
Blickbasierte Mensch-Computer-Interaktion mit Geoinformationssystemen

Peter KIEFER

Institut für Kartografie und Geoinformation · ETH Zürich · Stefano-Franscini-Platz 5 · 8093 Zürich, Schweiz

E-Mail: pekiefer@ethz.ch

Zusammenfassung

Die Schnittstelle zwischen Mensch und Computer steht zunehmend im Fokus der geoinformatischen Forschung. Aber auch in der Praxis wird der Aspekt der Nutzbarkeit immer wichtiger für die effiziente Einsetzbarkeit und breite Vermarktung eines Geoinformationssystems (GIS). Dieser Beitrag gibt einen Literaturüberblick zur *blickbasierten Interaktion* mit (mobilen oder Desktop-basierten) Geoinformationen. Bei dieser Interaktionsform wird mit Eye Tracking-Systemen das Blickverhalten des Nutzers gemessen, in Echtzeit verarbeitet, und zur Anpassung der Nutzerschnittstelle verwendet. Das Kapitel bietet eine generelle Einführung in die Eye Tracking-Technologie und verweist auf aktuelle Forschungsansätze im Bereich blickbasierter Interaktion mit Geoinformation.

1 Einführung

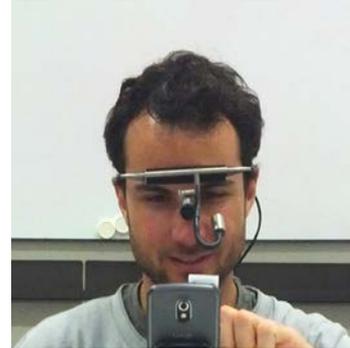
Der seit längerem anhaltende Trend hin zu webbasierten und mobilen Geoinformationssystemen (GIS) hat die Schnittstelle zwischen Mensch und Computer zunehmend in den Fokus der geoinformatischen Forschung rücken lassen (Hecht et al., 2013, Haklay, 2010). So steht als Nutzergruppe der heute entwickelten Systeme häufig nicht mehr der Domänen- oder Geodatenexperte im Vordergrund, sondern der durchschnittliche Internetnutzer, dessen Erwartungen an die intuitive Nutzbarkeit und Erlernbarkeit des Systems hoch sind. Aber auch beim Einsatz klassischer GIS in Unternehmen oder in der Verwaltung trägt die effiziente und effektive Nutzbarkeit ganz wesentlich zu deren Wirtschaftlichkeit bei (Haklay, 2010).

Eine besondere Herausforderung stellt die Mensch-Computer-Schnittstelle im Bereich mobiler GIS und ortsbezogener Dienste dar. Hier gilt es, Interaktionskonzepte zu entwickeln, die nicht nur den Einschränkungen der mobilen Hardware Rechnung tragen, sondern auch auf das jeweilige mobile Nutzungsszenario abgestimmt sind. So beschäftigt sich beispielsweise die mobile Kartografie mit der Visualisierung von Geoinformation auf kleinen Bildschirmen (Meng, 2005, Reichenbacher, 2001), während die Forschung zu adaptiven Assistenzsystemen darauf abzielt, die vom Nutzer benötigte Information aus dem Kontext abzuleiten und somit die Interaktion zwischen System und Nutzer zu minimieren (Raubal und Panov, 2009).

Einfache Kontextadaptionsregeln reichen jedoch häufig nicht aus, um eine intelligente und intuitive Nutzerschnittstelle zu realisieren. So ist beispielsweise die GPS-Position eines Nutzers letztlich nur ein Indikator für sein Informationsbedürfnis – die wirklichen Intentionen



Photo von SensoMotoric Instruments (Creative Commons Attribution 2.0 Lizenz)
<https://www.flickr.com/photos/smieyetracking/5470931994/>



Eigene Abbildung

Abb. 1: Remote-Eye Tracker im Fahrsimulator (links) und mobiler Eye Tracker bei der Interaktion mit einem Mobiltelefon (rechts).

nen bzw. kognitiven Prozesse bleiben vorerst unbekannt. Das *kognitive Engineering* von (mobilen oder Desktop-basierten) Geoinformationssystemen (Raubal, 2009) setzt an diesem Punkt an. Das Ziel ist hierbei die Erstellung von intelligenten Systemen, die auf Grund verschiedener Sensorinformationen („bottom-up“) und eines Nutzermodells („top-down“) den kognitiven Prozess des Nutzers zu erkennen und zu unterstützen versuchen.

Im folgenden Beitrag soll die *blickbasierte Interaktion* als ein Grundprinzip vorgestellt werden, das in der aktuellen Forschung zum kognitiven Engineering von geoinformatischen Nutzerschnittstellen diskutiert wird. Dieser Ansatz nutzt die technologischen Möglichkeiten der Blickerfassung (eng.: Eye Tracking) – einer Sensortechnologie, mit der sich die Blickbewegungen einer Person messen lassen – zur Realisierung aufmerksamkeitsbasierter Nutzerschnittstellen (Vertegaal, 2003). Grundannahme ist hierbei, dass die visuelle Aufmerksamkeit einer Person besonders gute Rückschlüsse auf ihre kognitiven Prozesse zulässt, da der größte Teil der entscheidungsrelevanten Information über die visuelle Sensorik aufgenommen wird.

Der folgende Abschnitt führt Grundbegriffe der Eye Tracking-Technologie ein und bietet einen Forschungsüberblick über den Einsatz von Eye Tracking in Studien zur Kartografie, zu GIS und in der Raumkognition. Hierbei werden die Blickdaten *nach* der Interaktion mit dem System analysiert. Abschnitt 3 geht einen Schritt weiter und diskutiert, wie die Eye Tracking-Daten *während* der Interaktion mit einem System verarbeitet und somit als Eingabemodalität für GIS verwendet werden können. Der letzte Abschnitt gibt einen Ausblick auf die zukünftigen Herausforderungen im Bereich blickbasierter geoinformatischer Nutzerschnittstellen.

2 Eye Tracking und Geoinformation

2.1 Einführung in die Blickerfassung

Dieser Abschnitt gibt eine kurze Einführung in Technologien zur Blickerfassung – im Folgenden Eye Tracking genannt. Für weitergehende Informationen seien die beiden Stan-



Abb. 2: *Links:* Aggregation von Gazes (rot) zu Fixationen (blau), verbunden durch einen Scanpath. *Mitte:* Visuelle Analyse des aggregierten Blickverhaltens auf einer Karte mit einer Heatmap. *Rechts:* Areas of Interest (blau) auf einer Karte im mobilen Eye Tracking. Blickposition beim rot-grünem Fadenkreuz.

dardwerke Duchowski (2007) und Holmqvist et al. (2011) empfohlen, auf denen auch die folgenden Absätze beruhen.

Grundsätzlich kommen heute zwei Arten von Eye Tracking-Systemen zum Einsatz (siehe Abb. 1): Am Monitor angebrachte „Remote“-Eye Tracker sowie am Kopf getragene mobile Eye Tracker. Beide Systeme nehmen durch ein Kamerasystem das Auge des Nutzers auf und ermitteln durch Bildverarbeitungsverfahren die Position der Pupille im Video. Über ein Modell des Auges lässt sich nach einer Kalibrierung die Blickposition der Person berechnen. Zur Verbesserung der Präzision wird die Pupille üblicherweise mit Infrarotilluminatoren beleuchtet und die dabei entstehenden Reflektionen auf der Hornhaut gemessen (sogenannte „Corneal Reflection Method“). Eine höhere Genauigkeit kann außerdem durch binokulare Eye Tracker erzielt werden, die im Gegensatz zu monokularen Systemen beide Augen verfolgen. Dieses Vorgehen vermeidet den Parallaxeffekt, der bei monokularen Systemen zu einer systematischen Ungenauigkeit in Abhängigkeit von der Distanz zwischen Auge und betrachtetem Objekt führt.

Die Ausgabedaten eines Eye Trackers bestehen aus einer mit Zeitpunkt versehenen geordneten Folge von (x,y)-Koordinaten im jeweiligen Referenzsystem des Eye Trackers. Bei Remote-Eye Trackern sind das also üblicherweise Pixelkoordinaten auf dem Bildschirm, während ein mobiler Eye Tracker Koordinaten im sich bewegenden Videobild der am Kopf befestigten Umgebungskamera ermittelt. Die automatische Ermittlung der betrachteten Information ist somit bei Remote-Eye Trackern weitaus einfacher, da hier meist bekannt ist, was an einer gegebenen Bildschirmkoordinate zu einem gegebenen Zeitpunkt sichtbar war. Im mobilen Eye Tracking hingegen ist es zunächst erforderlich, die (x,y)-Videokoordinate einem entsprechenden Objekt in der realen Welt zuzuordnen.

Die Bewegungen des Auges sind die schnellsten Bewegungen, zu denen der menschliche Körper in der Lage ist (bis zu 900 Winkelgrad pro Sekunde). Die Aufnahme rate eines Eye Trackers muss demnach entsprechend hoch sein – bei derzeit üblichen Systemen in der Größenordnung zwischen 30 Hz und 500 Hz. Sehr hohe Aufnahme raten sind hierbei nur für

spezielle Forschungsfragen erforderlich, während im Bereich der blickbasierten Interaktion 30 bis 100 Hz meist ausreichen.

Die menschliche Wahrnehmung geht weitaus langsamer von statten als die Bewegung des Auges; sprich: Nicht jeder gemessene Blick lässt den Rückschluss zu, dass zu diesem Zeitpunkt etwas bewusst wahrgenommen wurde. Daher werden die Blickdaten im nächsten Schritt raum-zeitlich zu sogenannten Fixationen aggregiert (Rayner, 1998). Eine Fixation liegt dann vor, wenn das Auge für einen bestimmten Zeitraum relativ still bleibt (siehe Abb. 2, links). Die zur Berechnung von Fixationen verwendeten Algorithmen und Parameter sind je nach Eye Tracker unterschiedlich und Gegenstand aktueller Forschung. Rayner (1998) beispielsweise listet Fixationsdauern zwischen 225 und 400 Millisekunden, während aktuellere Arbeiten auch kürzere Fixationsdauern um die 100 Millisekunden verwenden (siehe auch Kiefer et al., 2014, Abschnitt 3.2.2). Die Bewegungen zwischen verschiedenen Fixationen werden Sakkaden genannt (blaue Linien in Abb 2., links).

Die sich ergebende Sequenz aus Fixationen, Sakkaden und eventuellen Zeiträumen mit geschlossenem Auge (Zwinkern) ist die Basis für weitere Analysen, für die je nach Forschungsfrage unterschiedliche Ansätze zur Anwendung kommen:

- *Visuelle Analyse*: Zur Hypothesengenerierung oder Exploration werden die Daten häufig zunächst als „Heatmap“ (siehe Abb. 2, Mitte) oder „Scanpath“ dargestellt. Auch komplexere „Visual Analytics“-Methoden für Trajektorien können hier zum Einsatz kommen (Andrienko et al., 2012).
- *Inhaltsbasierte Analyse*: Der Stimulus wird mit sogenannten Areas of Interest (AOI) strukturiert (meist Polygone, siehe Abb. 2, rechts) und verschiedene Statistiken über der raum-zeitlichen Fixationssequenz bezüglich dieser AOI berechnet. Beispielsweise könnte das Interesse an einem AOI durch die prozentuale Betrachtungsdauer im Verhältnis zur gesamten Aufnahmedauer quantifiziert werden. Anstatt manuell vom Analysten gezeichneter Polygone können bei Karten auch direkt die zugrunde liegenden Geodaten verwendet werden, was die Automatisierung der Analyse erleichtert und gerade bei blickbasierter Interaktion mit interaktiven Karten unabdingbar ist (sog. „Gaze Map Matching“, Kiefer und Giannopoulos, 2012).
- *Analyse der Blickbewegungen*: Die Fixationssequenz wird – unabhängig vom betrachteten Inhalt – durch raum-zeitliche Kennzahlen beschrieben, wie beispielsweise der mittleren Fixationsdauer, der Anzahl an Zwinkern pro Minute oder Kennzahlen zur geometrischen Struktur der Polylinie. Auch hieraus lassen sich interessante Schlüsse ziehen, insbesondere auf die kognitive Last des Nutzers (Rayner, 1998) oder auf seine Suchprozesse (Kiefer et al., 2013).

Zwei Trends zeichnen sich derzeit auf dem Markt für Eye Tracking ab: Die steigende Outdoor-Fähigkeit der mobilen Eye Tracking-Systeme sowie die allmähliche Verbreitung von Eye Tracking-Technologien auf dem Massenmarkt. So gibt es heute neben den professionellen, sehr genauen und relativ teuren Eye Tracking-Systemen¹, auch Blickerfassung in

¹ Beispiele: Tobii (<http://www.tobii.com/>), SMI (<http://www.smivision.com/en.html>), Ergoneers (<http://www.ergoneers.com/>). [Alle Referenzen zuletzt besucht am 30.11.2014.]

Mobiltelefonen, Notebooks, sowie preiswerte Remotesysteme². Auch für Augmented Reality-Brillen – wie das derzeit in Europa noch nicht erhältliche Google Glass – ist Eye Tracking eine naheliegende und wahrscheinliche Erweiterung³.

Obwohl diese Systeme derzeit noch sehr ungenau sind, lassen die Marktentwicklungen erwarten, dass der Einsatz blickbasierter Interaktion bald nicht mehr eine Frage der Hardwareverfügbarkeit oder des Preises ist, sondern der intelligenten darauf aufbauenden Interaktionskonzepte sowie der Akzeptanz beim Nutzer. Die Entwicklung solcher Interaktionskonzepte ist Ziel der Forschungsarbeiten, die in Abschnitt 3 überblicksweise vorgestellt werden.

2.2 Eye Tracking zur Analyse von Karten und GIS

Die Erstellung kognitiv motivierter blickbasierter Nutzerschnittstellen beruht auf einem Verständnis der visuellen Wahrnehmung und kognitiven Prozesse des Menschen beim Umgang mit Karten, GIS und in mobilen Entscheidungssituationen. Ein solches Verständnis ist und war das Ziel von Eye Tracking-Studien, über die im Folgenden ein kurzer Literaturüberblick gegeben werden soll. Im Gegensatz zur blickbasierten Interaktion (Abschnitt 3) werden die Blickdaten bei diesen Studien nicht während der Interaktion, sondern erst im Nachhinein analysiert.

Steinke (1987) gibt einen Überblick über frühe Eye Tracking-Studien in der Kartografie. Sein Hauptinteresse gilt dem Zusammenhang zwischen kartografischen Designvariablen und den oben erwähnten raum-zeitlichen und inhalts-unabhängigen Maßzahlen für Blickverhalten als Ausdruck der kognitiven Belastung. Die Nutzerbarkeitsanalyse eines tatsächlichen GIS rückt erst sehr viel später in den Fokus der Forschung (Çöltekin et al., 2010, Çöltekin et al., 2009). Eye Tracking ist hier ein Mittel zum Auffinden möglicher Erklärungen von Effizienz- und Effektivitätsunterschieden zwischen verschiedenen Nutzergruppen. Die typischen funktionalen Schnittstellenbereiche eines GIS (z.B. der Kartenbereich, der Selektionsbereich etc.) werden hierbei als AOI in einer inhaltsbasierten Analyse verwendet. Über Clusteranalysen auf AOI-Sequenzen lassen sich beispielsweise interessante Muster zu effizienten bzw. ineffizienten Suchstrategien im GIS-Interface auffinden und beschreiben. Das Wissen hierüber ließe sich in einem interaktiven System nutzen, um ineffiziente Strategien des Nutzers in Echtzeit zu erkennen und entsprechend zu assistieren.

Wiederum eher kartografische Fragestellungen stehen im Mittelpunkt der Arbeiten von Ooms et al. (2012a) und Ooms et al. (2012b). Der Vergleich der Suchstrategien von kartografischen Experten und Novizen ergab, dass Experten signifikant mehr und kürzere Fixationen hatten als Novizen (2012a). Daraus lässt sich schlussfolgern, dass sie die Information aus dem Kartenmaterial schneller aufnehmen und schneller verarbeiten konnten. Eye

² Beispiele: Samsungs Galaxy-Serie (<http://www.cnet.com/news/eye-tracking-tech-in-the-samsung-galaxy-s4-say-what/>), HP Ultrabook (<http://blog.laptopmag.com/first-ultrabook-eye-tracking/>), Eye Tribe for 99\$ (<https://theeyetribe.com/>) [Alle Referenzen zuletzt besucht am 30.11.2014.]

³ Patent von Google: <http://www.techradar.com/news/portable-devices/google-patents-eye-tracking-for-google-glass-1091428> [Referenz zuletzt besucht am 30.11.2014.]

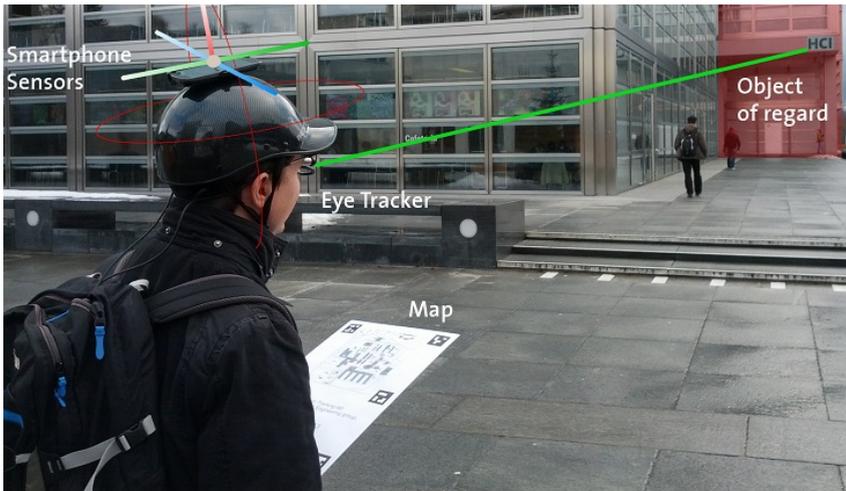


Abb. 3: Mobiles Eye Tracking-kombiniert mit einfachem Headtracking zur Berechnung des Blickvektors auf ein Gebäude.

Tracking-Analysen bei der Evaluation zweier Varianten eines Beschriftungsalgorithmus⁴ ergaben hingegen keine signifikanten Ergebnisse (2012b).

2.3 Visuelle Aufmerksamkeit bei der Wegsuche

Die Wegsuche ist wohl einer der häufigsten kognitiven Prozesse in der mobilen Entscheidungsfindung und derjenige, der am meisten durch mobile Systeme unterstützt wird – sei es durch das Autonavigationssystem oder die Handy-App für Fußgänger. Während bei der Wegsuche zwar (mobile) Karten zum Einsatz kommen, dienen zur Entscheidungsfindung ebenso Landmarken in der Umgebung – etwa ein markantes Gebäude oder eine Beschilderung, die vom Wegsuchenden wahrgenommen und mit der Kartenrepräsentation bzw. dem Gedächtnis verglichen werden. Die hierbei stattfindenden visuellen Suchprozesse und Suchstrategien sind Gegenstand verschiedener Untersuchungen.

Wiener et al. (2011) beispielsweise stellen fest, dass Personen, die ohne weitere Informationen oder Anhaltspunkte an einer Wegkreuzung eine simple links/rechts-Entscheidung treffen müssen, zwar zunächst beide Optionen visuell explorieren, aber gegen Ende des Entscheidungsprozesses tendenziell mehr visuelle Aufmerksamkeit auf die später gewählte Option investieren. Emo (2012) stellt in einem ähnlichen Szenario fest, dass die räumliche Geometrie einer Straßenflucht die räumliche Verteilung der Blicke beeinflusst. Beide verwenden in ihren Studien Remote-Eye Tracker vor einem Computerbildschirm, auf dem räumliche Szenen als Bilder eingeblendet werden. Auch vor einem Monitor arbeiten Gunzelmann et al. (2008), die sich mit kognitiven Strategien zur Rotation räumlicher Darstellungen beschäftigen. Sie konnten zwei Probandengruppen verschiedene Strategien zur Lösung der Rotationsaufgabe beibringen und fanden diese Strategien später im Blickverhalten wieder.

Die geschilderten Laborstudien sind jedoch in einem entscheidenden Aspekt eingeschränkt: Der Wegsuchende sitzt vor einem Rechner, anstatt sich physisch durch die Umgebung

fortzubewegen. Mobile Eye Tracking-Technologie hingegen erlaubt die Aufnahme des visuellen Suchverhaltens auch während der Wegsuche in realen urbanen Räumen. Durch Datensynchronisation mit einem Positionierungssensor entsteht dadurch ein reichhaltiger Datensatz, aus dem sich nicht nur herausfinden lässt, wohin jemand geschaut hat, sondern auch, welchen Weg er dabei zurückgelegt hat (Location-Aware Mobile Eye Tracking, Kiefer et al., 2012).

Zwei Studien im Stadtzentrum von Zürich demonstrieren die Möglichkeiten dieses Ansatzes: In einer Studie zur Selbstlokalisierung (Kiefer et al., 2014) wurden Personen in einer für sie fremden Umgebung mit der Aufgabe konfrontiert, auf einer Karte ihre aktuelle Position einzuzeichnen. Hierfür war ein Abgleich der real sichtbaren Landmarken in der Umgebung mit Symbolen auf der Karte notwendig. Die Studie ergab einen Zusammenhang zwischen den betrachteten AOI auf der Karte und den betrachteten Landmarken in der Umgebung, sowie der Korrektheit der eingezeichneten Position. Erfolgreiche Teilnehmer konnten offensichtlich die in der Situation hilfreiche Information aus der Karte und der Umgebung extrahieren und zur Lösung der Aufgabe verwenden. In einer weiteren Outdoor-Studie wurden Eye Tracking-Daten zur Messung der Komplexität von Entscheidungen an Straßenkreuzungen verwendet (Giannopoulos et al., 2014). Je mehr visuelle Aufmerksamkeit eine Person auf der Karte benötigte und je häufiger er oder sie zwischen Karte und Umgebung wechselte, desto höher war die Komplexität der Entscheidungssituation. Es zeigt sich also auch bei der Wegsuche, dass das Blickverhalten als Anzeichen für (Probleme bei) Suchstrategien oder als Prädiktor für gewählte Entscheidungen dienen kann.

Wie in Abschnitt 2.1 beschrieben, besteht eine der größten Herausforderungen bei mobilen Eye Tracking-Studien in der Abbildung der Eye Tracking-Daten auf die wahrgenommenen Objekte. In den Studien in der Innenstadt Zürichs wurden hierzu Ansätze aus der Bildverarbeitung verwendet: Blicke auf die Karte wurden voll-automatisch durch visuelle schwarze Marker verortet (vgl. Abb. 2, rechts), während Blicke auf die Landmarken in der Umgebung durch einen halb-automatischen Prozess bestehend aus Annotation und bildverarbeitenden Algorithmen analysiert wurden (vgl. hierzu Kiefer et al., 2014, Abschnitt 3.2). Mit Methoden der Bildverarbeitung erkennen auch Paletta et al. (2014) die farbige Smartphonehülle im Eyetrackingvideo und können dabei auf Marker verzichten. Für Landmarken wurde außerdem in einer Masterarbeit ein Ansatz vorgeschlagen, der Blickvektoren, entstanden durch Headtracking mit einfachen Smartphonesensoren, mit einem 3D-Stadtmodell in einem Geowebsservice verschneidet (Haesler, 2014) (vgl. Abb. 3).

3 Blickbasierte Interaktion mit Geoinformation

3.1 Explizite blickbasierte Interaktion

Viele der aktuellen Eye Tracking-Systeme erlauben den Zugriff auf die Daten in Echtzeit, d.h. während der Blickerfassung. Die Hersteller der Eye Tracker bieten Softwareentwicklern hierfür bestimmte Schnittstellen und Protokolle an, über die der Datenzugriff erfolgt. Ein interaktives Programm kann somit direkt auf die Blickeingabe reagieren und die Nutzerschnittstelle an das Blickverhalten anpassen. In der allgemeinen Forschung zur Mensch-Computer-Interaktion wird dieses Prinzip als *blickbasierte Interaktion* („gaze-based interaction“) in einer Reihe von Anwendungsgebieten verwendet, beispielsweise beim „Gaze

Typing“, das die Eingabe von Text durch den Blick ermöglicht (Majaranta et al., 2009). Dies ist nicht nur zur Assistenz für Computernutzer mit körperlichen Einschränkungen interessant, sondern auch in (mobilen) Einsatzszenarien, in denen Hand- oder Spracheingabe nicht möglich oder nicht adäquat sind.

Doch wie lässt sich blickbasierte Interaktion für Karten oder komplexe GIS umsetzen? Ein naheliegender Ansatz besteht darin, Blicke auf den Rand der Karte als Verschiebe-Operation zu interpretieren, während beim Zoomen der Blick den Zoommittelpunkt festlegt und durch eine weitere Aktion gestartet wird, etwas einen Klick, eine Spracheingabe oder ein Zwinkern. Einen ähnlichen Ansatz beschreiben Stellmach und Dachzelt (2012) für einen virtuellen Globus.

An diesem Beispiel lässt sich ein klassisches Problem der blickbasierten Interaktion verdeutlichen, das *Midas Touch Problem* (Jacob und Karn, 2003): Da der Blick bei der Karteninteraktion nicht primär zur Interaktion dient, sondern zur Informationsaufnahme, kann bei rein blickbasierter Interaktion schnell eine vom Nutzer nicht geplante Interaktion gestartet werden; etwa wenn er oder sie aus Versehen zwinkert oder längere Zeit an den Rand der Karte schaut. Daher wird blickbasierte Interaktion häufig *multimodal* eingesetzt, d.h. in Kombination mit anderen Eingabemodalitäten. Der Blick wird dabei typischerweise als Ersatz für einen Mauszeiger verwendet. Eine Herausforderung besteht hierbei in einer geeigneten Wahl der Kombination verschiedener Interaktionsmodalitäten, so dass das Gesamtsystem für den Nutzer verständlich, erlernbar und effizient sowie effektiv einsetzbar ist. Bisher findet sich kein Ansatz im GIS-Bereich, der dieses für komplexere Interaktionen als einfache Kartenoperationen umsetzt.

Das beschriebene Interaktionsprinzip fällt in die Klasse der *expliziten* blickbasierten Interaktion (Schmidt, 2000), die sich dadurch auszeichnet, dass der Nutzer seinen Blick gezielt zur Interaktion einsetzt, um bestimmte Adaptionen der Nutzerschnittstelle zu erreichen. Noch weitaus komplexere Interaktionsmöglichkeiten bietet hingegen die *implizite* blickbasierte Interaktion, die im folgenden Abschnitt diskutiert wird.

3.2 Implizite blickbasierte Interaktion

Implizite Interaktion (Schmidt, 2000) liegt dann vor, wenn das System das Verhalten des Nutzers aufzeichnet, interpretiert und zu einem späteren Zeitpunkt für eine Anpassung der Nutzerschnittstelle verwendet, ohne dass der Nutzer diese Adaption gezielt geplant hatte. Hier lässt sich eine Parallele zur Intentionserkennung aus Trajektorien (z.B. GPS-Spuren) ziehen (Kiefer, 2012), bei der an Hand eines Nutzermodells und dem aufgezeichneten Positionsverlauf des Nutzers gefolgert wird, welche Pläne dieser aktuell verfolgen könnte und welche Information daher hilfreich sein könnten (intelligente mobile Assistenzsysteme).

Ein Beispiel für implizite blickbasierte Interaktion ist der Ansatz GeoGazemarks (Giannopoulos et al., 2012). In diesem Interaktionsansatz wird der Blick des Nutzers auf der Karte aufgezeichnet, als Geo-Koordinaten gespeichert und zu einem späteren Zeitpunkt als aggregierte History dargestellt. Beim Herauszoomen sieht der Nutzer somit durch Symbole, in welchen Regionen der Karte er oder sie vorher aktiv war. Dies hilft der Orientierung auf kleinen Bildschirmen, auf denen es häufig schwierig ist, den Überblick über den Kartenkontext zu behalten und zu Kartenpositionen zurück zu finden, die man vorher ex-

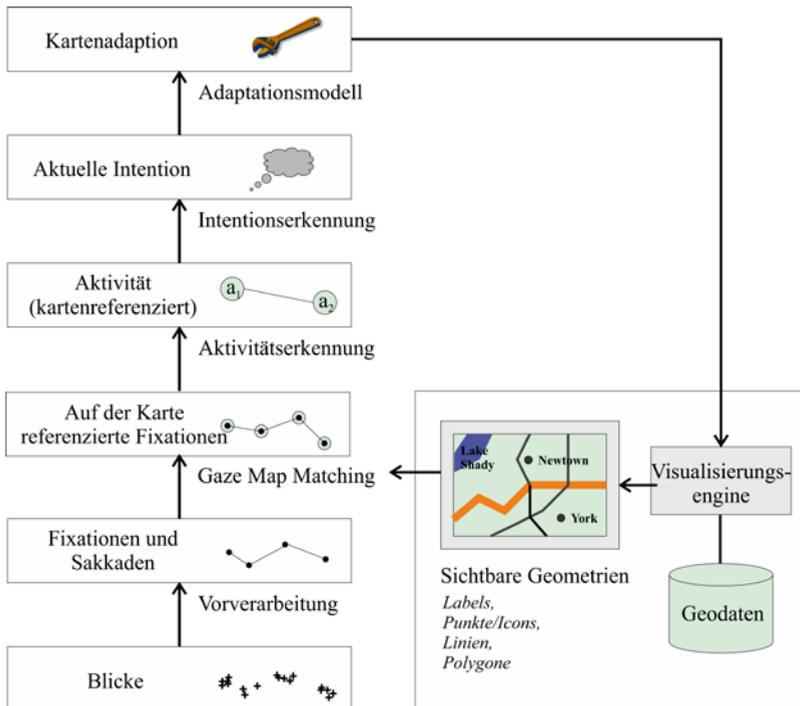


Abb. 4: Mehrschichtenarchitektur zur Erkennung von Intentionen aus dem Blickverhalten eines Kartennutzers

ploriert hatte. Ein zweiter Ansatz verwendet die raum-zeitlichen Eigenschaften des Blickverhaltens, um daraus die Aktivität des Nutzers zu erkennen (Kiefer et al., 2013). Technisch umgesetzt wurde dies mit einem Verfahren des maschinellen Lernens (Support Vector Machine). Der gelernte Klassifikator erreichte hierbei eine Genauigkeit von ca. 78% für die Aktivitäten freie Exploration, globale Suche, Routenplanung, fokussierte Suche, Verfolgen einer Linie und Vergleich von Polygonen.

4 Ausblick

Die erfolgsversprechenden Ergebnisse zur Aktivitätserkennung lassen hoffen, dass sich dieser Ansatz auch zur Erkennung höherer kognitiver Zustände, wie Intentionen oder Pläne, verwenden lässt. Eine mögliche Architektur in diese Richtung wird in Abbildung 4 vorgestellt: Ausgehend von den Messwerten – den Blicken – werden zunächst Fixationen und Sakkaden berechnet. Mit den oben erwähnten Ansätzen zum Gaze Map Matching (Kiefer und Giannopoulos, 2012) und den Informationen über den in jenem Moment sichtbaren Kartenfeatures werden diese Fixationen den Elementen der Karte zugeordnet, beispielsweise einer Straße oder einem See. Aus einer Sequenz mehrerer solcher kartenreferenzierter Fixationen könnte dann im nächsten Schritt durch Analyse ihrer raum-zeitlichen Abfolge eine kartenreferenzierte Aktivität ermittelt werden. Offen ist hier noch, wie die Fixationssequenz algorithmisch segmentiert werden kann, so dass tatsächlich separate Aktivitäten

entstehen. Der bisher noch nicht angegangene Schritt wäre dann die Intentionserkennung, welche mehrere Aktivitäten zu einer aktuellen Intention zusammenfasst. Basierend auf einem Adaptionmodell wird schließlich die Karte verändert, um dem Nutzer bei seiner aktuellen Intention behilflich zu sein. Eine interessante Frage ist diesbezüglich, wie die Kartenadaption im nächsten Schritt das visuelle Suchverhalten des Nutzers beeinflusst, und somit auch Einfluss auf alle im Anschluss erkannten Aktivitäten und Intentionen hat.

Zukünftige Schritte im Bereich Wegsuche werden sich damit beschäftigen, die Outdoor- und Echtzeitfähigkeit der beschriebenen Ansätze zur blickbasierten Interaktion mit 3D-Objekten sicherzustellen. Gerade im Bereich der Berechnung des betrachteten Objekts besteht hier noch Verbesserungspotenzial. Die Vision in diesem Bereich wäre ein blickbasierter Fußgängernavigationsassistent, der drei Arten von Informationen berücksichtigt: Das Blickverhalten auf der Karte, das Blickverhalten auf 3D-Objekte in der Umgebung (z.B. Gebäude) sowie die Position und Fortbewegung des Nutzers (Giannopoulos et al., 2013). Im Sinne des zu Beginn eingeführten kognitiven Engineering sind hierfür auch noch weitere Forschungsarbeiten im Bereich kognitiver Modelle bei der Wegsuche nötig, die speziellen Fokus auf das Blickverhalten des Wegsuchers legen.

Die Multimodalität der beschriebenen Ansätze ist ein weiteres Forschungsfeld. Beispielsweise kann blickbasierte Interaktion auch mit mehreren Bildschirmen gleichzeitig erfolgen (Zhang et al., 2013) oder durch haptisches Feedback unterstützt werden (Kangas et al., 2014).

Danksagung

Ich möchte mich herzlich bedanken bei Herrn Prof. Dr. Martin Raubal und Herrn Ioannis Giannopoulos für die gemeinsame spannende Forschung im GeoGazeLab am Lehrstuhl für Geoinformations-Engineering der ETH Zürich (<http://www.geogaze.org/>).

Literatur

- ANDRIENKO, G., ANDRIENKO, N., BURCH, M. & WEISKOPF, D. 2012. Visual analytics methodology for eye movement studies. *Visualization and Computer Graphics, IEEE Transactions on*, 18, 2889-2898.
- ÇÖLTEKIN, A., FABRIKANT, S. I. & LACAYO, M. 2010. Exploring the efficiency of users' visual analytics strategies based on sequence analysis of eye movement recordings. *International Journal of Geographical Information Systems*, 24, 1559-1575.
- ÇÖLTEKIN, A., HEIL, B., GARLANDINI, S. & FABRIKANT, S. I. 2009. Evaluating the Effectiveness of Interactive Map Interface Designs: A Case Study Integrating Usability Metrics with Eye-Movement Analysis. *Cartography and Geographic Information Science*, 36, 5-17.
- DUCHOWSKI, A. T. 2007. *Eye Tracking Methodology: Theory and Practice*, London, Springer.

- EMO, B. 2012. Wayfinding in real cities: experiments at street corners. In: STACHNISS, C., SCHILL, K. & UTTAL, D. (eds.) *Spatial Cognition VIII*. Berlin Heidelberg: Springer, 461-491.
- GIANNOPOULOS, I., KIEFER, P. & RAUBAL, M. 2012. GeoGazemarks: Providing gaze history for the orientation on small display maps. *Proceedings of the 14th International Conference on Multimodal Interaction (ICMI '12)*. New York, NY, USA: ACM, 165-172.
- GIANNOPOULOS, I., KIEFER, P. & RAUBAL, M. 2013. Mobile Outdoor Gaze-Based GeoHCI. *Proceedings of Geographic Human-Computer Interaction, Workshop at CHI 2013*. Paris, France, 12-13.
- GIANNOPOULOS, I., KIEFER, P., RAUBAL, M., RICHTER, K.-F. & THRASH, T. 2014. Wayfinding Decision Situations: A Conceptual Model and Evaluation. *Geographic Information Science*. Springer, 221-234.
- GUNZELMANN, G., DOUGLASS, S. & KHOOSHABEH, P. 2008. Learning to Orient Using a Map Display: Evidence from Eye Tracking In: HÖLSCHER, C. (ed.) *Spatial Cognition 2008: Poster Presentations (SFB/TR 8 Report No. 016 - 08/2008)*. Freiburg, Germany: Universität Bremen / Universität Freiburg, 85-88.
- HAESLER, S. 2014. *Erstellung und Evaluation einer service-orientierten Systemarchitektur zur ortsbezogenen blickbasierten Interaktion mit 3D-Geoobjekten*. M.Sc. Masterarbeit, ETH Zürich.
- HAKLAY, M. 2010. *Interacting with geospatial technologies*, Wiley Online Library.
- HECHT, B., SCHÖNING, J., HAKLAY, M., CAPRA, L., MASHHADI, A. J., TERVEEN, L. & KWAN, M.-P. Geographic human-computer interaction. CHI '13 Extended Abstracts on Human Factors in Computing Systems, 2013 Paris, France. 2479637: ACM, 3163-3166.
- HOLMQVIST, K., NYSTRÖM, M., ANDERSSON, R., DEWHURST, R., JARODZKA, H. & VAN DE WEIJER, J. 2011. *Eye Tracking - A Comprehensive Guide To Methods And Measures*, New York, Oxford University Press.
- JACOB, R. J. K. & KARN, K. S. 2003. Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises. In: HYONA, RACHACH & DEUBEL (eds.) *The Mind's Eye: Cognitive and Applied Aspects of Eye Movement Research*. Oxford, England: Elsevier, 573-605.
- KANGAS, J., RANTALA, J., MAJARANTA, P., ISOKOSKI, P. & RAISAMO, R. Haptic feedback to gaze events. Proceedings of the Symposium on Eye Tracking Research and Applications, 2014. ACM, 11-18.
- KIEFER, P. 2012. *Mobile Intention Recognition*, Springer.
- KIEFER, P. & GIANNOPOULOS, I. 2012. Gaze map matching: mapping eye tracking data to geographic vector features. *Proceedings of the 20th International Conference on Advances in Geographic Information Systems*. New York, NY, USA: ACM, 359-368.
- KIEFER, P., GIANNOPOULOS, I. & RAUBAL, M. 2013. Using Eye Movements to Recognize Activities on Cartographic Maps. *Proceedings of the 21th SIGSPATIAL International Conference on Advances in Geographic Information Systems*. New York, NY, USA: ACM, 498-501.
- KIEFER, P., GIANNOPOULOS, I. & RAUBAL, M. 2014. Where am I? Investigating map matching during self-localization with mobile eye tracking in an urban environment. *Transactions in GIS*, 18, 660-686.

- KIEFER, P., STRAUB, F. & RAUBAL, M. 2012. Towards Location-Aware Mobile Eye Tracking. *Proceedings of the 2012 Symposium on Eye-Tracking Research and Applications (ETRA 2012)*. New York, NY, USA: ACM, 313-316.
- MAJARANTA, P., AHOLA, U.-K. & SPAKOV, O. Fast gaze typing with an adjustable dwell time. Proceedings of the 27th international conference on Human factors in computing systems, 2009 Boston, MA, USA. ACM, 357-360.
- MENG, L. 2005. Egocentric design of map-based mobile services. *The Cartographic Journal*, 42, 5-13.
- OOMS, K., DE MAEYER, P., FACK, V., VAN ASSCHE, E. & WITLOX, F. 2012a. Interpreting maps through the eyes of expert and novice users. *International Journal of Geographical Information Science*, 26, 1773-1788.
- OOMS, K., DE MAEYER, P., FACK, V., VAN ASSCHE, E. & WITLOX, F. 2012b. Investigating the Effectiveness of an Efficient Label Placement Method Using Eye Movement Data *The Cartographic Journal*, 49, 234-246.
- PALETTA, L., NEUSCHMIED, H., SCHWARZ, M., LODRON, G., PSZEIDA, M., LULEY, P., LADSTÄTTER, S., DEUTSCH, S. M., BOBETH, J. & TSCHELIGI, M. Attention in mobile interactions: gaze recovery for large scale studies. CHI'14 Extended Abstracts on Human Factors in Computing Systems, 2014. ACM, 1717-1722.
- RAUBAL, M. 2009. Cognitive Engineering for Geographic Information Science 3(3). *Geography Compass*, 3, 1087-1104.
- RAUBAL, M. & PANOV, I. 2009. A formal model for mobile map adaptation. *Location Based Services and TeleCartography II*. Springer, 11-34.
- RAYNER, K. 1998. Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372-422.
- REICHENBACHER, T. 2001. Adaptive concepts for a mobile cartography. *Journal of Geographical Sciences*, 11, 43-53.
- SCHMIDT, A. 2000. Implicit human computer interaction through context. *Personal technologies*, 4, 191-199.
- STEINKE, T. R. 1987. Eye movement studies in cartography and related fields. *Cartographica: The International Journal for Geographic Information and Geovisualization*, 24, 40-73.
- STELLMACH, S. & DACHSEL, R. Investigating gaze-supported multimodal pan and zoom. 2012 Symposium on Eye Tracking Research & Applications (ETRA 2012), 2012 Santa Barbara, CA, USA. ACM, 29-36.
- VERTEGAAL, R. 2003. Attentive User Interfaces. *Communications of the ACM*, 46, 31-33.
- WIENER, J. M., HÖLSCHER, C., BÜCHNER, S. & KONIECZNY, L. 2011. Gaze behaviour during space perception and spatial decision making. *Psychological Research*, 1-17.
- ZHANG, Y., BULLING, A. & GELLERSEN, H. Sideways: A gaze interface for spontaneous interaction with situated displays. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, 2013. ACM, 851-860.