F. Mochnatzki, H.A. Mallot, Geographical slant facilitates navigation and orientation in virtual environments, Perception 33 (2004) 667–687.

[65] E.T. Rolls, A. Treves, R.G. Robertson, P. Georges-François, S. Panzeri, Information about spatial view in an ensemble of primate hippocampal cells, Journal of Neurophysiology 79 (1998) 1797–1813.

[66] Samual Rossel, Rüdiger Wehner, The bee's map of the e-vector pattern in the sky, Proceedings of the National Academy of Sciences 79 (1982) 4451–4455.

[67] A. Schaub, H.-U. Schnitzler, Flight and ecolocation behaviour of three vespertilionid bat species while commuting on flyways, Journal of Comparative Physiology A 193 (2007) 1185–1194.

[68] N.A. Schmajuk, A.D. Thieme, Purposive behaviour and cognitive mapping: a neural network model, Biological Cybernetics 67 (1992) 165–174.

[69] A. Schmolke, H.A. Mallot, Territory formation in mobile robots, in: R.K. Standish, M.A. Bedau, H.A. Abbass (Eds.), Artificial Life VIII, The MIT Press, Cambridge, MA, 2002, pp. 256–259.

[70] B. Schölkopf, H.A. Mallot, View-based cognitive mapping and path planning, Adaptive Behavior 3 (1995) 311–348.

[71] K. Schweizer, T. Herrmann, G. Janzen, S. Katz, The route direction effect and its constraints, in: C. Freksa, C. Habel, K.F. Wender (Eds.), Spatial Cognition: An Interdisciplinary Approach to Representing and Processing Spatial Knowledge, Number 1404 in Lecture Notes in Computer Science, Springer Verlag, Berlin, 1998, pp. 19–38.

[72] A.W. Siegel, S.H. White, The development of spatial representations of large-scale environments, in: H.W. Reese (Ed.), Advances in Child Development, vol. 10, Academic Press, New York, 1975, pp. 9–55.

[73] E.S. Spelke, K.D. Kinzler, Core knowledge, Developmental Science 10 (2007) 89–96.

[74] L.R. Squire, Memory and Brain, Oxford University Press, New York, Oxford, 1987.

[75] J. A Stamps, V.V. Krishnan, A learning-based model of territory establishment, Quarterly Review of Biology 74 (1999) 291–318.

[76] W. Stürzl, H.A. Mallot, Efficient visual homing based on Fourier transformed panoramic images, Robotics and Autonomous Systems 54 (2006) 300–313.

[77] S. Thrun, Learning metric-topological maps for indoor mobile robot navigation, Artificial Intelligence 99 (1998) 21–71.

[78] E.C. Tolman, Purposive Behavior in Animals and Men, The Century Co., New York, 1932.

[79] E.C. Tolman, Cognitive maps in rats and man, Psychological Review 55 (1948) 189–208.

[80] O. Trullier, S.I. Wiener, A. Berthoz, J.-A. Meyer, Biologically based artificial navigation systems: review and prospects, Progress in Neurobiology 51 (1997) 483–544.

[81] K. Vogeley, M. May, A. Ritzl, P. Falkai, K. Zilles, G.R. Fink, Neural correlates of first person perspective an one constituent of human self-consciousness, Journal of Cognitive Neuroscience 16 (2004) 815–827.

[82] K. von Frisch, Tanzsprache und Orientierung der Bienen (engl. version: The Dance Language and Orientation in Bees (1967)), Springer, Berlin, Heidelberg, New York, 1965.

[83] D. Waller, Y. Lippa, Landmarks as beacons and associative cues: their role in route learning, Memory & Cognition 35 (2007) 910–924.

[84] R. Wehner, Desert ant navigation: how miniature brains solve complex tasks, Journal of Comparative Physiology A 189 (2002) 579–588.

[85] R. Wehner, M. Boyer, F. Loertscher, S. Sommer, U. Menzi, Ant navigation: one-way routes rather than maps, Current Biology 16 (2006) 75–79.

[86] R. Wehner, M. Müller, Does interocular transfer occur in visual navigation by ants, Nature 315 (1985) 228–229.

[87] J.M. Wiener, G. Franz, N. Rosmanith, A. Reichelt, H.A. Mallot, H.H. Bülthoff, Isovist analysis captures properties of space relevant for locomotion and experience, Perception 36 (2007) 1066–1083.

[88] J.M. Wiener, H.A. Mallot, 'Fine-to-coarse' route planning and navigation in regionalized environments, Spatial Cognition and Computation 3 (2003) 331–358.

[89] J.M. Wiener, H.A. Mallot, Path complexity does not impair visual path integration, Spatial Cognition and Computation 6 (2006) 333–346.

[90] J.M. Wiener, A. Schnee, H.A. Mallot, Use and interaction of navigation strategies in regionalized environments, Journal of Environmental Psychology 24 (2004) 475–493.

[91] J. Zeil, M.I. Hoffmann, J.S. Chahl, Catchment areas of panoramic snapshots in outdoor scenes, Journal of the Optical Society of America A 20 (2003) 450–469.

[92] J. Zeil, A. Kelber, R. Voss, Structure and function of learning flights in bees and wasps, Journal of Experimental Biology 199 (1996) 245–252.

[93] S.W. Zhang, M.V. Srinivasan, T. Collett, Convergent processing in honeybee vision: multiple channels for the recognition of shape, Proceedings of the National Academy of Sciences 92 (1995) 3029–3031.

[94] M.B. Zugaro, A. Berthoz, S.I. Wiener, Background, but not foreground, spatial cues are taken as references for head direction responses by rat anterodorsal thalamus neurons, The Journal of Neuroscience 21 (RC154) (2001) 1–5.

Manuskript 5

Mental Travel Primes Place Orientation in Spatial Recall - a Novel Demonstration of Situatedness in Spatial Cognition

Ziele, Hauptergebnisse und Eigenanteil an der Publikation/dem Manuskript

Um sich in der Umwelt zu orientieren, bekannte Orte wiederzufinden und Wege zu planen, ist es hilfreich, die Eigenposition in Relation zu umliegenden Landmarken und benachbarten Orten zu kennen. Dabei ist nicht nur die eigene Position, sondern auch die Ausrichtung der eigenen Körperachse wichtig, um die Bewegungsrichtung zu einem Zielort zu bestimmen. Diese Orientierungsprozesse laufen in vertrauten Umgebungen, wie beispielsweise im eigenen Wohnort, weitgehen automatisch ab. Für die Orientierung wird dabei auf räumliches Wissen im Langzeitgedächtnis zurückgegriffen. Im Langzeitgedächtnis sind Orte und deren Relationen zueinander gespeichert. Allerdings ist nicht eindeutig geklärt, ob Orte in einer bevorzugten Orientierung oder Ausrichtung, das heißt in einem bestimmten Referenzrahmen, gespeichert sind und welchen Einfluß eine solche Vorzugsausrichtung beim Abruf des Gedächtnisinhaltes in das temporäre Arbeitsgedächtnis hat.

Der Frage nach der Orientierung von räumlichen Gedächtnisinhalten wurde in diesem Projekt nachgegangen. Im ersten Teil der Studie wurden Versuchspersonen aufgefordert, einen zentralen Platz ihres Wohnortes zu skizzieren. Im zweiten Teil sollten sich die Versuchspersonen jedoch vor dem Zeichnen vorstellen, gedanklich einem Weg zu folgen, der diesen Platz überquert (hier Routen-Priming genannt). Ausgewertet wurden dann die Orientierungen der Skizzen auf dem Fragebogen. Die Ergebnisse zeigen, dass die Skizzen ohne Routen-Priming bevorzugt in Richtung zu einer dominanten Landmarke des Platzes ausgerichtet waren. Hingegen waren die Skizzen in der Versuchsbedingung mit Routen-Priming in ihrer Orientierung in deutlich weniger Fällen zu dieser Landmarke hin orientiert. Vielmehr zeigten die Skizzen in diesem Teil des Experiments vorwiegend eine Ausrichtung entlang der zuvor vorgestellten Routenrichtung. Diese Ergebnisse zeigen eine neue Form des Primings und ein neuen Effekt situationsgebundener Kognition. Der Vorteil der gedanklichen 'Vor-Ausrichtung' von

Orten entlang einer Route könnte beim späteren Ablaufen der Route eine verbesserte Ortserkennung sein.

Der Aufbau und die Parameter des Experimentes wurden von mir eigenständig entwickelt und zunächst mit Studenten des tierphysiologischen Großpraktikums in Tübingen getestet. In Zusammenarbeit mit Herrn Dr. Tobias Meilinger vom Max-Planck-Institut für Biologische Kybernetik (Tübingen) wurden weitere Daten aufgenommen und ausgewertet. Das Manuskript wurde zusammen mit den Autoren verfasst und zur Veröffentlichung in *Psychonomic Bulletin & Review* eingereicht.

Referenz:
K. Basten, T. Meilinger, H.A. Mallot (submitted), *Mental Travel Primes Place Orientation in Spatial Recall - a Novel Demonstration of Situatedness in Spatial Cognition*

Mental Travel Primes Place Orientation in Spatial Recall – a Novel Demonstration of Situatedness in Spatial Cognition.

Kai Basten[1], Tobias Meilinger[2] & Hanspeter A. Mallot[1]

[1] University of Tübingen, Cognitive Neuroscience Lab, Faculty of Biology, Auf der Morgenstelle 28, 72076 Tübingen, Germany
[2] Max-Planck-Institute for Biological Cybernetics, Department of Human Perception, Cognition and Action, Spemannstr. 44, 72076 Tuebingen, Germany

Abstract

The interplay of spatial long-term and working memories and the respective role of oriented ("ego-centric") and orientation-independent representations is an important but poorly understood issue in spatial cognition. Using a novel priming paradigm, we demonstrate that spatial working memory codes of a given location depend on previous tasks such as oriented mental travels and are thus situated in behavioural context. In two experiments, 136 passersby were asked to sketch an image of a highly familiar city square. Experiment 1 showed that most participants preferred similar view orientations in doing so. In Experiment 2, prior to drawing the sketch, participants imagined to walk one of two routes crossing the same square in opposite directions. As compared to Experiment 1, participants drew the sketch more often in the orientation of the imagined routes and less often in the orientation assumed without prior mental travel. This indicates that participants adjusted or selected information from long-term memory according to the situational context. We suggest that orientation priming plays a role in path planning and may facilitate way-finding afterwards. Possible mechanisms of orientation priming are discussed with respect to theories of orientation dependence in spatial memory.

key words: *spatial cognition; priming; frame of reference; working memory; place recognition*

Introduction

Finding one's way in large-scale spaces is a core cognitive function in humans and animals. In this task, spatial knowledge from long-term memory has to be activated and transferred to a working memory stage where planning, reasoning, and verbalisation takes place.

The relation of spatial long-term and working memories is often discussed in terms of the distinction of allocentric and egocentric representations. Allocentric spatial knowledge comprises the geometric shape of an environment (Wang & Spelke, 2002) as well as object-to-object spatial relations either without a preferred orientation (Sholl, 2001) or within an environmental reference frame (Mou & McNamara, 2002; Mou, Fan, McNamara, & Owen 2008). In contrast, egocentric systems code self-to-object relations as in perceptual representations or in view-dependent snapshots of the environment

(Sholl, 2001; Wang & Spelke; 2002). While the recall of spatial memories in general may involve allocentric codes, visual place recognition will mainly work with egocentric codes (Gillner, Weiss, Mallot 2008; Wang & Spelke, 2002). Indeed, the anticipation of future viewpoints can ease later recognition of the according views (Amorim, 2003). The information processing required to generate egocentric working memories from allocentric long-term memory depends on the assumed type of long-term place representation: If places are represented in a view-independent way (Byrne, Becker & Burgess, 2007), orientation has to be imposed at the time of retrieval (transformation). If places are represented as a collection of views (Mallot & Basten, 2009), the retrieval process will have to select (or interpolate) the appropriate view (selection). For path planning, Wiener & Mallot (2003) hypothesised an egocentric working memory stage generated from the reference memory for each combination of ego position and local target locations along the path.

Imagery of places may also be influenced by the subject's current situation. Such situational interferences have been demonstrated within a subject's immediate surrounding where imagery is facilitated by the alignment of physical and imagined body orientation (Avraamides & Kelly, 2008; May, 1996). On the contrary, situatedness typically does not affect performance when reasoning about non-immediate (i.e., more distant) environments (Wang 2004).

Using a novel priming paradigm, we assessed the orientation specificity of the retrieval of spatial long-term knowledge (as required for the production of place sketches) and tested the hypothesis that this retrieval can be influenced by prior working memory activity such as path planning and mental (i.e., imagined) travel. This hypothesis predicts that the view orientation of place sketches is primed by the direction of prior mental travel and thus by prior egocentric working memory representations. We tested highly-familiar long-term memory contents (i.e., a central square in the participants' home town) which is likely to have been encountered in many different orientations. Place sketches produced with and without prior mental travel were analysed for view orientation.

Experiment 1: Orientation-Specific Memories of a Highly Familiar Place

Methods

Passersby at a University cafeteria were asked to sketch the "Holzmarkt" a well-known square in the medieval city centre of Tübingen (see Figure 1) within an 8x8cm square on a DIN A6 sheet of paper. The University cafeteria was located approximately 2.5 km northwest of the Holzmarkt. About 30% of the people addressed agreed to participate. If participants asked in which perspective they should draw the square they were told to choose the perspective which they felt was most appropriate. After drawing, participants were asked to write down on the same sheet of paper their age, gender, place of residence (i.e., city district), and years of residency in Tübingen. Participation took approximately five minutes and was rewarded with candy.

From the 56 sketches obtained, six were excluded for incomplete data. Data from 27 women and 23 men (average age 22 years, *SD* = 2.2 years) were analyzed; on average, they had lived in or near Tübingen for 3.2 years (*SD* = 4.5 years).

Three independent raters categorized the orientation of the sketches into north-up, south-up, east-up or west-up. They gave identical judgments for 49 of the 50 sketches (98%) corresponding to a very good chance corrected interrater-reliability of kappa = .96. Only the remaining 49 sketches were analyzed further.

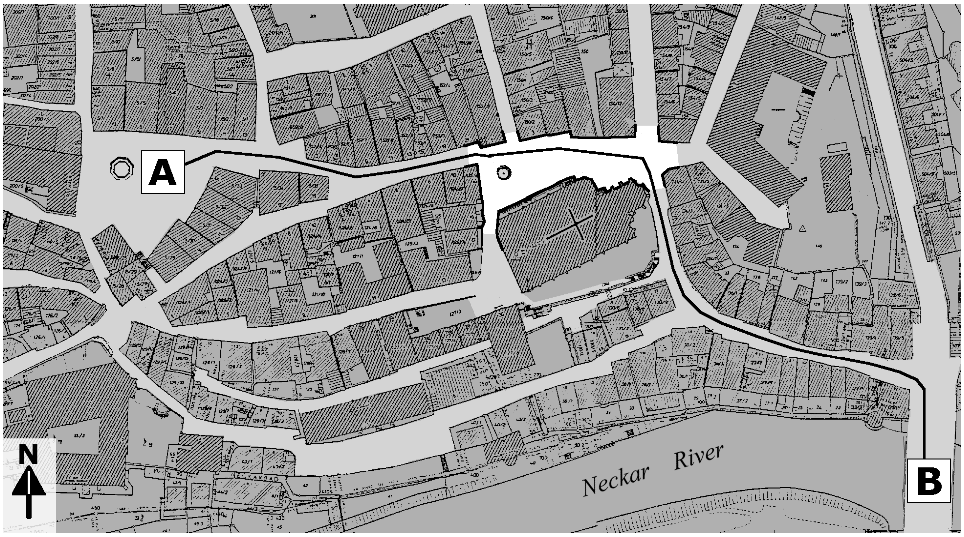


Figure 1. City map of the Holzmarkt (white) and its surroundings including a prominent church bordering the square on the south side. Participants were asked to draw the Holzmarkt in both experiments. In Experiment 2 (orientation priming) participants were additionally asked to imagine walking a route (black line) either from A (Market) to B (Neckar Bridge, eastward route) or from B to A (westward route) before drawing. No map was shown to the participants.

Results

As shown in Figure 2 (left) participants sketched the square with a preferred orientation (χ^2 test against a uniform distribution: χ^2 (3, N = 49) = 78, $p < .001$). Eighty percent drew the sketches south-up in contrast to 25% expected by an equal distribution (one-tailed binomial test with $\pi = .25$ and N = 49: $p < .001$). Other orientations were drawn less often than 25% (north-up: 6.1%, $p = .001$; west-up: 10.2%, $p = .009$; east-up: 4.1%, $p < .001$). Individual differences (gender, age, time and place of residency), did not show any effects, neither here nor in the second experiment and thus are not further reported. Results are discussed in detail in conjunction with Experiment 2 below.

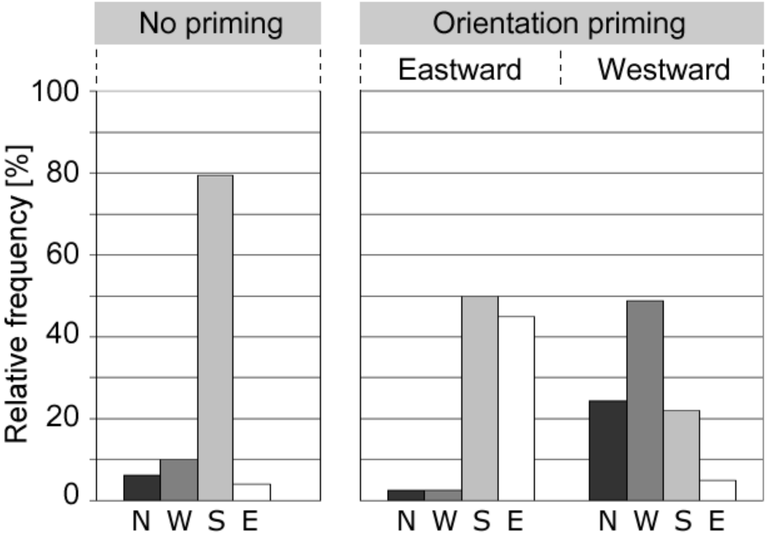


Figure 2. Frequency of sketch orientations without (Experiment 1) or with orientation priming (Experiment 2). Sketch orientations: N = North-up, W = West-up, S = South-up, E = East-up.

Experiment 2: Orientation Priming

Methods

General methods were same as in Experiment 1, but before sketching participants were asked to imagine walking a route across the Holzmarkt either in eastward or westward direction (Fig. 1). In the eastward condition, the written request 'Imagine to walk from the Market to the Neckar Bridge by crossing the Holzmarkt,' was handed to the participants. In the westward condition, Market and Neckar Bridge were exchanged. No participant of Experiment 2 participated also in Experiment 1. The experiment was run in multiple eastward/westward blocks, and participants were assigned to one of the two conditions in the sequence of recruitment.

Sketches of 81 participants could be analyzed (41 westwards, 40 eastwards, 42 women, 39 men). On average participants were 26 years old (*SD* = 9.5) and had lived for 6.7 years in Tübingen (*SD* = 9.1 years). Again only sketches with identical orientation ratings from all three raters were used (interrater-reliability of kappa =1.0).

Results

The distributions of sketch orientations differed between the two priming conditions ($\chi^2(3, N = 81) = 41.5$, $p < .001$, Cramérs $V = .72$) and from Experiment 1 conducted without priming (see Figure 2; route westwards: $\chi^2(3, N = 90) = 31.5$, $p < .001$, Cramérs $V = .59$; route eastwards: $\chi^2(3, N = 89) = 21.9$, $p < .001$, Cramérs $V = .50$). Participants were more likely to orient their sketches according to the primed direction. In the westwards priming condition, participants more often drew a sketch west-up (49% tested with a one-tailed binomial test against a probability of 10.2% as observed in Experiment 1, $N = 41$: $p < .001$) and north-up (24% tested against 6.1%, $N = 41$, $p < .001$). Less often they drew a sketch south-up (22% vs. 79.6%, $N = 41$, $p < .001$) which was the orientation mainly drawn without priming. We observed no differences in the frequency of sketching east-up (4.9% vs. 4.1%, $N = 41$, $p = .505$). In the eastwards priming condition, participants drew a sketch east-up more often than in the no-priming condition (45% vs. 4.1%, $N = 40$, $p < .001$). The south-up orientation was chosen less frequently (50% vs. 79.6%, $N = 40$, $p < .001$), but still was the most frequent one in this condition. No differences were observed for the north-up (2.5% vs. 6.1%, $N = 40$, $p = .290$) and west-up orientations (2.5% vs. 10.2%, $N = 40$, $p = .075$).

Discussion

Orientation Priming

In Experiment 1, participants chose to draw sketches of the square mostly with the south-up orientation. As all participants were highly familiar with the area (i.e. had experienced the square many times from multiple perspectives), the physical structure of the square itself likely determined the orientation of its recall (Mou & McNamara, 2002). Orientation preference may be derived from the geometric layout (east-west) as well as from the salient landmark (church) at the south side, or the south-up geographical slant of the square (see Restat, Steck, Mochnatzki, & Mallot, 2004). If the place representation relies on one intrinsic orientation, our data suggest that this is the southward orientation, despite the fact that the long axis of the square is east-west. Alternatively, place representations could contain views with multiple orientations as has been suggested for place recognition (Valiquette & McNamara 2007) and route following (Mallot & Gillner 2000). In this case, we need to assume that one view (i.e. southwards) is preferred..

In Experiment 2, when primed by imagining a route crossing the square, participants' sketches were more often oriented along the direction of the imagined route and less often in the orientation preferred without priming. Thus, mentally walking a route, as might also be done during route planning (Byrne et al., 2007), primed the orientation in which a location was recalled. That is to say, orientation priming changed the accessibility of the orientation of a place memory.

Standard priming procedures affect the accessibility of a stimulus by presenting semantically related, often co-occurring, or perceptually similar stimuli (Neelly, 1991), by presenting an object located on a route before versus after the target object (Janzen, 2006), or by presenting objects located close by versus further away from the target or within the same versus different spatial regions (McNamara, 1986). The orientation priming procedure used in the present experiment differs from other forms of spatial priming in that the primed item is not a particular place but the orientation in which this place is sketched or imagined. Also, priming is triggered not by the spatial, perceptual, or conceptual proximity of a stimulus, but by the (assumed) proximity of a working memory representation used during prior mental travel.

The orientation priming effect reported here affects the process of retrieving egocentric working memories of places from long-term memory (LTM). This process depends on the structure of long-term memory which may be view-dependent, aligned to an intrinsic reference direction, or independent of orientation. If LTM place representations are sets of views taken from various viewpoints (view-dependent memory, Gillner et al., 2008; Wang & Spelke, 2002), retrieval amounts to a selection process that picks one particular view to represent the place in working memory. In this case, orientation priming is the pre-activation of the view in LTM resulting in its subsequent selection. If, on the contrary, the environment is stored relative to an intrinsic reference direction and accessed more easily in that orientation (McNamara, Sluzenski, Rump, 2008; Mou & McNamara, 2002), imagining it in a different perspective requires a transformation such as a mental rotation into that perspective. In orientation priming, the transformed perspective might persist in working memory for subsequent recall. In orientation-independent memory (Byrne et al, 2007; Sholl, 2001), the retrieval process may select a particular landmark object (rather than a specific view) and assumes the perspective under which this object appears when looking from the square. Orientation priming will then result from pre-activated object representations. In summary, different mechanisms will be responsible for orientation priming in the different long-term memory models: priming could result from the prior selection of views or landmark objects, or it might originate from the orientation of persisting working memory contents.

The distinction between the selection and transformation mechanisms for view-dependent long-term memory is akin to the distinction between representations for place recognition and representations for locating a goal relative to a given position, as suggested by Valiquette & McNamara (2007). It seems therefore possible that the two hypothesized mechanisms, selection and transformation, may even co-exist and support performance in different tasks.

Situated and Embodied Cognition

Orientation priming affects the orientation of the recall of places. Thus, the retrieved orientation is not (only) dependent on the long-term memory content, but rather depends on the situational planning task. To our knowledge, this situational effect in spatial cognition has not been described before; it adds to situational effects, such as the interference effect between body orientations in current and imagined environments (Avraamides & Kelly, 2008; May, 1996), and the alignment of the current body orientation in an environment and the orientation of its paper map or memory representation, which results in better way-finding performance (Levine, Jankovic, & Palij, 1982; Mou & McNamara, 2002; Wang & Spelke, 2002).

Orientation priming is also in line with embodied cognition approaches which propose that representations, and in particular short-term representations, are based on sensorimotor and thus orientation-dependent representations (Barsalou, 2008; Wilson, 2002). Neuronal correlates supporting such view-dependent representation of scenes can be found in the parahippocampal place area, which is activated during mental imagery of places and in mental navigation (Epstein, 2008).

One final question concerns the function of orientation priming. The purpose of processing information about a specific route will generally be to follow this route afterwards. Recognition of locations along the route should be facilitated, if the representation is aligned with the upcoming perspective (Amorim, 2003). In this sense off-line planning of routes might facilitate later online-cognition while walking the route. Orientation priming might thus help to effectively prepare for anticipated situations in way-finding tasks.

Acknowledgements

We would like to thank Sandra Holzer and participants of the spatial cognition course 2008/2009 for help with data collection and analysis. Helpful comments on earlier versions of the text have been provided by Isabelle and Heinrich Bülthoff.

References

Amorim, M.-A. (2003). ‘What is my avatar seeing?’: The coordination of ‘out of body’ and ‘embodied’ perspectives for scene recognition across views. *Visual Cognition*, *10(2)*, 157-199.

Avraamides, M.N., & Kelly, J.W. (2008). Multiple systems of spatial memory and action. *Cognitive Processing*, *9*, 93-106.

Barsalou, L.W. (2008). Grounded cognition. *Annual Review of Psychology*, *59*, 617-645.

Byrne, P., Becker, S., & Burgess, N. (2007). Remembering the Past and Imagining the Future: A Neural Model of Spatial Memory and Imagery. *Psychological Review*, *114*, 340-375.

Epstein, R.A. (2008). Parahippocampal and retrosplenial contributions to human spatial navigation. *Trends in Cognitive Sciences*, *12*, 388-396.

Giller, S., Weiss, A., Mallot, H.A.(2008). Visual place recognition and homing in the absence of feature-based landmark information. *Cognition, 109, 105-122.*

Janzen, G. (2006). Memory for object location and route direction in virtual large-scale space. *The Quarterly Journal of Experimental Psychology*, *59*, 493-508.

Levine, M., Jankovic, I.N., & Palij, M. (1982). Principles of spatial problem solving. *Journal of Experimental Psychology General*, *111*, 157-175.

Mallot, H.A., & Basten, K. (2009). Embodied Spatial Cognition: Biological and Artificial Systems. *Image and Vision Computing*, *27*, 1658-1670.

Mallot, H.A., Gillner, S.D. (2000) Route navigation without place recognition. What is recognized in recognition-triggered responses? *Perception, 29:*43-55

May, M. (1996). Cognitive and embodied modes of spatial imagery. *Psychologische Beiträge*, *38*, 418-434.

McNamara, T.P. (1986). Mental representations of spatial relations. *Cognitive Psychology*, *18*, 87-121.

McNamara, T. P., Sluzenski, J., & Rump, B. (2008). Human Spatial Memory and Navigation. In H. L. Roediger, III (Ed.), *Cognitive Psychology of Memory. Vol. 2 of Learning and Memory: A Comprehensive Reference, 4 vols* (pp. 157-178). Oxford: Elsevier.

Mou, W., Fan, Y., McNamara, T.P., & Owen, C.B. (2008). Intrinsic frames of reference and egocentric viewpoints in scene recognition. *Cognition*, *106*, 750-769.

Mou, W., & McNamara, T. P. (2002). Intrinsic frames of reference in spatial memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, *162-170.*

Neely, J.H. (1991). Semantic priming effects in visual word recognition: A selective review of current findings and theories. In D. Besner & G.W. Humphreys (Eds.), *Basic processing in reading: visual word recognition* (pp. 264-336). Hillsdale: Erlbaum.

Restat, J.D., Steck, S.D., Mochnatzki, H.F., & Mallot, H.A. (2004). Geographical slant facilitates navigation and orientation in virtual environments. *Perception*, *33(6)*, 667-687.

Sholl, M.J. (2001). The Role of a Self-Reference System in Spatial Navigation. In D.R. Montello (Ed.), *COSIT 2001* (pp. 217–232). Berlin: Springer.

Valliquette, C., & McNamara, T.P. (2007). Different mental representations for place recognition and goal localiosation. *Psychonomic Bulletin & Review*, 14(4), 676-680.

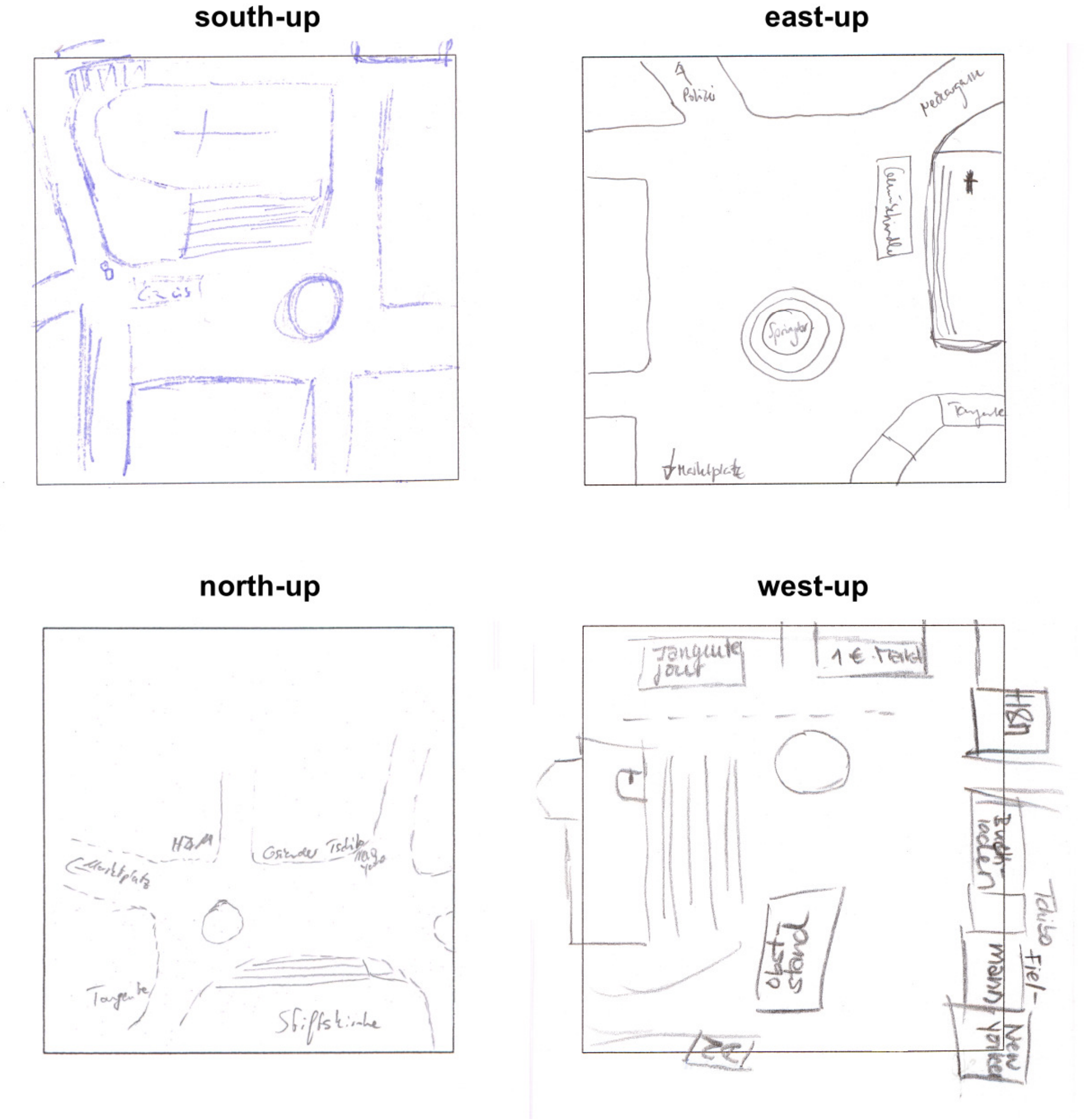
Wang, R. (2004). Between reality and imagination: when is spatial updating automatic? *Attention, Perception & Psychophysics*, *66*, 68-76.

Wang, F.R., & Spelke, E.S. (2002). Human spatial representation: insights from animals. *Trends in Cognitive Sciences*, *6*, 376-382.

Wiener, J., & Mallot, H.A. (2003). 'Fine-to-Coarse' Route Planning and Navigation in Regionalized Environments. *Spatial Cognition and Computation*, *3*, 331-358.

Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, *9(4)*, 625-636.

Supplementary Material



Supplementary material: Sample sketches for each orientation. Original size of drawing area was 8x8cm.

Zusammenfassung

Die vorliegende Arbeit untersuchte die Fähigkeit von Tieren und Menschen, sich in ihrem Lebensraum zu orientieren. Der Schwerpunkt der Arbeit lag dabei auf der Ortserkennung und der Einbettung von Ortswissen in Routen- oder Kartenrepräsentationen im Langzeitgedächtnis.

In der wissenschaftlichen Auseinandersetzung mit verschiedenen Orientierungsmechanismen und deren komplexem Zusammenspiel wurden bereits weitreichende Erkenntnisse über die Raumwahrnehmung, die Verarbeitung der Rauminformationen und der daraus folgenden Interaktion mit der Umwelt gewonnen.
Viele dieser Erkenntnisse stammen aus Experimenten, die das Orientierungsverhalten in Laborumgebungen untersuchen, in denen oft nur wenige kontrollierte Orientierungsmerkmale (Landmarken) präsentiert werden, die den Versuchspersonen oder Versuchstieren in der Regel bis zum Experiment unbekannt. Neben Laborexperimenten werden vor allem Tiere auch in ihrer natürlichen Umwelt beobachtet und auf ihre Orientierungsfähigkeiten hin getestet. Die Vielzahl der Landmarken in natürlichen Umgebungen erschwert allerdings die Interpretation der Daten, da die verwendeten Orientierungsmechanismen und Strategien nicht immer eindeutig isoliert werden können.

Im Rahmen dieser Doktorarbeit wurde zum einen der Erwerb und die Verwendung von Ortswissen untersucht, und zum anderen die Verknüpfung von Orten in Form von Routen- und Übersichtswissen betrachtet. In den Experimenten dazu wurden landmarkenreiche virtuelle Umgebungen verwendet oder bereits vorliegendes Ortswissen aus vertrauten natürlichen Umgebungen abgefragt, was zum einen zu natürlicheren Orientierungsbedingungen führen sollte, aber im Gegenzug eine Herausforderung für die Datenanalyse darstellte. Aus einzelnen Projekten der Doktorarbeit sind die fünf vorliegenden Manuskripte hervorgegangen:

Das erste Projekt beschäftigte sich mit der Ortserkennung bei Wüstenameisen. Es wird angenommen, dass Wüstenameisen, die in halbtrockenen Gebieten ihre Nahrung suchen, durch visuelle Merkmale entlang ihrer individuellen Routen geführt werden. Welche visuellen Merkmal das sind, ist bislang unklar. In diesem Projekt konnte mittels einer quantitativen Computersimulation gezeigt werden, dass es den Tieren auf Grund ihrer visuellen Verarbeitung möglich ist, einzelne Orte entlang ihrer

Routen anhand der Silhouette umliegender Objekte gegen den Himmelshintergrund (d.h. der Skyline) zu unterscheiden. Gegenüber anderen Verfahren zur Ortserkennung ist diese Skyline-Orientierung weitgehend beleuchtungsinvariant, benötigt geringe Verarbeitungs- und Gedächtnisressourcen und ist zudem biologisch plausibel.

Wie sich Menschen in Umgebungen mit vielen, unregelmäßig verteilten Landmarken oder Hindernissen (ähnlich der Umwelt der Wüstenameisen) orientieren können wurde bisher kaum untersucht. Der Erwerb von Routen- und Übersichtswissen in solchen Umgebungen war Thema des zweiten Projektes dieser Arbeit.
In einem Computerexperiment lernten Versuchspersonen in einer virtuellen Umgebung, charakterisiert durch zahlreiche ähnliche Hindernisse, sich zwischen zwei Zielpunkten hin und her zu bewegen. Während einige Versuchspersonen dabei immer wieder gleiche Routen nutzten, variierten die Routen bei anderen mit jedem Hin- und Rücklauf zwischen den Zielpunkten. Die Ergebnisse des Experiments weisen darauf hin, dass zum Lernen der Routen sowohl eine lokale (gleiche Routen) wie auch eine globale (variierende Routen) Navigationsstrategie angewandt wurde. Allerdings nutzten die einzelnen Versuchspersonen nicht ausschließlich die eine oder andere Strategie, sondern verwendeten individuell beide Strategien jeweils in unterschiedlichen Anteilen.

Aus der Auswertung der Laufspuren (Trajektorien) aus dem Experiment zum Routenlernen beim Menschen entwickelte sich ein eigenständiges Projekt zur allgemeinen Analyse von Trajektorien. In diesem dritten Projekt wurden Laufspuren als das Ergebnis zahlreicher lokaler Richtungsentscheidungen betrachtet, da Richtungsentscheidungen in der Regel nicht kontinuierlich entlang einer Route getroffen werden können, sondern nur an Orten (z.B. Strassenkreuzungen), die auch Richtungsentscheidungen erlauben. Durch diese Betrachtungsweise konnte eine Route für die Auswertung auf eine Sequenz von Richtungsentscheidungen reduziert werden. In einem zweiten Schritt wurden die Sequenzen unterschiedlicher Routen miteinander verglichen und im dritten Schritt statistisch ausgewertet. Die einzelnen Auswerteschritte können dabei jeweils flexibel an die vorliegenden Versuchsbedingungen angepaßt werden. Dadurch ist das Verfahren für zahlreiche Anwendungsmöglichkeiten geeignet und ist zudem vereinbar mit Theorien der Raumkognition, die das Routengedächtnis in seiner Struktur als Folge von diskreten Richtungsentscheidungen betrachten.

Im vierten Projekt der Arbeit wurde ein Modell zur Einbettung von Umgebungswissen in ein übergeordnetes Ortsgedächtnis dargestellt. In diesem Modell wurde das Ortsgedächtnis mittels einer Graphenstruktur repräsentiert, in der einzelne Orte, in Form von Knotenpunkten, durch Kanten mit benachbarten Orten verbunden sind und somit ein Netzwerk bilden. Der Graph wiederum kann durch Regionen in größere übergeordnete Bereiche strukturiert werden. Zusätzlich können einem Ort oder der Verbindung zwischen Orten räumliche (z.B. metrische Informationen oder lokale Ansichten) und nicht-räumliche (z.B. vorhandene Nahrungsquellen) Eigenschaften zugewiesen werden.

Das Graphenmodell vereint durch seine Flexibilität verschiedene Theorien der Raumkognition und damit verschiedene Ausprägungen des Ortsgedächtnises, die auf linearer Verkettung von Orten, verknüpften Routen und/oder Karten aufbauen. Zusätzlich bietet das Modell auf Grund seiner Graphenstruktur einen guten Rahmen, um Ortsgedächtnisse unterschiedlicher Komplexität in künstlichen Systemen zu implementieren.

Informationen aus dem übergeordneten Ortsgedächtnis müssen in der Planungsphase einer Route aus dem Langzeitgedächtnis abgerufen werden. In welcher Perspektive Ortsinformationen im Langzeitgedächtnis vorliegen, beziehungsweise ob diese Perspektive beim Abruf aus dem Langzeitgedächtnis beeinflußt werden kann, stand im Fokus des letzten Projekts dieser Arbeit. Versuchspersonen skizzierten dazu einen vertrauten zentralen Platz ihres Wohnortes. Untersucht wurde nun, in welcher perspektivischen Orientierung der Platz skizziert wurde.
Die Ergebnisse der Skizzenauswertung weisen darauf hin, dass zunächst markante Landmarken, wie ein auffälliges Gebäude, die bevorzugte Orientierung eines Ortes bei dessen Abruf aus dem Langzeitgedächtnis bestimmen können. Sollten sich die Versuchspersonen in einen zweiten Versuch hingegen vor dem Skizzieren vorstellen, einer bestimmten Route über den gefragten Ort zu folgen (Routen-Priming), drehten sich die dargestellten Orientierungen weg von der markanten Landmarke hin zu der Orientierung, mit der zuvor der Platz gedanklich überquert wurde. Damit konnte eine neue Form von Priming und ein neuer Effekt situationsgebundener Kognition gezeigt werden. Ein Vorteil der perspektivischen Bereitstellung von Ortswissen durch die Routenplanung könnte in einer erleichterten Ortserkennung bei anschließendem Folgen der Route liegen, da die zu erwartenden Perspektiven entlang einer Route schon durch den Planungsprozess voraktiviert wurden.

In den vorgestellten Projekten dieser Arbeit wurde das Ortsgedächtnis im Kontext der Ortserkennung, der Integration einzelner Orte in Routen, der Darstellung der Rauminformationen im Langzeitgedächtnis und des Abrufs von gespeichertem Ortswissen betrachtet. Die im Rahmen dieser Arbeit verwendeten Methoden und gewonnenen Ergebnisse machen deutlich, dass die Sensorik, die Informationsverarbeitung und die daraus folgende Interaktion mit der Umwelt zwar isoliert untersucht werden können, aber nur in der Betrachtung als einheitliches Konzept verständlich werden. Zudem konnte auch hier gezeigt werden, dass raumkognitive Prozesse zustands- bzw. situationsabhängig sind. Aus dieser Sicht ist Raumkognition verankert in einer physikalischen Interaktion mit der Umwelt basierend auf situativem Handel.

Anhang

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die mich in meiner Promotionszeit unterstützt und begleitet haben.

Besonderes bedanken möchte ich mich bei Herrn Prof. Dr. Hanspeter A. Mallot für die Aufnahme am Lehrstuhl Kognitive Neurowissenschaft, die Bereitstellung des Themas und die Betreuung meiner Doktorarbeit. Die ausgezeichnete fachliche Unterstützung und das Gewähren von Freiräumen war stets fruchtbar für meine Arbeit.

Großer Dank gilt Herrn Prof. Dr. Heinrich H. Bülthoff für die wissenschaftliche Begutachtung dieser Arbeit als Zweitgutachter.

Für sowohl kritische wie auch konstruktive Kommentare und zahlreiche anregende Diskussionen bedanke ich mich herzlich bei allen Mitarbeitern des Lehrstuhls 'Kognitive Neurowissenschaft'. Die vielen gehaltvollen Stunden waren durch individuelle Sichtweisen und Interdisziplinarität von besonderem Wert für meine Arbeit.

Herrn Dr. Jan Wiener und Frau Rebecca Hurlebaus danke ich für ihr außerordentliches Engagement und die belebende Zusammenarbeit in dem Projekt: *Route Learning Strategies in a Virtual Cluttered Environment* (Manuskript 2).

Für seine Unterstützung und die 'Psychologenbrille' im Projekt *View point priming in spatial recall a novel demonstration of situatedness in spatial cognition.* (Manuskript 5) danke ich herzlich Herrn Dr. Tobias Meilinger.

Viele Diskussionen und Ideen brachten das Projekt *Trajectory discretisation and sequence analysis for wayfinding experiments in spatial cognition* (Manuskript 3) auf den Weg. Dafür danke ich Herrn Dr. Johannes Thiele.

Ich danke Herrn Dr. Wolfgang Hübner für sein Engagement im Bereich der Programmierung und seine immer hilfreiche Unterstützung in diesen Fragen. Für zahlreiche kritische und belebende Diskussionen danke ich Herrn Dr. Gregor Hardießund Herrn Dr. Hansjürgen Dahmen.

Für allgemeine Anregungen und Korrekturen des Rahmentextes dieser Arbeit bedanke ich mich vor allem bei Frau Beate Dreher.

Ein großes Dankeschön geht an meine Freundin, meine Familie, und meine Freunde, die mich während meines Studiums und dieser Doktorarbeit immer unterstützt haben und für den nötigen Ausgleich sorgten.[3]

Ein weiterer Dank für finanzielle Unterstützung während meiner Doktorandenzeit geht an die Landesgraduiertenförderung Baden-Württemberg, das EU-Projekt *Wayfinding*, die Wilhelm-Schuler-Stiftung Tübingen und die Reinhold-und-Maria-Teufel-Stiftung Tuttlingen.

Danke

---

[3] Und ein herzliches Dankeschön an alle, die ich vergessen haben sollte.

Publikationsliste

Papers

Mallot HA, Basten K (2009): *Embodied Spatial Cognition: Biological and Artificial Systems.* Image and Vision Computing, 27, 1658-1670

Basten K, Mallot HA (accepted): *Simulated visual homing in desert ant natural environments: efficiency of skyline cues.* Biological Cybernetics

Basten K, Meilinger T, Mallot HA (submitted): *View point priming in the recall of a familiar city square - a novel demonstration of situatedness in spatial cognition*

Basten K, Mallot HA (in prep.): *Trajectory discretisation and sequence analysis for wayfinding experiments in spatial cognition*

Hardiess G, Basten K, Mallot HA (in prep.): *Trade-off between locomotion and working memory use in a block copying paradigm*

Reviewed Conference Papers

Hurlebaus R, Basten K, Mallot HA, and Wiener JM (2008): *Route learning strategies in a virtual cluttered environment.* In C. Freksa et al., editor, Spatial Cognition (SC) 2008, number 5248 in Lecture Notes in Artificial Intelligence, 104-120, Springer

Schmolke A, Basten K, and Mallot HA (2004): *Effects of intruders and environments on spatial behavior: A simulation study.* In: Ilg UJ, Bülthoff HH, and Mallot HA, editors, Dynamic Perception, 39-44, Akad. Verl. Ges. Aka, Berlin

Abstracts

Hardiess G, Basten K, Mallot HA (2009): *Task complexity modulates trade-off between locomotion and working memory usage in a large copying paradigm.* Cognitive Processing, 10, S154-S154

Hardiess G, Basten K, Mallot HA (2009): *The role of the trade-off between locomotion and memory use in a block copying paradigm.* Interdisciplinary College, Günne

Basten K, Mallot HA (2008): *Skyline cues for visual outdoor navigation: learning from desert ants.* Workshop at Int. Conference Spatial Cognition '08, Freiburg

Abstracts (2)

Basten K, Mallot HA (2008): *Using skyline information for spatial cognition: a simulation study of route learning in desert ants.* PENS/Hertie Winterschool: The design of Neuronal Networks: Contributions from Invertebrates

Basten K, Mallot HA (2007): *Desert Ant Navigation in Cluttered Environments: Which visual features are used?* Proceedings of the 10th Tübinger Perception Conference, Bülthoff et al. (Eds.)

Aschoff M, Basten K, Mallot HA (2007): *Path Integration and Obstacle Avoidance: Design and Test of a Virtual Agent.* Proceedings of the 10th Tübinger Perception Conference, Bülthoff et al. (Eds.)

Hurlebaus R, Basten K, Mallot HA (2007): *Human Route Learning in Virtual Cluttered environments.* Proceedings of the 10th Tübinger Perception Conference, Bülthoff et al. (Eds.)

Basten K, Mallot HA (2006): *Does integration of odometry and obstacle avoidance explain ant trajectories in cluttered environments?* German Workshop on Artificial Life (GWAL-7), Jena, Artmann and Dittrich (Eds.)

Basten K, Mallot HA (2006): *Desert Ants in Cluttered Environments: which mechanisms control trail selection?* Interdisciplinary College (IK 2006), Guenne, Germany, Chairs Malaka R and Spitzer M

Basten K, Schmolke A, Mallot HA (2004): *A Graph Representation Applied on Miniature Robots for Modelling Territorial Behavour.* In Proceedings of the 7th Perception Conference, Bülthoff, Mallot, Ulrich, Wichmann (Eds.), p. 64, Knirsch Verlag, Kirchentellinsfurt

Schmolke A, Basten K, Mallot HA (2003): *The Comparison of Different Spatial Representations in a Robot Model of Territoriality.* In Procceedings of the European Cognitive Science Conference 2003, p. 435, LEA Publisher/London

Diploma Thesis

Basten, K (2004): *Eine Graphenstruktur als Kognitive Karte für Territoriales Verhalten: Eine Simulation mit Miniatur-Robotern.* Diplomarbeit, Universität Tübingen