



MASTERARBEIT

Geo-spezifische Metadaten in Bildern Gewinnung, Speicherung und Anwendungsszenarien

Ausgeführt am Institut für
Multimedia Information Systems
der Universität Wien

unter der Anleitung von Prof. Dr. Wolfgang Klas und Mag. Wolfgang Jochum als
verantwortlich mitwirkendem Universitätsassistenten

durch

Robert Neuner
Alserstraße 22 / 13
1090 Wien

Datum

Unterschrift (Student)

Eidesstattliche Erklärung

Ich erkläre an Eides statt, dass ich die vorliegende Arbeit selbständig und ohne fremde Hilfe verfasst, andere als die angegebenen Quellen nicht benützt und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

<Ort, Datum>

<Eigenhändige Unterschrift>

Abstract

The number of internet platforms which allow their users to search for their pictures according to the location where the picture was taken is constantly increasing. However, with this kind of search users can only determine where a picture was taken, but have no possibility to directly search for objects on the photos. In order to eliminate this limitation, this thesis analyses in how far a picture's field of view can be used in order to identify an object depicted on a photo. Therefore FoVA – the Field of View Analyzer – was developed. FoVA allows searching for the location where the picture was taken and also utilizes the field of view of a picture in order to identify the depicted objects. A prerequisite for FoFA is, that the location where the picture was taken and the direction of photographing have already been recorded and linked. The linking of these two sets of data, also called geo-referencing, is mainly done manually and therefore quite time intensive. This thesis also examines in how far this process can be automated and which technical conditions have to be fulfilled to be able to do so. For this reason, the application GPSTool was developed which enables users to utilize their PDA to record pictures including geodata in a simple way.

Zusammenfassung

Immer mehr Internetplattformen ermöglichen es, Fotos anhand ihres Aufnahmestandortes zu suchen. Die ortsbasierte Suche ermöglicht jedoch nur festzustellen, wo ein Bild gemacht wurde. Sie bietet keine Möglichkeit, nach den Objekten selbst zu suchen, die auf dem Foto zu sehen sind. Um diese Einschränkung aufzuheben, untersucht diese Arbeit, inwieweit das Sichtfeld eines Fotos dazu verwendet werden kann, um auf dem Foto abgebildete Objekte zu identifizieren. Im Zuge dieser Arbeit wurde FoVA entwickelt, der Field of View Analyzer. Mittels FoVA ist es möglich, das Sichtfeld eines Fotos zu nutzen um nicht nur nach dem Aufnahmeort, sondern auch nach den auf dem Foto abgebildeten Objekten zu suchen. Voraussetzung dafür ist, dass die Fotos bereits im Vorfeld mit ihrem Standort und ihrer Aufnahmerichtung verknüpft worden sind. Die Verknüpfung, auch Geo-Referenzierung genannt, geschieht meist manuell und ist dadurch sehr zeitaufwändig. Diese Arbeit untersucht deshalb auch, inwieweit dieser Prozess automatisiert werden kann und welche technischen Voraussetzungen dafür erfüllt sein müssen. Aus diesem Grund wurde im Zuge dieser Arbeit auch die Applikation GPSTool entwickelt, mit der ein Anwender seinen PDA nutzen kann, um auf bequeme Weise Fotos inklusive ihrer Geodaten aufzunehmen.

Inhaltsverzeichnis

1. Einleitung	7
2. Metadaten	11
2.1 Definition.....	11
2.1.1 Semantische Daten	12
2.2 Geografische Metadaten	13
2.3 Bildspezifische Metadaten.....	14
2.3.1 Interne Speicherung	15
2.3.2 Externe Speicherung	17
2.3.3 Evaluierung	19
3. Technische Voraussetzungen zur Gewinnung geografischer Metadaten	21
3.1 Das NAVSTAR GPS-System	22
3.1.1 Aufbau und Funktionsweise.....	22
3.1.2 Signaleigenschaften und Störeinflüsse.....	24
3.1.3 Das NMEA Protokoll.....	27
3.1.4 Evaluierung	29
3.2 Elektronischer Kompass.....	30
3.2.1 Aufbau und Funktionsweise.....	30
3.2.2 Einschränkungen bei der magnetischen Richtungsmessung.....	30
3.2.3 Evaluierung	31
4. Metadaten zur Verbesserung des Suchprozesses	33
4.1 Geo-Referenzierung	33
4.1.1 Manuelle Referenzierung	34
4.1.2 Automatische Referenzierung.....	34

4.1.3	Semiautomatische Referenzierung.....	35
4.1.4	Evaluierung	37
4.2	Das Kamerasichtfeld	38
4.2.1	Evaluierung	43
5.	Beispielapplikation zur Verbesserung des Suchprozesses	44
5.1	Annahmen und Limitationen.....	45
5.2	FoVA – The Field of View Analyzer.....	46
5.2.1	Programmfunktionen	46
5.2.2	Grundlagen zur Nutzung des Sichtfeldes.....	50
5.2.3	Gewonnene Erkenntnisse	51
5.2.4	Verbesserungsmöglichkeiten	55
5.3	GPSTool	55
5.3.1	Programmbeschreibung.....	56
5.3.2	Gewonnene Erkenntnisse	60
5.4	Geo-Referenzierung	63
5.4.1	Modifizierungen.....	64
5.5	Datenspeicherung	64
5.5.1	MetadatenSpeicherung von Bildern.....	64
5.5.2	MetadatenSpeicherung von Sehenswürdigkeiten	66
6.	Zusammenfassung.....	69
7.	Ausblick.....	72
	Abbildungsverzeichnis.....	74
	Tabellenverzeichnis.....	75
	Literaturverzeichnis.....	76
	Anhang	78

Kapitel 1

Einleitung

Für uns Menschen wird es immer selbstverständlicher, Erinnerungen in digitaler Form ablegen zu können. Digitale Foto- und Videokameras helfen uns dabei, die schönen und wichtigen Momente des Lebens einzufangen, bieten uns jedoch kaum Unterstützung bei der Organisation der aufgenommenen Daten. Um die Erinnerungen auch noch nach Jahren abrufen zu können, genügt es nicht, sie einfach nur auf einer Festplatte oder einer DVD abzuspeichern. Viel entscheidender ist es, Methoden und Techniken zu entwickeln, die es uns auf einfache Weise ermöglichen, die digitalen Objekte, die wir mit einer Erinnerung verknüpfen, in Sekundenschnelle wiederzufinden.

Diese Arbeit entstand aus dem Bedürfnis heraus, die eigenen digitalen Erinnerungen, insbesondere digitale Fotos, die sich im Laufe der Jahre angesammelt haben, besser zu organisieren und die Suche nach bestimmten Fotos zu erleichtern. Studien haben gezeigt, dass Menschen ihre Erinnerungen nach vier Kriterien einordnen: dem Ereignis, dem Ort des Geschehens, dem damit assoziierten Thema und der Zeit [1]. Da die meisten Aufnahmegерäte den Zeitpunkt der Aufnahme erfassen können, besteht die einfachste Möglichkeit zur Organisation in der chronologischen Sortierung der Daten. Um eine bestimmte digitale Erinnerung zu finden, genügt es, den Zeitraum, oder besser den genauen Zeitpunkt, des Ereignisses zu kennen. Doch für viele von uns ist es schwierig, sich an ein genaues Datum zu erinnern, besonders, wenn es schon mehrere Jahre zurück liegt. Der nächste Schritt, zu einer besseren Organisation der Erinnerungen, besteht anschließend in der Verwendung von

Schlagwörtern, die das Ereignis, das Thema und den Ort näher beschreiben können. Bei einem umfangreichen Datenbestand ist die manuelle Vergabe von Schlagwörtern ein sehr mühsamer Prozess und für den Einzelnen mit hohem Aufwand verbunden. Wird die Arbeit jedoch auf einen großen Benutzerkreis aufgeteilt, dann zeigen Internetplattformen wie Flickr¹, dass die Vergabe von Schlagwörtern auch bei umfangreichen Datenmengen innerhalb einer großen Benutzergruppe gut funktionieren kann. Das Konzept der Arbeitsteilung eignet sich zwar für viele Daten, im Bereich der digitalen Erinnerungen ist es aber ungeeignet. Bei digitalen Erinnerungen handelt es sich vor allem um private Daten, die man vielleicht einem kleinen Benutzerkreis, aber nicht der Öffentlichkeit zugänglich machen will.

Da, wie bereits erwähnt, auch der Ort eines Geschehnisses ein wichtiges Kriterium zur Einteilung von Erinnerungen darstellt, ist derzeit eine weitere Möglichkeit zur Organisation von Bildern im Entstehen. Fotodienste im Internet, wie FlickrMaps² oder Panoramia³, ermöglichen es den Nutzern, ihre Fotos anhand des Aufnahmeortes auf einem virtuellen Globus zu platzieren. Sie ermöglichen dadurch die Suche nach Bildern anhand ihrer Aufnahmeposition. Diesen Anwendungen fehlt aber eine Möglichkeit zur automatischen Verknüpfung von Foto und Standort. Die Referenzierung muss manuell erfolgen, was bei einem großen Bildbestand zu einem beträchtlichen Arbeitsaufwand führt. Die Möglichkeiten, die sich durch Verknüpfung von Bild und Ortsdaten für den Anwender eröffnen sind jedoch vielfältig. Die Fotos der letzten Urlaubsreise können auf anschauliche Weise mit einer Stadt, einem Strand oder anderen Orten verknüpft werden und erlauben dadurch eine ortsbasierte Suche und eine übersichtliche Darstellung der Bilder.

Aufgrund der Tatsache, dass der Suchprozess nur den Aufnahmeort des Bildes berücksichtigt, findet genau genommen, nur eine Umkreissuche statt, die alle Fotos in einem bestimmten Umkreis des gesuchten Ortes zurückliefern kann. Diese Art der Suche berücksichtigt jedoch in keinster Weise die Objekte, die auf dem Foto selbst abgebildet sind. Gerade bei der Suche nach Fotos geht es dem Anwender aber in

¹ Flickr: <http://www.flickr.com>

² FlickrMaps: <http://www.flickr.com/maps>

³ Panoramia: <http://www.panoramia.com>

erster Linie darum festzustellen, was auf dem Foto zu sehen ist und nicht wo es aufgenommen wurde.

Angesichts dieser bestehenden Einschränkungen, ist es das Ziel dieser Arbeit, zu untersuchen, inwieweit die Suche nach bestimmten Fotos durch die Auswertung zusätzlicher Informationen des Bildes verbessert werden kann. Um dieses Ziel zu erreichen, muss innerhalb dieser Arbeit eine Basis geschaffen werden, auf der die Untersuchungen aufsetzen können. Zu Beginn wird daher zuerst nach Techniken gesucht, die es ermöglichen, geografische Informationen zu einem Foto speichern zu können. Anschließend wird aufgezeigt, wie diese Informationen gewonnen werden können und welche technischen Voraussetzungen dafür erfüllt sein müssen. Sobald Gewinnung und Speicherung der Daten besprochen wurden, kann untersucht werden, in welcher Weise diese Informationen mit einem Foto verknüpft werden können. Neben der bereits erwähnten manuellen Referenzierung wird an dieser Stelle auch auf automatische Methoden eingegangen. Schlussendlich wird anhand einer Beispielapplikation untersucht, inwieweit sich die Suche nach bestimmten Fotos durch die Auswertung zusätzlicher Daten im Vergleich zu einer reinen Umkreissuche verbessern lässt.

Um die untersuchten Methoden zur Speicherung, Gewinnung und Referenzierung beurteilen zu können, ist es notwendig, bereits im Vorfeld einige Anforderungen und Limitationen zu definieren:

- Der Begriff der digitalen Erinnerungen kann weit gefasst werden. Diese Arbeit beschäftigt sich ausschließlich mit digitalen Bildern, manche der behandelten Konzepte und Ergebnisse können aber eventuell auch für die bessere Organisation von Video- oder Audiodaten genutzt werden.
- Das Mindestmaß an geografischer Information, welche zu einem Foto gespeichert werden sollen, ist seine Aufnahmeposition und seine Aufnahmerichtung. Die Position sollte als Koordinatenpaar in Längen- und Breitengrad gemäß dem WGS-84 Standard definiert werden können. Dieses Bezugssystem ermöglicht einen weltweiten Einsatz und ist stark verbreitet. Die Aufnahmerichtung sollte in Grad definiert werden können.

- Die verwendeten Testbilder werden ausschließlich unter freiem Himmel aufgenommen und zeigen Aufnahmen von berühmten Gebäuden, Kirchen und anderen Sehenswürdigkeiten. Das abgebildete Objekt sollte sich in der Bildmitte befinden und aus Gründen der Sichtbarkeit aus nicht weniger als 10 bzw. nicht mehr als 200 Metern Entfernung fotografiert werden. Abbildung 1 (A) zeigt diesen Sachverhalt in einer Grafik.
- Die Genauigkeit der Positionsbestimmung sollte aufgrund der Anforderungen an die Testdaten unter 10 Metern liegen. Damit ist sichergestellt, dass die ermittelte Position auch im schlechtesten Fall niemals hinter dem fotografierten Objekt liegt, siehe Abbildung 1 (B).
- Die Genauigkeitsabweichung der Richtungsbestimmung sollte, ausgehend von einem 360° Kreis, nicht über 2° Grad liegen. Bei einer maximalen Entfernung von 200 Metern, ergibt sich dadurch für das Objekt in der Bildmitte eine maximale Verschiebung nach links oder rechts von ca. 7 Metern. Diese Anforderung soll sicherstellen, dass die Aufnahmerichtung immer in Richtung Bildmitte zeigt und dadurch optimal mit dem Foto ausgerichtet ist, siehe Abbildung 1 (C).

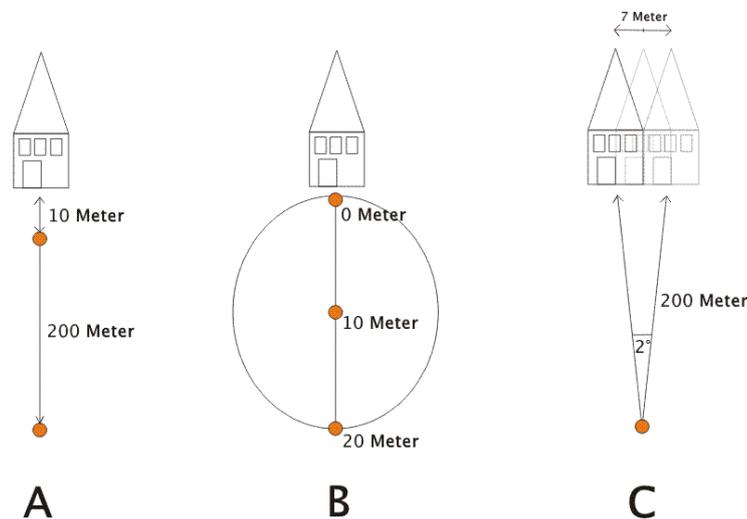


Abbildung 1: Grafische Darstellung der Anforderungen an das Fotomotiv, die Positionsbestimmung und die Richtungsmessung

Die genannten Anforderungen gelten als wichtige Kriterien, um die in dieser Arbeit untersuchten Methoden und gewonnenen Testbilder beurteilen zu können. Im Laufe dieser Arbeit wird daher immer wieder auf sie verwiesen.

Kapitel 2

Metadaten

Um digitale Fotos organisierbar zu machen, hilft es, neben dem Bild selbst zusätzliches Wissen über das Foto und dessen Aufnahmeprozess zu kennen. Wann und wo wurde es aufgenommen? Welche Kamera wurde dafür benutzt? Was ist auf dem Bild zu sehen? Antworten auf diese Fragen sind Informationen, die es erleichtern, ein Foto besser klassifizieren oder beschreiben zu können. Diese zusätzlichen Informationen, die es ermöglichen eine Ressource näher zu beschreiben, werden auch als Metadaten bezeichnet.

Dieses Kapitel definiert zu Beginn die Begriffe „Metadaten“, „geografische Metadaten“ und geht in weiter Folge genauer auf bildspezifische Metadaten ein. Um eine den Anforderungen dieser Arbeit entsprechende Speicherung der notwendigen Metadaten zu realisieren, werden schließlich die derzeit gebräuchlichsten Metadatenstandards der digitalen Fotografie auf ihre Eigenschaften zur Speicherung von geografischen Informationen hin untersucht und bewertet.

2.1 Definition

Metadaten werden oft als Daten über Daten bezeichnet und dienen dazu, eine bestimmte Ressource näher zu beschreiben und zu erklären. Sie sind strukturierte, maschinenlesbare Informationen, die einem Anwender das Auffinden, Suchen und Verwalten von Ressourcen erleichtern. [2], [3]

Man kann beschreibende, strukturierte und administrative Metadaten unterscheiden [2].

- **Beschreibende Metadaten:** dienen zum Auffinden bzw. Identifizieren einer Ressource. Unter diese Gruppe fallen beispielsweise Elemente wie Autor, Titel und Stichwörter.
- **Strukturierende Metadaten:** beschreiben wie verschiedene Teile einer Ressource zusammenhängen können, zum Beispiel wie einzelne Seiten angeordnet werden müssen, um ein Kapitel zu formen.
- **Administrative Metadaten:** bieten Informationen zur Verwaltung einer Ressource, beispielsweise wann und wie sie erstellt wurde und um welchen Dateityp es sich handelt.

Der Begriff Metadaten wird in verschiedenen Anwendungsbereichen oft unterschiedlich verwendet. Bei der Verwendung von Metadaten innerhalb einer Gruppe bzw. Anwendungsdomäne ist es daher sinnvoll, sich zuvor auf eine gemeinsame Bedeutung des Begriffes zu einigen.

2.1.1 Semantische Daten

Die Begriffe Metadaten und semantische Daten werden wegen ihrer Verwandtschaft oft verwechselt bzw. falsch verwendet. Der folgende Vergleich zeigt den Unterschied der beiden Begriffe auf.

Metadaten sind, ganz allgemein gesehen, Daten über Daten. Laut der Definition von Tim Berners-Lee [3], sollten sie strukturiert und vor allem maschinenlesbar sein. Metadaten alleine haben aber noch nicht unbedingt eine Semantik, bedeuten also, für sich gesehen, noch nichts. Ein Beispiel: Wenn jemand die Breite eines Bildes in Zentimetern angibt und nur als Zahl zum Bild speichert, dann ist diese Zahl ein Metadatum. Ohne jedoch die Einheit zu kennen, könnte sie von einem anderen Benutzer fälschlicherweise als Anzahl der Pixel interpretiert werden.

Ganz anders verhält es sich mit semantischen Daten. Diese dienen im Allgemeinen dazu, eine einheitliche Bedeutung zu schaffen. Wenn Metadaten um semantische Informationen angereichert werden, zum Beispiel, wenn sie durch ein Schema eindeutig definiert werden, dann sind sie auch semantische Daten. So drückt

beispielsweise das Metadatum „Autor: Robert Neuner“ zu diesem Dokument den Verfasser dieser Arbeit aus.

2.2 Geografische Metadaten

Unter geografischen Metadaten versteht man Metadaten, die einen geografischen Bezug zwischen der Ressource und ihrem Standort in der realen oder einer virtuellen Welt herstellen können. Man kann dabei zwischen geo-referenzierten Daten und Daten mit geografischem Bezug unterscheiden.

Unter geo-referenzierten Daten versteht man Daten, die direkt mittels Metadaten um eine oder mehrere geografische Komponenten angereichert wurden. Prinzipiell können die verschiedensten Daten geo-referenziert werden. Beispielsweise könnte der Standort aller Tankstellen innerhalb Österreichs standardisiert erfasst werden, um einem Benutzer die nächstgelegenen Tankstellen auf einer Karte anzeigen zu können. Auch die Adresse auf der Visitenkarte eines Restaurants stellt einen geografischen Bezug zwischen der Karte und dem Standort der Gaststätte her. In dieser Arbeit, handelt es sich bei geo-referenzierten Daten um Fotos, die um geografische Informationen, beispielsweise dem Aufnahmeort, ergänzt werden.

Daten mit geografischem Bezug sind hingegen Daten, die bereits durch ihren Inhalt einen Aufschluss über ihren Standort ermöglichen. Sie müssen daher nicht wie geo-referenzierte Daten erst um Koordinaten angereichert werden, sondern ermöglichen bereits durch das reine Lesen bzw. Bearbeiten eine geografische Zuordnung. Es handelt sich also um eine indirekte Geo-Referenzierung. Daten, die von sich aus bereits einen geografischen Bezug aufweisen, gibt es viele, beispielsweise berühmte Sehenswürdigkeiten oder Orte. An dieser Stelle sei aber angemerkt, dass diese Zuordnung nur mit einem Vorwissen, welches besagt dass sich eine bestimmte Sehenswürdigkeit in einer bestimmten Stadt befindet, funktionieren kann. Ein Österreicher wird beispielsweise das Wort Stephansdom sofort mit seinem Standort in Wien assoziieren, ein Ausländer oder gar ein Computer können mit dem Begriff alleine jedoch nichts anfangen.

Direkte und indirekte Geo-Referenzierung unterscheiden sich also im dafür nötigen Vorwissen. Während die direkte Referenzierung neben den Metadaten selbst keine zusätzliche Informationen benötigt um eine eindeutige Referenz herstellen zu

können, ist für die indirekte Referenzierung Vorwissen nötig. Die indirekte Referenzierung hat außerdem den Nachteil, dass es leichter zu Mehrdeutigkeiten kommen kann. Eine Kirche mit dem Namen „Karlskirche“ gibt es beispielsweise einmal in Kassel und einmal in Wien.

2.3 Bildspezifische Metadaten

Bei Bildmetadaten handelt es sich um beschreibende Metadaten, die dem Anwender zusätzliche Informationen über ein Foto liefern können. Tabelle 1 zeigt einige typische Metadaten für Bilder bzw. Fotos. Man sieht, dass zwischen manuell oder automatisch erfassbaren Daten unterschieden werden kann.

Manuell	Automatisch
Bildbeschreibung	Uhrzeit und Datum
Copyright	Bildgröße
Kommentar	Kameramodell
Fotograf	Details zur Kameraeinstellung wie Blende, Autofokus, Blitz
Stichwörter	Farbmodell
Titel	Dateityp

Tabelle 1: Typische Metadaten für Bilder und Fotos

Um die Metainformationen eines Bildes effizient nutzen zu können, ist nicht nur die Qualität der Daten, sondern auch die Speicherung derselben von entscheidender Bedeutung. Man kann zwei Arten von Speichermethoden unterscheiden: die interne Speicherung innerhalb der Nutzdaten und die externe Speicherung in einer separaten Datei bzw. Datenbank. Die beiden Methoden haben Vor- und Nachteile, auf die in den folgenden Abschnitten eingegangen wird. Diese gelten nicht nur für die Arbeit mit geografischen oder bildspezifischen Metadaten, sondern auch für den Umgang mit Metadaten im Allgemeinen.

2.3.1 Interne Speicherung

Die nachstehend aufgeführten Standards erlauben das Speichern von Metadaten innerhalb der Nutzdaten und werden heutzutage von vielen Geräten und Anwendungen unterstützt. Die Speicherung erfolgt dabei innerhalb des Dateih-Headers, einem speziellen Bereich am Anfang einer Datei. Der größte Vorteil dieser Methode besteht in der Tatsache, dass Bild und Metadaten immer in der gleichen Datei vorliegen und der Zusammenhang der beiden Elemente nicht so leicht verloren gehen kann. Dies birgt aber auch Gefahren: Wenn ein Programm beispielsweise neue Metadaten an die falsche Stelle innerhalb der Datei schreibt, kann es dabei wichtige Bildinformationen überschreiben, die im schlimmsten Fall die ganze Datei zerstören können. Die Qualität der Standards und der eingesetzten Tools sind zur sicheren Bearbeitung solcher Bilder daher von entscheidender Bedeutung. Bewährte Standards zur internen Speicherung sind IIM/IPTC, exif und XMP die nachfolgend näher beschrieben werden.

Der IIM-Standard (Information Interchange Model) [4] wurde 1990 von der IPTC-NAA (International Press Telecommunications⁴ -Newspaper Association of America⁵) entwickelt und lebt heute unter der Bezeichnung IPTC weiter. Der Standard erlaubt die textbasierte Speicherung von Metadaten innerhalb eines vordefinierten Sets von Feldern, also beispielsweise: Titel, Beschreibung, Bildrechte und ist einer der Urväter der metadatenbasierten Beschreibung. Einige Felder wurden bereits 1979 definiert, aber erst seit 1994 existiert ein Standard zur direkten Speicherung der Daten innerhalb des Dateih-Headers. Heutzutage wird IPTC vor allem innerhalb großer Bilddatenbanken, wie sie im Verlagsbereich vorkommen, verwendet und nimmt dort einen hohen Stellenwert für das Ablegen und Auffinden von Bildern ein, denn er unterstützt die einfache Verwendung von Schlagwörtern zur Beschreibung eines Bildes. IPTC stellt dem Anwender viele vordefinierte Felder zur Verfügung, ist aber nicht besonders flexibel und auch nicht erweiterbar. Die Felder für geografische Informationen beschränken sich auf Angaben wie Stadt oder Land. [5]

⁴ <http://www.iptc.org>

⁵ <http://www.naa.org/>

Das Exif-Format (Exchangeable Image File Format) [6] wurde von der Japan Electronic and Information Technology Industries Association (JEITA⁶) spezifiziert, um einen Standard für die Speicherung von Bildern, die mit Hilfe von Digitalkameras oder ähnlichen Geräten gemacht wurden, zu schaffen. Definiert wird dabei nicht nur das Dateiformat, in dem die Bilddaten abgelegt werden, sondern auch ein Set von Feldern, die Metainformationen wie Bildgröße, Aufnahmedatum und Blende enthalten können. Bei unkomprimierten Bildern kommt das TIFF-Format, bei komprimierten Bildern das JPEG-Format zum Einsatz.

Metadaten werden innerhalb des Standards entsprechend dem folgenden Codeschema gespeichert:

- **Tag:** Bezeichnung des Tags, wird durch eine eindeutige, 2-byte große Nummer repräsentiert
- **Type:** der Typ des gespeicherten Wertes wie BYTE, SHORT, LONG
- **Count:** die Anzahl der gespeicherten Werte
- **Value Offset:** das Offset vom Anfang des Dateiheders bis zum gespeicherten Wert

Neben den in der Spezifikation definierten Tags, gibt es für die Kamerahersteller auch die Möglichkeit eigene Tags innerhalb der *Maker Note Information* zu definieren. Hier werden Daten gespeichert, deren Information über die typischen exif-Daten hinaus geht und alle Einstellungen der verwendeten Kamera berücksichtigen. Das geschieht aber in fast allen Fällen proprietär und, wie im Falle von Nikon, sogar teilweise verschlüsselt⁷. Viele Programme zum Bearbeiten oder Auslesen der exif-Daten haben folglich Probleme mit den nicht standardisierten Feldern der Hersteller und so kann es im Zuge der Bearbeitung im schlimmsten Fall zu Datenverlust kommen. Für geografische Informationen stellt der Standard ein umfangreiches Set an Feldern zur Verfügung, die es unter anderem erlauben Längengrad, Breitengrad und Höhe des Aufnahmeortes, die Aufnahmerichtung und spezifische Daten der Positionsbestimmung zu speichern. Die exif-Spezifikation definiert noch weitere Felder, wie Geschwindigkeit bei der Aufnahme oder

⁶ <http://www.jeita.or.jp/english/>

⁷ Nikon Forum: http://photo.net/bboard/q-and-a-fetch-msg?msg_id=00Dkyp

spezifische Kontrolldaten der Positionsbestimmung, die jedoch für diese Arbeit nicht benötigt werden.

Der XMP-Standard (Extensible Metadata Platform) [7] definiert hingegen ein Containerformat, innerhalb dessen Schemata generiert werden können um Metadaten standardisiert abzulegen. Dies kann sowohl innerhalb der Nutzdaten, als auch über eine externe Speicherung geschehen. XMP wurde 2005 in Kooperation der Firma Adobe Systems⁸ und dem IPT Council der Öffentlichkeit präsentiert. Die Basiselemente, die zur MetadatenSpeicherung zur Verfügung stehen, orientieren sich am Dublin Core Standard⁹ und ermöglichen so die Verwendung von einfachen Feldern wie Title, Description und Keywords. Die Speicherung selbst erfolgt in XML-Syntax, dem RDF Schema¹⁰ des World Wide Web Consortiums entsprechend. XMP unterstützt nicht nur Metadaten innerhalb von JPEG oder TIFF Dateien, sondern auch viele andere Bildformate sowie Formate von Adobe selbst wie PDF, Adobe Illustrator oder Adobe Photoshop. Weiters bietet der Standard vordefinierte Schemata an, mit denen der gesamte exif- oder IIM/IPTC Standard in XMP abgebildet werden kann. XMP bietet mit dem exif-Schema [7] das gleiche umfangreiche Set an, um geografische Metainformationen eines Fotos speichern zu können. Im Gegensatz zum exif-Format bietet XMP durch seine Erweiterbarkeit jedoch die Möglichkeit, zusätzliche Felder zu definieren und so auf einer standardisierten Basis die Semantik eigener Metadaten weiter zu steigern. [8]

2.3.2 Externe Speicherung

Bei der externen Speicherung werden die Metainformationen getrennt von den Nutzdaten gespeichert. Dies kann entweder in separaten Dateien oder in Datenbanken geschehen. Die externe Speicherung spielt ihre Vorzüge besonders bei großen Datenbeständen aus. Anstatt in jedem Dateiheader nach einem bestimmten Metadatum suchen zu müssen, kann beispielsweise innerhalb einer Datenbank, in der die Metadaten als Referenz zur Datei gespeichert wurden, gesucht werden.

⁸ <http://www.adobe.com/>

⁹ <http://dublincore.org>. Dublin Core stellt ein einfaches Metadaten-Schema mit nur 15 Kernelementen zur Beschreibung einer Ressource zur Verfügung.

¹⁰ <http://www.w3.org/TR/rdf-schema/>

Ein Nachteil der externen Speicherung von Metadaten liegt jedoch auf der Hand. Jedes Programm, welches die Nutzdaten in einer Weise verändert, die eine Anpassung der Metadaten notwendig macht, muss diese Änderung auch innerhalb der Datenbank bzw. der externen Datei durchführen. Wenn dies nicht geschieht, sind Daten und Metadaten nicht mehr synchron und es besteht die Gefahr, dass mit fehlerhafter Information weitergearbeitet wird. Die Synchronisation der Daten kann nicht nur direkt durch das Programm erfolgen das die Änderung hervorgerufen hat, sondern auch über einen externen Prozess der beispielsweise die Nutzdaten in regelmäßigen Abständen auf Modifizierungen überprüft und gegebenenfalls die extern gespeicherten Metadaten aktualisiert.

Ein Metadatenstandard der auch die externe Speicherung unterstützt ist MPEG-7 (Multimedia Content Description Interface-7) [9]. Das Hauptaugenmerk von MPEG-7 liegt im Unterschied zu den anderen MPEG Standards 1, 2 und 4, die sich hauptsächlich mit der Komprimierung der Nutzdaten beschäftigen, in der standardisierten Beschreibung verschiedener multimedialer Inhalte. Ziel des Standards ist die manuelle bzw. automatische Generierung verschiedenster Metadaten, die das effiziente Organisieren und Suchen innerhalb audiovisueller Daten ermöglichen. Um dies zu erreichen, werden aus den Multimediaobjekten bestimmte Merkmale extrahiert und abgespeichert. Dieser Vorgang geschieht möglichst ohne das Zutun eines Anwenders, da die Merkmale meist auf Informationen beruhen, die die Daten selbst liefern können. Ist ein Merkmal einmal extrahiert, muss es zur weiteren Verarbeitung in einen semantischen Zusammenhang gebracht werden. Dies geschieht über so genannte *Descriptors*, die gleichzeitig auch die Syntax für das Merkmal festlegen. MPEG-7 stellt eine Vielzahl an Deskriptoren bereit, die wiederum durch *Description Schemes* in Anwendungsbereiche gruppiert werden. Zusätzlich stellt der Standard eine eigene Sprache, die Description Definition Language (DDL) zur Verfügung, mit Hilfe derer vorhandene Schemata erweitert oder zusätzliche Schemata beschrieben werden können. Obwohl MPEG-7 selbst keine geografischen Metadaten vorsieht, erlaubt es dessen Erweiterbarkeit eigene Deskriptoren für geografische Information als Metadaten innerhalb von MPEG-7 zu definieren.

Die Verwendung von Datenbanken zur Speicherung von Metadaten ist hingegen eine Methode wenn der Fokus auf Geschwindigkeit und Langzeitarchivierung liegt. Prinzipiell kann zur Speicherung jede relationale bzw. objektorientierte Datenbank verwendet werden. Der Vorteil einer Datenbank liegt in der Tatsache, dass der Zugriff auf bestimmte Datenfelder über einen Index geschieht. Im Vergleich zur sequentiellen Suche, bei der nacheinander alle Dateien geöffnet, gelesen und wieder geschlossen werden müssen, muss bei dieser Methode nur der Index durchsucht werden um anschließend mit einer einzigen Leseoperation auf die gesuchten Daten zugreifen zu können.

Da es sich bei geografischen Informationen wie dem Längen- und Breitengrad eines Standpunktes um zwei verschiedene Werte handelt, müssen für die Suche nach dem Standort zwei Indizes erstellt werden. Dies wird in den meisten Fällen jedoch zu keinem Geschwindigkeitseinbruch führe. Für den Fall, dass es doch nötig sein sollte mit nur einem Index zu arbeiten, stellen die Autoren des Papers „*Location Tags on Digital Images*“ [10] eine effiziente Methode zur Speicherung von Längen- und Breitengrad innerhalb einer relationalen Datenbank dar. Sie verwenden dazu eine Funktion, die die beiden Werte mit wählbarer Genauigkeit zu einer einzigen Zahl kombiniert. Da diese Zahl anschließend als Index verwendet werden kann, führt diese Methode zu einer besonders performanten Suche.

2.3.3 Evaluierung

Tabelle 2 vergleicht wichtige Eigenschaften der Standards.

Eigenschaften	IIM/IPTC	Exif	XMP	MPEG-7
Textbasiert	Ja	Ja	Ja, XML-Syntax	Ja, XML-Syntax
Textlängenbeschränkung	Ja	Ja	Nein	Nein
Thumbnails	Nein	Ja, binär	Ja, Base64 Form	Ja
Erweiterbar	Nein	Ja, proprietär	Ja	Ja
Felder für geografischen Bezug	Nur Auswahl von Stadt, Land möglich	Ja, beispielsweise LAT und LONG	Wie exif, aber auch durch Erweiterbarkeit gegeben	Durch eigene Deskriptoren möglich

Tabelle 2: Vergleich von IIM/IPTC, Exif, XMP und MPEG-7

Es zeigt sich, dass die in Kapitel 1 beschriebenen Anforderungen für die Speicherung von geografischen Informationen innerhalb der genannten Speicherverfahren nur von IIM/IPTC nicht erfüllt werden. Dieser Standard erlaubt keine Speicherung von Koordinaten für den Aufnahmeort und bietet auch keine Unterstützung für die Aufnahmerichtung des Fotos. Exif und XMP erfüllen hingegen die genannten Anforderungen und ermöglichen die Speicherung von Längen- und Breitengrad und der Aufnahmerichtung. Auch im sehr mächtigen MPEG-7 können dank der flexiblen Erweiterbarkeit und der Möglichkeit, eigene Schemen und Deskriptoren entwickeln zu können, geografische Information abgespeichert werden. Aufgrund des großen Umfangs des Standards, der dafür notwendigen Einarbeitungszeit und vor allem wegen der geringen Verbreitung, fällt die Entscheidung für den in dieser Arbeit zu verwendenden Standard jedoch nicht auf MPEG-7.

Die Entscheidung zwischen exif und XMP fällt hingegen schwer. Beide Formate unterstützen die genannten Anforderungen und bieten notfalls Methoden zur Erweiterung des Formates an. Letztlich ist es jedoch sinnvoll, dem exif-Standard den Vorzug zu geben. Im Vergleich zu XMP ist er bereits weit verbreitet und durch den geringeren Umfang leichter verwaltbar. Noch viel wichtiger ist aber, dass er von nahezu allen digitalen Kameras unterstützt wird und Fotos dieser Geräte bereits alle einen exif-Header besitzen.

Die zusätzliche datenbankbasierte Speicherung von Metadaten ist sinnvoll, wenn die Suche nach Metadaten im Vordergrund steht. Dazu können bereits bestehende Metadaten innerhalb einer Datei ausgelesen und anschließend als Kopie in einer Datenbank gespeichert werden. Diese Methode kombiniert die Vorteile aus beiden Welten und verbindet etablierte interne Metadatenstandards mit einer indexbasierten Suche. In diesem Fall darf jedoch auf die notwendige Synchronisation der Daten nicht vergessen werden.

Kapitel 3

Technische Voraussetzungen zur Gewinnung geografischer Metadaten

Dieses Kapitel beschäftigt sich mit den technischen Voraussetzungen, die gegeben sein müssen, um die geografischen Informationen, die ein Foto aufweisen kann, gewinnen zu können. Es gibt verschiedene Möglichkeiten, die notwendigen Daten zur Aufnahmeposition bzw. Aufnahmerichtung eines Fotos zu bestimmen. Das Paper „Geographic Location Tags on Digital Images“ [10] bietet eine Übersicht über verschiedene Techniken der Positionsbestimmung. Aufgrund der Tatsache, dass sich einige der dort genannten Techniken, die beispielsweise das Mobilfunk- oder WLAN-Netz zur Berechnung des Standortes nutzen, noch in Entwicklung befinden, konzentriert sich diese Arbeit auf die satellitengestützte Positionsbestimmung.

Das NAVSTAR-GPS der USA ist derzeit das einzige System, das weltweit verfügbar ist und zur zivilen Nutzung geöffnet wurde. Konkurrenzsysteme wie das russische GLONASS, das chinesische COMPASS und das europäische GALILEO erreichen erst in den nächsten Jahren ihre letzte Ausbaustufe [11], [12]. Alle vier genannten Positionierungssysteme funktionieren auf ähnliche Weise, weshalb sich diese Arbeit auf die Beschreibung des NAVSTAR-GPS Systems beschränkt.

Neben der Bestimmung des Aufnahmeortes kommt im Zuge dieser Arbeit der Aufnahmerichtung eines Fotos eine entscheidende Bedeutung zu. Um diese zu messen, wird ein Kompass verwendet, der die Ausrichtung der Kamera feststellen

kann. Im zweiten Teil dieses Kapitels werden die Funktionsweise und Eigenschaften von elektromagnetischen Kompassen diskutiert.

3.1 Das NAVSTAR GPS-System

Das Global Positioning System wurde Anfang der 70er Jahre von den USA für militärische Zwecke entwickelt. Hauptziel war es, eine möglichst genaue Positionsbestimmung auf der Erde zu ermöglichen. Der Genauigkeitsbereich von GPS liegt beim gleichzeitigen Empfang von mindestens drei Satelliten unter 10 Metern. Dieser Bereich lässt sich aber mit mehr oder weniger aufwendigen Methoden in den Millimeterbereich senken. Ursprünglich für das Militär entwickelt, erkannten die Verantwortlichen wenig später auch das Potential, welches GPS für die zivile Nutzung haben könnte und es wurde unter der Amtszeit von US-Präsident Reagan schlussendlich auch für diesen Bereich freigegeben. Das geschah jedoch mit dem Vorbehalt, die Signale bzw. die Genauigkeit jederzeit mit einem künstlichen Fehler verschlechtern zu können. Dies sollte die militärische Nutzung eines potentiellen Feindes einschränken. Erst im Mai 2000 wurde diese künstliche Verschlechterung, die so genannte Selective Availability (SA), endgültig abgeschaltet und seit dem erlebt die zivile Nutzung des Systems einen großen Aufschwung. [13], [14]

3.1.1 Aufbau und Funktionsweise

GPS ist ein weltumspannendes Netzwerk von ursprünglich 24 Satelliten¹¹, das einem Anwender auf der Erde eine relativ genaue Positionsbestimmung ermöglicht. Die Satelliten kreisen in sechs verschiedenen Orbitalbahnen, in ca. 20.000 km Höhe. Die Bahnen werden auch als Slots bezeichnet und beherbergen jeweils mindestens 4 Satelliten. Diese Slots sind so gewählt, dass auf jedem Punkt der Erde, sogar auf den Polen, immer mindestens 4 Satelliten gleichzeitig sichtbar sind. [13]

¹¹ Heute sind es ca. 30. Die aktuelle Anzahl der Satelliten findet man beispielsweise unter:
<http://www.kowoma.de/gps/gpsmonitor/gpsstatus.htm>

Das GPS-System setzt sich aus 3 verschiedenen Segmenten zusammen: [13]

- **Weltraumsegment:** Das Weltraumsegment besteht aus den zuvor erwähnten 24 Satelliten, die ca. zweimal am Tag die Erde umkreisen und dabei immerfort ihre Position übermitteln. Die erwartete Lebensdauer eines Satelliten beträgt zwischen 5 und 7 Jahren. Manche funktionieren aber auch noch viele Jahre darüber hinaus.
- **Kontrollsegment:** Das Kontrollsegment besteht aus mehreren, über die Welt verteilten Bodenstationen, die die Satelliten ständig beobachten. Dabei werden ihre Bahndaten und die Uhrengenauigkeit analysiert und bei etwaigen Abweichungen Korrekturdaten an die Satelliten gesendet.
- **Benutzersegment:** Unter das Benutzersegment fallen prinzipiell alle Geräte und Anwendungen, die das GPS-Signal zur Positionsbestimmung oder Vermessung nutzen.

Jeder Satellit überträgt folgende Daten: [15]

- **Identifikationsnummer:** Zahl zwischen 1 und 32, die den Satelliten eindeutig identifiziert
- **Ephemeridendaten:** Beinhalten den Satellitenstatus, die genaue Uhrzeit, Information zu seiner eigenen Bahn und eventuelle Korrekturdaten
- **Almanachdaten:** Enthalten die ungefähren Positions- und Bahndaten von allen anderen funktionierenden Satelliten

Damit ein GPS-Empfänger seine eigene Position auf der Erdoberfläche bestimmen kann, benötigt er die Signale von mindestens 3 Satelliten. Wird auch die Höhe gewünscht sind es mindestens 4. Um die Position zu berechnen, vergleicht der Empfänger die Zeit, zu der das Signal vom Satelliten gesendet wurde, mit der Zeit als es empfangen wurde. Die Differenz dieser Zeiten wird als Laufzeit bezeichnet. Mit Hilfe der Laufzeit und den Informationen über die aktuelle Position des Satelliten kann nun die Entfernung zwischen Satellit und Empfänger errechnet werden. Sobald die Entfernung bekannt ist, kann mittels trigonometrischer Gleichungen die Position bestimmt werden. Abbildung 2 veranschaulicht diesen Prozess in einer Grafik. Um jeden Satelliten wird eine gedankliche Kugel mit dem Radius der berechneten Entfernung gelegt und ein gemeinsamer Schnittpunkt gesucht. Je mehr Satelliten beteiligt sind, desto genauer wird die Berechnung. [16]

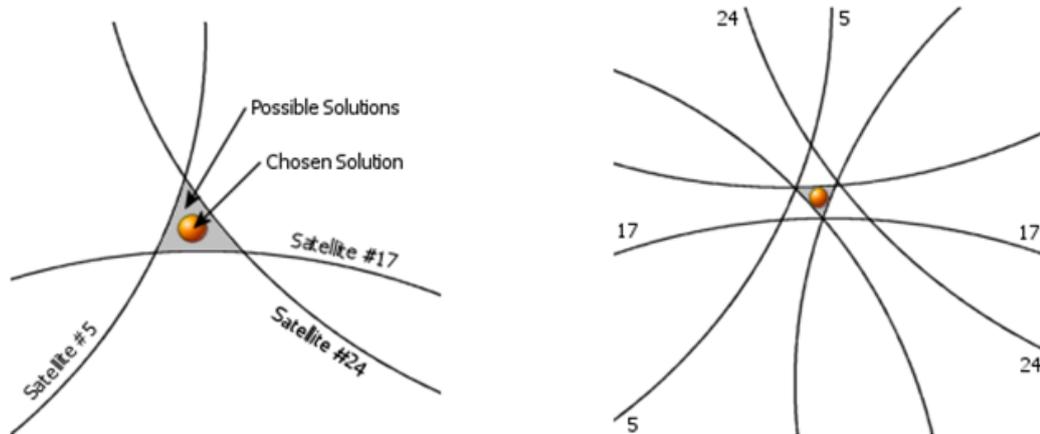


Abbildung 2: Auswirkung der Anzahl und der Verteilung der sichtbaren Satelliten auf die Güte der ermittelten Position [16]

Idealerweise würden sich alle Kugeln in einem Punkt schneiden. In der Praxis gibt es aber immer wieder Störungen und Reflexionen, die das Signal beeinflussen können. In den meisten Fällen wird sich also eine ähnliche Form wie in Abbildung 2 dargestellt ergeben innerhalb der sich der Empfänger seine Positionen wählen muss. [16]

3.1.2 Signaleigenschaften und Störeinflüsse

Die Sendeleistung der Satelliten ist im Allgemeinen nicht besonders hoch und ermöglicht generell nur dann einen sehr guten Empfang zwischen Satellit und Empfänger, wenn ein direkter Sichtkontakt vorherrscht. Ein Empfang innerhalb von Gebäuden und Tunneln ist daher ohne zusätzliche Technik grundsätzlich nicht möglich. Für eine gute Positionsbestimmung ist aber auch eine gute Verteilung der empfangbaren Satelliten am Himmel, die so genannten Satellitengeometrie, wichtig. Die Ausrichtung und Position der Satelliten ist nämlich von entscheidender Bedeutung. Senden beispielsweise alle Satelliten aus Sicht des Empfängers aus einer ähnlichen Himmelsrichtung, dann wird der Spielraum für die exakte Bestimmung der Position relativ groß. Die gedachten Kugeln schneiden sich dann sehr ungünstig. Ein sehr gutes Ergebnis wird nur dann erzielt, wenn möglichst viele Signale aus möglichst verschiedenen Himmelsrichtungen empfangen werden können. [16]

Grundsätzlich können auftretende Fehler in 3 Kategorien eingeteilt werden [13]:

- **Satellitenfehler:** Die Genauigkeit der Bahndaten, auch Ephemeriden genannt, trägt einen Großteil zur Güte der Positionsbestimmung bei. Bewegt sich der Satellit auch nur ein klein wenig abseits seiner berechneten Bahn, kann dies auf Grund der trigonometrischen Berechnung zu großen Abweichungen führen.

Die Genauigkeit der Uhrzeit ist ein weiteres wichtiges Kriterium. Die in den Satelliten verbauten Atomuhren, müssen ständig überwacht werden, um ein zeitlich synchrones System zu schaffen. Aufgrund der Höhe und der Geschwindigkeit der die Uhren ausgesetzt sind, müssen sogar relativistische Effekte¹² bei der Uhrentaktung berücksichtigt werden.

- **Signalausbreitungsfehler:** Wie alle funkbasierten Anwendungen unterliegt auch das GPS-Signal in der Praxis umweltbedingten Störeinflüssen. Zwar gilt das Signal als relativ wetterunempfindlich, reagiert aber empfindlich auf magnetische Effekte, die in der Ionosphäre entstehen können und auf Reflexionen, die durch Gebäude hervorgerufen werden.

Gerade die Ionosphäre ist einer der größten Störfaktoren für das GPS-Signal. Das Problem besteht darin, dass sich ein elektrisches Signal im Vakuum mit Lichtgeschwindigkeit ausbreitet und beim Eindringen in die Atmosphäre gebremst wird. Dieser „Brems-Effekt“ ist in elektrisch geladenen Schichten der Atmosphäre zwar nicht höher, das Signal wird aber leicht abgelenkt und muss dadurch einen geringfügigen längeren Weg zurücklegen. Genau das wirkt sich aber negativ auf die Qualität der Messung aus, da zu einer genauen Positionsbestimmung die Laufzeit des Signals entscheidend ist.

Abhilfe für dieses Problem schaffen unter anderem spezielle Differential-GPS-Empfänger, die eine so genannte Ionosphärenkorrektur durchführen können. Darunter versteht man Korrekturdaten, die nach Messungen der Atmosphärenverzögerung von einer Bodenstation an den Empfänger

¹² Laut Einsteins Relativitätstheorie laufen bewegte Uhren langsamer als ruhende Uhren und Uhren, die sich in einem schwachen Gravitationsfeld befinden schneller als welche, die starker Gravitation ausgesetzt sind.

übertragen und für eine genauere Berechnung der Position genutzt werden können.

Gebäudereflexionen sind hingegen nur in Städten mit hohen Häusern störend. Das Problem für den Empfänger besteht darin, dass er durch die Reflexionen ein und dasselbe Signal eines Satelliten mehrmals kurz hintereinander mit unterschiedlichen Laufzeiten empfängt. Abbildung 3 verdeutlicht diesen Sachverhalt.

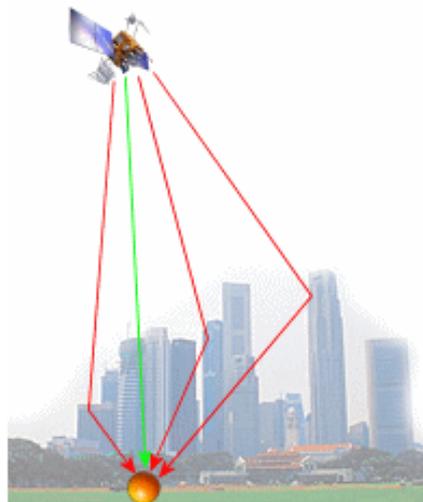


Abbildung 3: Reflexionsproblematik eines GPS-Signals [16]

Dieses Problem, meist als *Multipath-* oder *Mehrwegephänomen* bezeichnet, kann von modernen GPS-Empfängern aber leicht behoben werden. Die Lösung besteht in der Verwendung des ersten eintreffenden Signals eines Satelliten. Dieses ist aufgrund der kürzeren Laufzeit höchst wahrscheinlich das Direktteste und alle später eintreffenden Reflexionen des Ursprungssignals können verworfen werden.

- **Empfängerfehler:** Unter Empfängerfehler fallen das so genannte Messrauschen und die hardwaretechnische Verzögerung bei der Signalverarbeitung. Beide Störeinflüsse wirken sich aber sehr gering auf die Güte der Positionsbestimmung aus und können daher für die meisten Anwendungsgebiete vernachlässigt werden.

Wie stark sich welcher Fehler auf die Abweichung zur korrekten Position auswirkt kann aus Tabelle 3 abgelesen werden. [17]

Störungsart	Geschätzte Abweichung
Störungen durch die Ionosphäre	± 5 Meter
Schwankungen der Satellitenumlaufbahnen	± 2.5 Meter
Uhrenfehler der Satelliten	± 2 Meter
Mehrwegeeffekt	± 1 Meter
Störungen durch die Troposphäre	± 0.5 Meter
Rechnungs- und Rundungsfehler	± 1 Meter

Tabelle 3: Auswirkung der behandelten Fehlern auf die Güte der GPS-Position [17]

3.1.3 Das NMEA Protokoll

Um einen möglichst reibungslosen und standardisierten Austausch von GPS-bezogenen Daten zu gewährleisten, wurde in den 80er Jahren der NMEA (National Marine Electronics Association) Standard [18] entwickelt. Von den Versionen NMEA-0180, 0182, 0183, 2000 findet man heutzutage hauptsächlich den NMEA-0183 in Verwendung. Dieser Standard wird auch von den meisten am Markt befindlichen GPS-Empfängern unterstützt. [14]

Der Standard sieht zur Kommunikation zwischen GPS-Empfänger und Software ein serielles Protokoll nach RS-232 Spezifikation vor. Die Daten selbst werden als ASCII-Zeichen in Sätzen bis maximal 80 Zeichen übermittelt. Jeder neue Satz beginnt mit einem Dollarzeichen \$, gefolgt von einer Sender-ID, beispielsweise GP¹³, und einer Satz-ID, beispielsweise RMC¹⁴. Danach folgen die jeweiligen Werte in kommaseparierter Darstellung. Die beiden IDs legen genau fest um welche Art von Satz es sich handelt und sind in verschiedene Gruppen eingeteilt. Eine detaillierte Übersicht aller Gruppen, inklusive deren Beschreibung, ist im Standard [19] einsehbar.

¹³ GP, Abkürzung für GPS. Weist den Sender als GPS-Empfänger aus.

¹⁴ RMC, Abkürzung für „Recommended Minimum Sentence“. Dieser stellt die wichtigsten GPS-Daten wie Position und Uhrzeit bereit.

Die in dieser Arbeit verwendeten Sätze werden in den entsprechenden Praxiskapiteln genauer erklärt und aufgeschlüsselt. Abbildung 4 und Abbildung 5 sollen aber jetzt schon als Einblick in den Standard dienen.

```
$GPRMC,121255.000,A,4812.9061,N,01620.9111,E,1.48,25.39,070507,,A*57
$GPGGA,121256.000,4812.9067,N,01620.9117,E,1.03,3.8,253.3,M,43.4,M,0000*5F
$GPGSA,A,2,25,13,20,,,,,,,,,3.9,3.8,1.0*34
$GPGSV,3,1,12,23,79,047,21,13,64,247,33,25,52,151,29,04,46,293,20*74
$GPGSV,3,2,12,20,42,107,23,27,14,189,,02,14,318,,31,08,034,09*71
$GPGSV,3,3,12,16,08,079,15,17,06,236,14,01,04,058,,11,02,165,*75
```

Abbildung 4: Typische NMEA-0183 Satzfolge¹⁵

Abbildung 4 zeigt einen typischen NMEA Block. Jede neue Zeile enthält einen kompletten Satz, der nun weiter aufgeschlüsselt werden kann. Anhand der ersten Zeile dieses Blockes, zeigt Abbildung 5, wie beispielsweise ein \$GPRMC Satz interpretiert werden muss, um Uhrzeit, Koordinaten, Geschwindigkeit und Kurs auszulesen.

```
$GPRMC,121255,A,4812.9061,N,01620.9111,E,1.48,25.39,070507,,A*57
```

Uhrzeit der Positionsbestimmung: 12:12:55 UTC-Zeit

Empfängerwarnung, A = Daten OK

48° 12.9061' nördliche Breite

16° 20.9111' östliche Länge

Geschwindigkeit über Grund (Knoten)

wahrer Kurs (ohne Bewegung 0)

Datum: 07.05.07

Modus, A = Autonome Messung

Prüfsumme

Abbildung 5: Interpretation eines \$GPRMC Satzes

Nach der Identifikationsnummer folgt die Uhrzeit im UTC Format. Anschließend enthält der Satz ein Feld, das angibt, ob es Probleme bei der Messung gab oder ob die Daten fehlerfrei gewonnen wurden. Danach folgen Felder für Längengrad, Breitengrad, Hemisphäre, Geschwindigkeit, Kurs und Datum. Abschließend folgen noch ein Feld, das angibt ob die Messung autonom oder differentiell durchgeführt

¹⁵ Die Sätze wurden unter Microsoft Windows mit dem integrierten Hyperterminal von einem USB GPS-Empfänger in der Privatwohnung aufgenommen.

wurde, und eine Prüfsumme, mit der die Fehlerfreiheit des Satzes selbst überprüft werden kann.

3.1.4 Evaluierung

GPS ist ein leistungsfähiges System zur Positionsbestimmung auf der Erde. Eine wichtige Anforderung dieser Arbeit an GPS ist dessen Genauigkeit. Wenn man die möglichen Abweichungen der verschiedenen Fehlerarten aufsummiert, ergibt sich eine maximale Gesamtabweichung von +/- 12 Metern. Da manche Fehler wie der Mehrwegefehler oder etwaige Rechenfehler selten auftreten und die Abweichung nicht nur in eine Richtung geschieht, gehen die Autoren Bauer [13] und Görrisch [14] von einer Positionsgenauigkeit aus die in den aller meisten Fällen unter 10 Metern liegt. Dies ist für diese Arbeit ausreichend, da Aufgrund der Anforderungen dieser Arbeit davon ausgegangen werden kann, dass Fotos von Gebäuden oder Sehenswürdigkeiten von mehr als 10 Metern Entfernung gemacht werden.

Eine weitere Anforderung an das System ist eine schnelle Betriebsbereitschaft. Dies ist gerade beim Fotografieren wichtig, denn niemand möchte dabei minutenlang auf die Positionsbestimmung warten müssen. Der in dieser Arbeit genutzte GPS-Empfänger kann dieses Kriterium nur zum Teil erfüllen. Wenn der Empfänger mehrmals am Tag benutzt wird, dann liefert er bereits nach ca. 8 Sekunden eine gültige Position, siehe Anhang III. Ein Kaltstart des Gerätes, benötigt immerhin schon 48 Sekunden. Falls der Empfänger mehrere Wochen nicht benutzt wurde, kann es sein, dass die Almanachdaten nicht mehr gültig sind und neu heruntergeladen werden müssen. Dieser Vorgang dauerte beim Testgerät ca. 30 Minuten.

Die Unterstützung des NMEA Protokolls für eine standardisierte Ausgabe der Messdaten erweist sich in der Praxis als ausreichend. Einen Computer vorausgesetzt, können damit sekundlich aktualisierte Positionsdaten ausgelesen werden.

Ein wichtiges Kriterium für den mobilen Einsatz ist auch der Stromverbrauch. Der getestete GPS-Empfänger lief im Dauereinsatz ca. 10 Stunden. Bei integrierten Lösungen, in denen sich Kamera und GPS-Empfänger einen Akku teilen müssen, ist also mit einer geringeren Laufzeit zu rechnen.

3.2 Elektronischer Kompass

Um nicht nur den Standort des Fotografen, sondern auch die Richtung, in die die Aufnahme gemacht wurde, feststellen zu können, reicht ein GPS-Gerät alleine nicht aus. Zwar können die meisten Empfänger anhand von hintereinander folgenden Positionsmessungen eine Bewegungs- bzw. Marschrichtung errechnen und daraus die geografische Nordrichtung ableiten, allerdings ist dies aus Prinzip nur bei einer Bewegung des Empfängers möglich [20].

Die Anforderungen die diese Arbeit an die Richtungsbestimmung stellt, erfordern jedoch auch eine Messung bei Stillstand des Fotografen. Für die Bestimmung der Aufnahmerichtung kommt daher ein elektronischer Kompass zum Einsatz, der die Ausrichtung der Kamera gemäß einer 360° Einteilung in Richtung Norden messen können soll. Solche Kompass gibt es heutzutage in verschiedensten Formen und Größen, vom kleinen tragbaren Gerät bis hin zu stabilen Festeinbauten in Fahrzeugen. Die Genauigkeit beträgt meist $\pm 1^\circ$.

3.2.1 Aufbau und Funktionsweise

Ein elektronischer Kompass ist seinem analogen Pendant in der Funktionsweise sehr ähnlich. Auch er benutzt die magnetischen Feldlinien der Erde, um daraus seine Ausrichtung errechnen zu können. Was bei einem Analogkompass die Nadel, sind bei einem digitalen Gerät entweder so genannte Feldplattensensoren oder millimeterdünne, magnetsensitive Folien wie MU-Metall bzw. Permalloy. Diese Materialien sind in der Lage ihren Widerstand anhand von magnetischen Einflüssen zu ändern und eignen sich daher gut für die Messung des Erdmagnetfeldes. [21], [22]

3.2.2 Einschränkungen bei der magnetischen Richtungsmessung

Digitale wie analoge Kompass haben ähnliche Probleme wenn es um die genaue Richtungsbestimmung geht. Im Folgenden werden daher die Begriffe Deklination, Inklination und Deviation kurz erläutert werden [23]:

Durch die Neigung der Erdachse befindet sich der geografische Nordpunkt ca. 2000 Kilometer entfernt vom magnetischen Nordpunkt. Dieser ist zusätzlich ständigen Schwankungen unterworfen und wandert im Laufe der Zeit. Die Deklination beschreibt die Winkeldifferenz zwischen den beiden Nordpunkten. In vielen Teilen

Mitteleuropas beträgt die Missweisung unter 1° Abweichung und kann daher meist vernachlässigt werden. In Gegenden wie auf den Baffin Islands im Nordosten Kanadas kann die Deklination aber auch bis zu 50° West betragen und muss natürlich dementsprechend in die Berechnungen einbezogen werden. [20]

Die Inklination bezeichnet den Winkel zwischen den magnetischen Feldlinien und der Erdoberfläche. Da sich die Nadel eines Kompasses immer anhand des kürzesten Weges in Richtung magnetischem Norden ausrichtet, würde sie bei uns in Europa „in die Erde hinein“ zeigen. Um das zu verhindern werden die Nadeln auf der anderen Seite künstlich beschwert, was dazu führt, dass sie für verschiedene Breitengrade angepasst werden müssen.

Kompassreagieren aus Prinzip extrem störanfällig auf magnetische Objekte oder Induktionen, die sich in ihrer Nähe befinden bzw. stattfinden. Dieser Umstand wird als Deviation bezeichnet und sollte bei jeder Messung bedacht werden. Der Kompass sollte sich also niemals neben Elektrogeräten oder gar Magneten befinden, da diese das Ergebnis maßgeblich verändern können.

3.2.3 Evaluierung

Große Vorteile der digitalen Messung sind zum einen die Geschwindigkeit und Stoßfestigkeit der verwendeten Sensoren da keine beweglichen Teile verwendet werden und andererseits die Möglichkeit zur softwaremäßigen Weiterverarbeitung der Daten. Ein erwähnenswerter Nachteil besteht jedoch in der Tatsache, dass ein elektronischer Kompass vor seinem Einsatz kalibriert werden sollte. Eine Kalibrierung ist immer dann empfehlenswert, wenn sich der Standort der Messung stark geändert hat und dient dazu, den empfindlichen Sensor auf den neuen Ort einzustellen. Der in dieser Arbeit verwendete Kompass muss zu diesem Zweck zweimal um die eigene Achse gedreht werden – ein Prozess der auf Dauer doch umständlich ist.

Das Gerät verfügt außerdem über keinerlei Neigungskompensation und muss daher bei einer Messung waagrecht gehalten werden. Diese Einschränkung stellt innerhalb dieser Arbeit kein Problem dar, da der Kompass nicht physisch mit der Kamera verbunden ist, sollte jedoch bei einer etwaigen Kombination von Kamera und Kompass bedacht werden.

Die Genauigkeit von $\pm 1^\circ$ ist ausreichend. Wenn beim Fotografieren von einer maximalen Distanz von 200 Metern ausgegangen wird, dann beträgt die „Verschiebung“ des fotografierten Objekts in dieser Entfernung ca. $\pm 3,5$ Meter.

Deklination und Deviation stellen zwar Probleme bei der Richtungsbestimmung dar, sind aber relativ leicht zu korrigieren bzw. vernachlässigbar. Da man davon ausgehen kann, dass sich das Erdmagnetfeld nicht innerhalb von wenigen Jahren wesentlich verschiebt, kann man die Deklination für einen gewissen Zeitraum und eine bestimmte Region als konstant betrachten. Viele Handheld-GPS-Geräte bieten zu diesem Zweck eine automatische Missweisungskorrektur an. Bei der Deviation hingegen reicht es schon, darauf zu achten, dass sich bei der Messung keine potentiellen Störobjekte in der Nähe befinden. Inwieweit dies jedoch bei der Integration eines elektronischen Kompasses in eine Digitalkamera berücksichtigt werden kann, muss weiter untersucht werden.

Die Inklination kann hingegen nur durch aufwendige, mechanische Lagerung eliminiert werden. Für beide Kompass-Typen ist eine in alle Richtungen bewegliche Lagerung zu empfehlen, wie sie bei Kreiselkompassen aus der Schifffahrt bekannt ist. Für digitale Kompassse empfiehlt sich auch die Verwendung eines 3-Achsen Kompasses, inklusive Roll- und Neigungskompensierung, wie er in [22] beschrieben wird.

Eine wesentliche Einschränkung in der Praxis ist auch der ständig notwendige Kalibrierungsprozess. Sollte für dieses Problem in Zukunft keine Lösung gefunden werden, ist es bei einer bestehenden Kombination von Kamera und Kompass denkbar, den Sensor beim Einschalten des Gerätes innerhalb des Gehäuses rotieren zu lassen. Nach dieser kurzen internen Kalibrierung wäre die Kamera dann voll einsatzbereit.

Kapitel 4

Metadaten zur Verbesserung des Suchprozesses

Dieses Kapitel beschäftigt sich mit der Gewinnung der notwendigen geografischen Informationen und deren Verknüpfung mit einem bestimmten Bild. Die Autoren des Papers „*Geographic Location Tags on Digital Images*“ [10] beschreiben in ihrer Arbeit mehrere Verfahren für diesen Referenzierungsprozess. Einige davon sind auch für diese Arbeit relevant und werden in Kapitel 4.1 behandelt.

Der zweite Teil dieses Kapitels beschäftigt sich mit einer neuen Methode, die nur anhand geografischer und aufnahmespezifischer Informationen feststellen kann, welcher Bildausschnitt der realen Welt auf dem Foto abgebildet wurde. Ergebnisse dieser Untersuchung sollen später in die Verbesserung des Suchprozesses einfließen.

4.1 Geo-Referenzierung

Unter dem Begriff der Geo-Referenzierung versteht man das direkte Verknüpfen von beliebigen Nutzdaten mit standortbezogenen Daten, beispielsweise das Verknüpfen von Bildern mit ihrem Aufnahmeort oder auch die Angabe einer Adresse auf einer Visitenkarte. Der Prozess funktioniert natürlich auch umgekehrt. Wenn man beispielsweise nur den Namen eines Restaurants kennt und über diesen nach der Adresse sucht, dann ist das eine indirekte Geo-Referenzierung, auch Reverse Geocoding genannt. Vor- und Nachteile der direkten bzw. indirekten Geo-Referenzierung wurden bereits in Kapitel 2.2 beschrieben.

Bei der Georeferenzierung von Bildern kann zwischen Verfahren unterschieden werden, welche die Nutzdaten bei ihrer Erstellung referenzieren, oder Verfahren die eine nachträgliche Referenzierung ermöglichen.

4.1.1 Manuelle Referenzierung

Der technisch einfachste Weg Positions- und Richtungsdaten für ein Bild zu gewinnen, ist die manuelle Zuordnung bzw. Referenzierung. Dieses Vorgehen ist von der Speicheremethode unabhängig, es ist nur wichtig, dass die Geo-Referenz zwischen Aufnahmeort und Foto einmal hergestellt wird. Da die manuelle Zuordnung stets einen Aufwand für den Anwender bedeutet, sollten die dafür verwendeten Verfahren auf möglichst einfache und intuitive Bedienung ausgelegt sein. Der Referenzierungsprozess könnte dazu beispielsweise durch Karten-Programme wie Google Earth¹⁶ unterstützt werden. Denkbar wäre es, das Bild direkt über eine digitale Landkarte zu ziehen, es über dem Aufnahmeort abzulegen und die Richtung mit einem kleinen Pfeil festzulegen. Die ermittelten Koordinaten und Richtungsdaten könnten dann anschließend im Bild gespeichert werden.

4.1.2 Automatische Referenzierung

Einige mobile Geräte, wie beispielsweise der im Zuge dieser Arbeit verwendete HP iPAQ 6915¹⁷, besitzen heute neben Telefon und Kamera auch einen GPS-Empfänger, mit dem sich der Standort des Anwenders auf wenige Meter genau bestimmen lässt. Durch die Verknüpfung von Kamera und GPS-Empfänger ist es möglich, die aktuellen Koordinaten gleichzeitig mit einem Foto zu speichern.

Mit dem iPAQ gestaltete sich dieser Prozess als sehr einfach. Der Benutzer muss lediglich vor dem Fotografieren die GPS-Funktion des Gerätes aktivieren und warten, bis die eigene Position bestimmt wurde. Danach kann er wie gewohnt Fotos machen, die Speicherung der Positionsdaten erfolgt automatisch innerhalb des Bildes.

¹⁶ <http://earth.google.de>

¹⁷

<http://h20000.www2.hp.com/bizsupport/TechSupport/DocumentIndex.jsp?&lang=en&cc=us&content Type=SupportManual&docIndexId=179166&prodTypeId=215348&prodSeriesId=1822499&lang=en&cc=us>

Einige Kamerahersteller¹⁸ bieten unter ihren Top-Modellen auch Geräte mit kombinierbarem GPS-Modul an. Diese können die Positionsdaten direkt im exif-Header des Bildes speichern, sind aber im Vergleich noch verhältnismäßig teuer.

Die Aufnahme­richtung lässt sich durch den GPS-Empfänger alleine jedoch nicht bestimmen. Für eine automatische Weiterverarbeitung der Richtungsdaten, müsste dieser zusätzlich mit einem elektronischen Kompass ausgerüstet sein. Der im Zuge dieser Arbeit verwendete GPS-Empfänger hat einen solchen Kompass integriert und kann daher auch für die automatische Geo-Referenzierung eingesetzt werden.

4.1.3 Semiautomatische Referenzierung

Unter einem GPS-Tracker versteht man ein Gerät, das die eigene Position in regelmäßigen Abständen ermitteln und speichern kann. Je nach Speicherintervall können alle dazwischen liegenden Positionen bei Bedarf interpoliert werden. Oft nicht größer als eine Streichholzschachtel, protokollieren diese Geräte die Position des Trägers solange genug Satelliten im Empfangsbereich liegen. Falls die Verbindung zwischenzeitlich abreißen sollte, nehmen die Geräte normalerweise ihre Arbeit automatisch wieder auf, sobald sie ein gültiges Signal erhalten. Die so gewonnenen Wegpunkte können anschließend in einem standardisierten Format gespeichert werden. Die Tatsache, dass zu jedem Wegpunkt auch der Zeitpunkt der Messung gespeichert ist, eröffnet dem Anwender nun die Möglichkeit, die Bilder nachträglich zu geo-referenzieren. Hierbei ist es wichtig, dass der Zeitstempel bei dem das Bild aufgenommen wurde, möglichst dem des gespeicherten Wegpunktes des GPS-Trackers entspricht. Da es sich um zwei verschiedene Uhrzeiten handelt, der Uhrzeit der Kamera und der Uhrzeit des GPS-Trackers, ist es notwendig, diese zuvor zu synchronisieren. Wenn dies nicht geschieht, muss bei der Geo-Referenzierung ein Time-Offset, also die Differenz der beiden Uhrzeiten, angegeben werden.

Mittels eines Programms, wie beispielsweise *RoboGEO*¹⁹, lassen sich die beiden Zeitstempel bequem miteinander abgleichen. Die Bilder werden somit auf einfache

¹⁸ Ricoh: http://www.ricohpmmc.com/uk/products/camera/caplio_500SE/caplio_500SE.htm Nikon: http://nikon.de/product/de_DE/products/broad/799/overview.html

¹⁹ <http://www.robogeo.com>

Weise geo-referenziert, obwohl Kamera und GPS-Tracker niemals physisch miteinander verbunden waren.

Die Genauigkeit dieser Methode hängt stark von den Speicherintervallen des GPS-Trackers ab. Bei Intervallen im Minutenbereich, wird die Interpolation bei rascher Positionsänderung nur ungenaue Ergebnisse liefern können. Wird jedoch ein Intervall von wenigen Sekunden gewählt, dann ist zwar der Speicher des Trackers schneller gefüllt, die Aufnahmeposition des Bildes kann aber genauer zwischen zwei Wegpunkten interpoliert werden.

Die Messung der Aufnahmeorientierung funktioniert bei der Nutzung von Intervallen nur bedingt. Da ein Tracker normalerweise separat von der Kamera aufbewahrt wird, beispielsweise in der Tasche eines Rucksacks, kann dieser, selbst wenn er einen elektronischen Kompass integriert hätte, die Aufnahmeorientierung der Kamera nicht messen. Funktionieren würde diese Lösung nur, wenn der Tracker stets mit der Kamera ausgerichtet wird, beispielsweise an der Unterseite befestigt. Gleichzeitig müsste sich der Anwender beim Fotografieren mindestens einen Speicherintervall Zeit lassen um sicherzugehen, dass auch die aktuellen Richtungsdaten im Wegpunkt abgelegt werden.

Um diese Einschränkung zu umgehen, bietet sich die folgende Lösung an. Der Jelbert Geo-Tagger²⁰, siehe Abbildung 6, kann beim Auslösen der Kamera mittels eines aufsteckbaren GPS-Empfängers die Koordinaten und die Aufnahmeorientierung des Gerätes ermitteln und anschließend in einer separaten Datei abspeichern.



Abbildung 6: Jelbert Geo-Tagger¹⁴

²⁰ http://www.geotagger.co.uk/geotagger_website/pics.html

Positions- und Richtungsdaten werden dazu in einem standardisierten Format in einer Textdatei gespeichert. Mittels RoboGeo kann diese anschließend wiederum zur Referenzierung herangezogen werden.

Im Vergleich zur GPS-Tracker basierten Speicherung hat der Jelbert Geo-Tagger den Vorteil, die geografischen Metadaten zeitgleich mit dem Foto messen zu können. Damit sind nahezu idente Zeitstempel möglich und eine Interpolation kann entfallen.

4.1.4 Evaluierung

Die drei vorgestellten Methoden zur Gewinnung von geografischen Metadaten unterscheiden sich hauptsächlich im Aufwand, der für den Benutzer entsteht. Für das Resultat, also ein geo-referenziertes Foto, ist das verwendete Verfahren egal. Am Ende besteht kein Unterschied zwischen einem Bild, das manuell, automatisch oder semiautomatisch referenziert wurde.

Tabelle 4 bietet einen Überblick über wichtige Eigenschaften der genannten Methoden.

Methode	Hardware / Infrastruktur	Genauigkeit Standort	Genauigkeit Richtung	Aufwand für Anwender	Verfügbarkeit
Manuelle Referenzierung	verfügbar	anwender- und anwendungs-abhängig	anwender- und anwendungs-abhängig	hoch	anwender- und anwendungs-abhängig
Automatische Referenzierung mittels GPS	im Entstehen	< 10m	< 1°	niedrig	nur im Freien
Semiautomatische Referenzierung mittels GPS	verfügbar	abhängig vom Aufnahmeintervall	nur möglich wenn Kompass ausgerichtet, dann < 1°	mittel	nur im Freien

Tabelle 4: Gegenüberstellung der genannten Verfahren zur Gewinnung von geografischen Metadaten [10]

Eine Anforderung dieser Arbeit ist die möglichst automatische Geo-Referenzierung der Bilder. Da die manuellen Methoden stets einen hohen Benutzeraufwand verlangen, sind sie, obwohl Techniken dafür bereits verbreitet sind, für diese Arbeit nicht geeignet.

Eine Methode, die für den Benutzer ohne zusätzlichen Aufwand verbunden ist, ist die direkte Verbindung eines GPS-Empfängers und eines elektronischem Kompasses

mit der Kamera. In diesem Fall werden die Positions- und Richtungsdaten gleich beim Auslösen der Kamera in den exif-Header des Bildes geschrieben. Dies ist sicher ein optimaler Ansatz, entsprechende Geräte sind derzeit aber noch rar. Die Einsatzmöglichkeiten und die Genauigkeit der Messungen hängen in diesem Fall nur von GPS-Empfänger und Kompass ab.

Die semiautomatische Methode der nachträglichen Geo-Referenzierung ist hingegen weit verbreitet [24]. Sie hat den Vorteil, dass man außer einem GPS-Tracker mit integriertem, elektronischem Kompass keine neuen Geräte benötigt. Die alte Kamera kann also weiterverwendet werden. Den Abgleich zwischen Wegpunkten und Bildern können Programme wie RoboGeo übernehmen, der zusätzliche Aufwand für den Anwender hält sich daher in Grenzen. Im Vergleich zur automatischen Geo-Referenzierung, hängt die Genauigkeit dieser Methode zusätzlich vom Speicherintervall der Wegpunkte ab. Um ein gutes Ergebnis ohne Interpolation zu bekommen, sollte der Intervall daher nicht mehr als ein paar Sekunden betragen. Wie bereits beschrieben, eignet sich diese Methode auch für die Richtungsmessung, jedoch nur, wenn der Tracker auch immer mit der Kamera ausgerichtet wird.

4.2 Das Kamerasichtfeld

Durch die Verknüpfung von geografischen Metainformationen mit Bildern, wurde nun eine Basis geschaffen, auf der neue Verfahren aufsetzen können, um den Suchprozess nach digitalen Bildern zu verbessern.

Diese Arbeit konzentriert sich auf die Berechnung des Sichtfeldes, also den Bereich vor der Kamera, den der Fotograf festhalten wollte. Die Motivation dazu erklärt sich aus der Tatsache heraus, dass Objekte, die sich innerhalb des Sichtfeldes befinden, viel Information über das Bild selbst preis geben und so zu einer besseren Suche beitragen können.

Unter dem Sichtfeld versteht man den Bereich, der von einer Kamera erfasst und abgespeichert werden kann. Das Sichtfeld wird ausgehend vom Brennpunkt der Kamera entweder durch einen horizontalen und vertikalen bzw. einem diagonalen Feldwinkel aufgespannt. Abbildung 7 verdeutlicht diesen Sachverhalt.

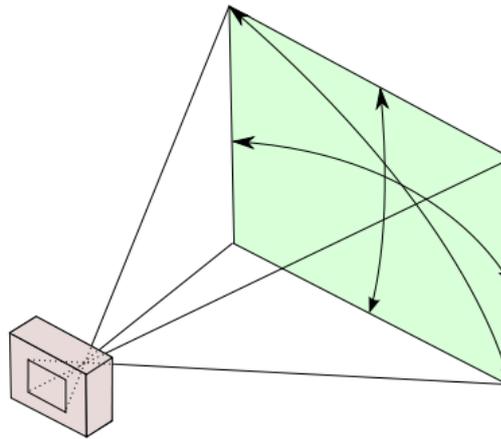


Abbildung 7: Sichtfenster das durch horizontalen, vertikalen und diagonalen Sichtwinkel aufgespannt wird. [25]

Die Größe des Sichtfeldes einer Kamera hängt im Allgemeinen von drei Faktoren ab, der Brennweite, der Größe des Bildsensors und der Linsenkrümmung [25]:

Die Brennweite entspricht dem Abstand zwischen dem Brennpunkt der Linse und dem Bildsensor. Mittels der Brennweite kann der Fotograf den Bildausschnitt einer Aufnahme verändern, hohe Brennweiten führen zu einem kleinen, niedrige Brennweiten zu einem großen Bildausschnitt. Tabelle 5 zeigt einige typische Brennweiten und die damit zusammenhängende Veränderung des Bildausschnittes.

Der Bildsensor ist der Teil einer Kamera, auf den die Strahlen des Objektivs gelenkt werden. Sein Format und seine Größe beeinflussen den Bereich, den eine Kamera abbilden kann, maßgeblich. In der Fotografie beziehen sich viele Angaben auf das in der Analogfotografie verwendete Kleinbildformat²¹. Da die Bildsensoren in der Digitalfotografie aber meist viel kleiner sind als ihre analogen Pendanten, wurde der Begriff des *Digital Multipliers* eingeführt, der das Verhältnis zwischen den beiden Bildsensorgößen ausdrückt. Der Sensor einer 6 Megapixel Kamera ist beispielsweise ca. 1,5 mal so klein wie der einer Kleinbildkamera, der eines Handys sogar ca. 4 mal so klein.²²

Die Linsenkrümmung beschreibt die Art wie eine Linse geschliffen wurde. Fotografen benutzen hauptsächlich zwei Arten von Linsen: *Rectilinear*- und *Fisheye*-Linsen, die sich im Krümmungsverhalten unterscheiden. Während erstere die

²¹ Das Kleinbildformat sieht einen Film mit der Größe von 36mm x 24mm vor.

²² siehe: <http://www.dpreview.com/learn/?/key=focal+length+multiplier>

Strahlen ungebrochen durch die Linse leiten, bricht zweitens alle Strahlen, die nicht direkt durch das Zentrum der Linse verlaufen [26]. Aufgrund der verzerrten Darstellung, die Fotos von Fisheye-Linsen aufweisen, beschränkt sich diese Arbeit mit der Untersuchung des Sichtfeldes von Rectilinear-Linsen.

Da die Größe des Sichtfeldes eines Bildes ausschließlich von der Größe des Feldwinkels abhängt, ist es notwendig diesen genau berechnen zu können. Der diagonale Feldwinkel $2w$ errechnet sich bei der Verwendung einer Rectilinear-Linse aus dem Aufnahmeformat d und der Brennweite f der Kamera: [27]

$$2w = 2 * \arctan \left(\frac{d}{2 * f} \right)$$

Auffallend ist, dass der Feldwinkel ausschließlich von der Brennweite und der Größe des Bildsensors abhängt. Da innerhalb einer Kamera immer der gleiche Bildsensor verwendet wird, hängt das Sichtfenster daher ausschließlich von der benutzten Brennweite ab. Um einen Vergleich zwischen digitaler und analoger Fotografie ziehen zu können, ist auch hier der *Digital Multiplier* wichtig, da sich die Angaben zu den Brennweiten eines Objektivs immer auf das Kleinbildformat beziehen.

Tabelle 5 bietet einen Überblick über die Veränderung des Blickwinkels abhängig von der verwendeten Brennweite. Alle Angaben beziehen sich auf das Kleinbildformat.

Brennweite (mm)	13	15	18	21	24	28	35	50	85	105	135	180
Diagonaler Winkel	118	111	100	91.7	84.1	75.4	63.4	46.8	28.6	23.3	18.2	13.7
Vertikaler Winkel	85.4	77.3	67.4	59.5	53.1	46.4	37.8	27.0	16.1	13.0	10.2	7.63
Horizontaler Winkel	108	100.4	90.0	81.2	73.7	65.5	54.4	39.6	23.9	19.5	15.2	11.4

Tabelle 5: Veränderung des Blickwinkels abhängig von der Brennweite (Kleinbildformat), [25]

Sobald der Feldwinkel bekannt ist, kann man mit der Berechnung des Sichtfeldes beginnen. Obwohl das Sichtfeld eigentlich im dreidimensionalen Raum liegt, begnügt sich diese Arbeit mit der zweidimensionalen Betrachtung. Dies resultiert aus der Tatsache, dass die im Zuge dieser Arbeit verwendeten Geräte über keinerlei Neigungssensoren verfügen und dies auch keine Anforderung an die Arbeit darstellt. Das Sichtfeld wird also durch einen ebenen Feldwinkel, den Aufnahmeort und die

Aufnahmerichtung bestimmt. Abbildung 8 zeigt, wie das 3-dimensionale Sichtfeld in den 2-dimensionalen Raum überführt wird.

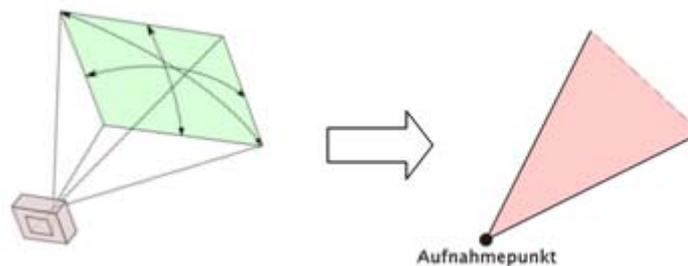


Abbildung 8: Überführung eines 3-dimensionalen Sichtfeldes in den 2-dimensionalen Raum

Wenn man die Koordinaten des Aufnahmepunktes und die Aufnahmerichtung kennt, dann befinden sich alle Objekte, wie Gebäude oder Sehenswürdigkeiten, die auf dem Foto zu sehen sind, innerhalb des durch den Feldwinkel aufgespannten Sichtfeldkegels. Damit dieser nicht bis in die Unendlichkeit reicht, muss eine Distanz angegeben werden, die die Entfernung zwischen der Basis des Kegels und dem Aufnahmepunkt beschränkt. Die Distanz sollte immer so gewählt werden, dass das Sichtfeld gerade noch alle auf dem Foto abgebildeten Objekte abdeckt. Ist die Distanz festgelegt, kann das Sichtfelddreieck vollständig berechnet werden. Ausgehend von den Koordinaten des Aufnahmepunktes des Bildes, müssen für das Dreieck noch die Koordinaten zweier weiterer Punkte gefunden werden. Diese Arbeit nennt sie Zielpunkt links und Zielpunkt rechts. Dazu wird im ersten Schritt ein Hilfspunkt berechnet, von dem aus sich die zwei unbekannt Punkte leichter errechnen lassen. Dieser liegt genau die angegebene Distanz in Aufnahmerichtung vom Aufnahmepunkt entfernt. Von diesem Hilfspunkt aus, werden nun der linke und rechte Zielpunkt berechnet. Diese liegen auf einer Linie, die orthogonal zur Aufnahmerichtung ausgerichtet ist. Die Entfernung a vom Hilfspunkt aus kann über den Sinussatz berechnet werden:

$$a = \frac{\text{Distanz } d}{\sin\left(90 - \frac{\text{Feldwinkel } w}{2}\right)} * \sin\left(\frac{\text{Feldwinkel } w}{2}\right)$$

Abbildung 9 verdeutlicht noch einmal das Vorgehen anhand eines Beispiels. Die 0° Achsen beim Aufnahme- bzw. Hilfspunkt geben die Nordrichtung an - das Bild wurde also in Richtung Nord-Ost (45°) aufgenommen.

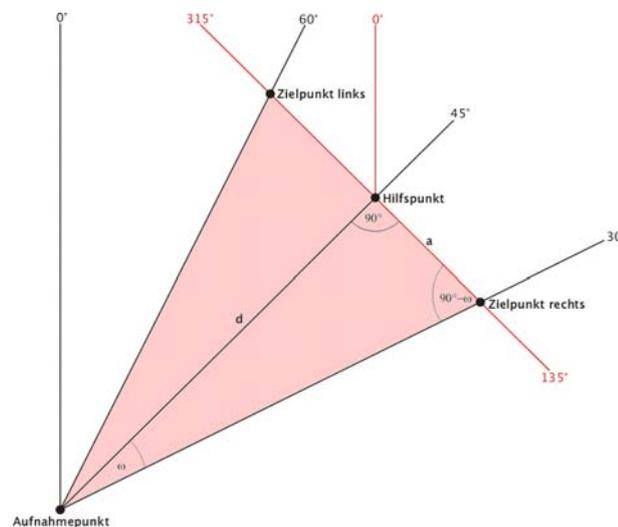


Abbildung 9: Berechnung des Sichtfeldes

Die beiden Zielpunkte können nun, ausgehend vom Hilfspunkt, in Richtung 135° bzw. 315° mit jeweils der Länge a bestimmt werden.

Um das Sichtfeld automatisch berechnen zu können sind also vier Parameter notwendig: die Koordinaten des Aufnahmeortes, die Aufnahmerichtung, der Feldwinkel und die Distanz, die die Länge des Dreiecks festlegt. All diese Parameter können in bereits vorhandenen Metadatenstandards, wie exif oder XMP, gespeichert oder daraus gewonnen werden.

Einen ähnlichen Ansatz verfolgen auch die Autoren des Papers „*Beyond Location Based – The Spatially Aware Mobile Phone*“ [28]. In ihrer Arbeit nutzen sie ein eigenes Visibility Model, mit dem festgestellt werden kann, wohin in der realen Welt ein Anwender mit seinem Mobiltelefon zeigt. Diese Daten werden in Echtzeit verarbeitet und können dem Benutzer Informationen über das Objekt, auf das er gerade zeigt, liefern. Die Echtzeitfähigkeit des Systems und die Tatsache, dass für die Identifikation von Objekten das Mobiltelefon nicht als Kamera sondern als Zeigestab verwendet wird, grenzen deren Paper jedoch von dieser Arbeit ab.

4.2.1 Evaluierung

Das größte Problem bei der Bestimmung des Feldwinkels ist, dass man dafür das Aufnahmeformat der Kamera, also die Bildsensorgröße, kennen muss. Diese variiert im Gegensatz zu Analogkameras die meistens das Kleinbildformat verwenden bei digitalen Geräten stark und man ist auf die Angaben des Herstellers angewiesen. Der exif Standard definiert zu diesem Zweck die Felder *FocalPlaneXResolution*, *FocalPlaneYResolution* und *FocalPlane-ResolutionUnit*. Diese geben die Länge bzw. Breite des Bildsensors an. Leider werden diese Informationen nicht von allen Kameras bereitgestellt, siehe „Vergleich von drei Digitalkameras“, Anhang I. In diesem Fall kann eventuell noch das technische Datenblatt weitere Details liefern. Ist die Größe des Sensors unbekannt, kann der Feldwinkel nur geschätzt werden.

Sind Feldwinkel, Aufnahmeort und Aufnahmerichtung bekannt, ergibt sich daraus ein Sichtfeldkegel der die auf dem Foto abgebildeten Objekte beinhaltet. Dies ist bereits ein wichtiger Schritt in der Verbesserung des Suchprozesses. Sind nämlich die Koordinaten der gesuchten Objekte bekannt, kann nun überprüft werden ob sie sich wirklich innerhalb des Kegels befinden oder nicht. Im Vergleich zu einer reinen Umkreissuche, bei der nur die Nähe der Objekte zum Aufnahmeort des Fotos überprüft werden kann, kann nun zusätzlich festgestellt werden ob sie im Sichtfeld des Fotos liegen oder nicht.

Damit die Größe des Sichtfeldes nicht im Unendlichen liegt, sollte dieses durch eine Distanz begrenzt werden. Erst durch diese Begrenzung kann sichergestellt werden, dass nicht auch gänzlich verdeckte Objekte im Hintergrund, fälschlicherweise als Bildinhalt identifiziert werden. Manche Kameras können dazu die gemessene Distanz des Autofokus im exif-Header speichern, siehe Anhang I. Dies wäre ein erster Ansatz für die automatische Bestimmung der Länge des Sichtfelds, denn der Autofokus stellt sich normalerweise auf den Entfernungsbereich scharf, der in der Mitte des Bildes liegt. Problematisch könnten sich jedoch Fotos mit einer hohen Tiefenschärfe erweisen, da hier der Schärfebereich im Unendlichen liegt und somit keine zuverlässige Distanzmessung erfolgen kann. Bevor nicht geklärt ist, wie gut die Distanzmessung in der Praxis funktioniert, muss die Entfernung daher noch manuell bestimmt werden.

Kapitel 5

Beispielapplikation zur Verbesserung des Suchprozesses

In diesem Kapitel wird nun anhand einer Applikation die im Zuge dieser Arbeit entwickelt wurde, konkret untersucht, ob das Suchergebnis für den Anwender verbessert werden kann, wenn bei der Suche nach einem bestimmten Foto dessen Sichtfeld, wie es in Kapitel 4.2 beschrieben wird, berücksichtigt wird. Da die dafür verwendeten Bilder bereits alle geografische Informationen aufweisen müssen wird außerdem gezeigt, wie sich geo-referenzierte Fotos in der Praxis gewinnen lassen.

Im Zuge dieser Arbeit wurden die Applikationen *FoVA* und *GPSTool* entwickelt, die den gesamten Arbeitsfluss, also Gewinnung, Speicherung und Auswertung von geografischen Metadaten abbilden kann. Ziel von FoVA ist es, die folgenden beiden Anwendungsfälle der Motiverkennung und der Motivsuche zu erfüllen.

Bei der Motiverkennung fotografiert der Anwender eine ihm unbekannte Sehenswürdigkeit und würde gerne wissen, um was es sich dabei handelt. Um das herauszufinden, lädt er das Bild über eine Website auf einen Server hoch, der anschließend die Aufnahmeposition und –richtung ausliest. Nach einem Vergleich mit den in einer Datenbank gespeicherten Koordinaten der Sehenswürdigkeiten kann ihm das System nun mitteilen, welche Sehenswürdigkeit(en) er höchstwahrscheinlich fotografiert hat.

Die Motivsuche gibt einem Anwender die Möglichkeit, Fotos einer bestimmten Sehenswürdigkeit zu suchen. Dazu gibt der Benutzer den Namen einer

Sehenswürdigkeit ein und startet den Suchprozess. Das System sucht nun die Koordinaten der Sehenswürdigkeit und vergleicht sie mit den Aufnahmeorten und der Aufnahmerichtung der gespeicherten Bilder. Zurückgeliefert werden alle Bilder, die mit hoher Wahrscheinlichkeit die gesuchte Sehenswürdigkeit darstellen.

Große Bedeutung für die Realisierung der beiden Anwendungsfälle kommt der Bestimmung des Sichtfeldes eines Fotos zu, denn nur mit ihm kann überhaupt festgestellt werden, ob sich eine Sehenswürdigkeit auf einem Bild befindet oder nicht.

Der folgende Abschnitt gliedert sich in drei Teile. Zu Beginn werden Annahmen und Limitationen besprochen, die während der Entwicklung von FoVA Bestand hatten. Danach folgt die Beschreibung von FoVA selbst. Dabei wird gezeigt, wie FoVA zu einer Verbesserung des Suchprozesses beitragen kann. Der letzte Abschnitt beschreibt ausführlich, wie die um geografische Metadaten angereicherten Bilder und die Standorte bestimmter Sehenswürdigkeiten als Testdaten gewonnen wurden.

5.1 Annahmen und Limitationen

Zur Bestimmung der Positionsdaten wird ausschließlich das GPS-System verwendet. Das hat den Nachteil, dass unter schlechten Empfangsbedingungen die ermittelte Position um bis zu 10 Meter verschoben sein kann. Der Empfang ist außerdem nur im Freien möglich, was für die Realisierung der Anwendungsfälle jedoch genügt.

Innerhalb dieser Arbeit ist es nicht Ziel von FoVA, die Position aller Sehenswürdigkeiten zu kennen. Es wurde nur eine kleine Auswahl, die jedoch zeigt, dass die beiden Anwendungsfälle grundsätzlich realisiert werden können.

Aufgrund fehlender Angaben zur Größe des Bildsensors in den beiden verwendeten PDA-Modellen, konnte der Feldwinkel nur geschätzt werden. Bei echten Digital – bzw. Spiegelreflexkameras ist diese jedoch oft bekannt und kann benutzt werden um den Feldwinkel genauer zu berechnen.

Da die Browserunabhängigkeit keine Anforderung für dieses Projekt war, wurde sie auch nicht in den Vordergrund gestellt. Während der Entwicklung kam

ausschließlich der Firefox-Browser in Version 2.0²³ zum Einsatz, andere Programme wurden nicht getestet.

5.2 FoVA – The Field of View Analyzer

FoVA wurden mittels eines Web-Frontends realisiert. Als Back-End fungieren ein Apache Webserver inklusive PHP und MySQL. Zur Darstellung genügt ein aktueller Browser mit aktiviertem Javascript. Das verwendete Kartenmaterial stammt von Google Maps²⁴, welches über entsprechende JavaScript-Funktionen eingebunden werden darf.

5.2.1 Programmfunktionen

Ziel von FoVA ist es, den Anwender bei der Suche nach bestimmten Sehenswürdigkeiten oder Gebäuden einer Stadt zu unterstützen. Das Programm bietet ihm dazu folgende Funktionen:

Die Funktion *Upload Image* ermöglicht es dem Anwender ein Bild mit geografischen Metadaten auf den Server hochzuladen und zu speichern. Nach dem Upload hat der Anwender die Möglichkeit, die aus dem exif-Header ausgelesenen Positionsdaten zu überprüfen und das als rotes Dreieck dargestellte Sichtfeld zu beeinflussen. Dieser Zwischenschritt ist notwendig, da aufgrund der fehlenden Bildsensorgröße der verwendeten Aufnahmegereäte, der Feldwinkel nur geschätzt werden kann. Aufgrund der baugleichen Linsenoptik, wurde für die beiden Geräte ein Winkel von ca. 60° angenommen. Dieser wird auch als Vorgabewert verwendet. Die Entfernung zur Sehenswürdigkeit muss mangels Entfernungsmesser ebenfalls geschätzt und manuell eingetragen werden. Da es sich bei den Fotos um keine Landschaftsaufnahmen handelt, wurde hier ein Vorgabewert von 200 Meter gewählt.

Abbildung 10 zeigt wie das Sichtfelddreieck durch die Veränderung des Feldwinkels vom Benutzer angepasst werden kann.

²³ <http://www.firefox-browser.de/windows.php>

²⁴ <http://maps.google.com>



Abbildung 10: Auswirkung verschiedener Feldwinkel auf die Größe des Sichtfeldes

Im linken Bild beträgt der Feldwinkel 60° , im rechten 30° . Die Entfernung von der Basis des Dreiecks zum Aufnahmepunkt beträgt jeweils 200 Meter. An diesem Beispiel ist deutlich zu sehen, wie stark sich der Feldwinkel auf die Größe des daraus resultierenden Sichtfeldes auswirkt.

Die Funktion *Show Map* zeigt alle in der Datenbank gespeicherten Bilder und Sehenswürdigkeiten auf einer großen Karte an. Abbildung 11 zeigt einen Kartenausschnitt mit vier Fotos.



Abbildung 11: Kartenausschnitt mit vier Bildern

Search for images using the database liefert dem Anwender alle Bilder, auf denen eine bestimmte Sehenswürdigkeit zu sehen ist. Nach der Eingabe eines Suchbegriffes wird dieser in einer Datenbank, die die Namen und Standorte ausgewählter Sehenswürdigkeiten enthält, gesucht. Wenn ein entsprechender Eintrag gefunden wurde, dienen dessen geografische Koordinaten als Basis für den nächsten Schritt. Dabei werden alle Bilder dahingehend überprüft, ob die Sehenswürdigkeit innerhalb des Sichtfeldes des Fotos liegt. Wenn ja, dann wird sie dem Suchergebnis als Treffer hinzugefügt. Abbildung 12 zeigt eine typische Suchanfrage nach Fotos einer bestimmten Sehenswürdigkeit.

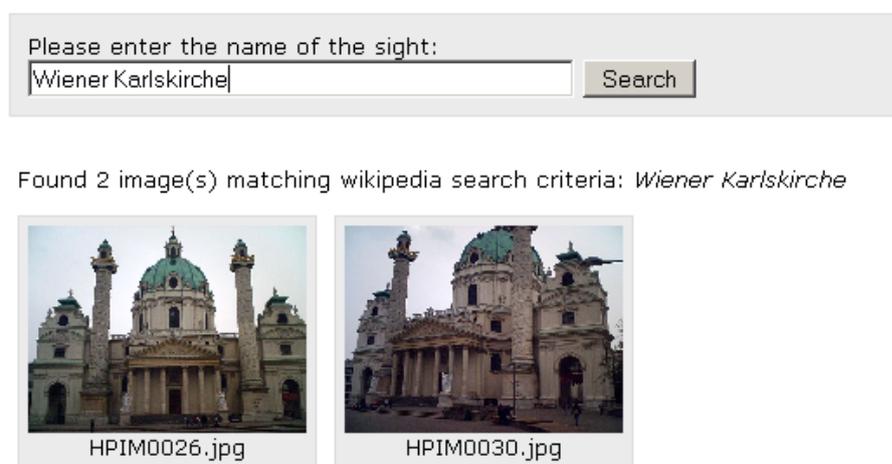


Abbildung 12: Suche nach Fotos von der Wiener Karlskirche

Search for images using Wikipedia funktioniert genau so wie die Suche mittels Datenbank, verwendet aber als Datenbasis die Wikipedia. Wie diese als Datenbasis verwendet werden kann, wird in Kapitel 5.5.2 näher beschrieben. Nach der Eingabe des Suchbegriffes wird seine geografische Position mit Hilfe der Wikipedia ermittelt. Anschließend wird, wie zuvor schon beschrieben, das Sichtfeld aller Bilder analysiert, um dem Anwender nur die relevanten Fotos zu präsentieren. Eine solche Suche kann mehrere Resultate hervorbringen, da es beispielsweise neben der Karlskirche in Wien auch eine in Kassel gibt. Angezeigt werden aber nur die Resultate von denen es auch Bilder gibt.

Identify image using the database ermöglicht es dem Anwender festzustellen, welche Sehenswürdigkeit er fotografiert hat. Dazu muss er zuerst das Bild auf den Server hochladen. Anschließend werden die Positions- und Richtungsdaten aus dem exif-

Header ausgelesen und das Sichtfeld erstellt. Nun wird für alle Sehenswürdigkeiten, die in der Datenbank gespeichert sind überprüft, ob sie innerhalb des Sichtfeldes liegen. Wenn ja, dann wird auf der Karte neben dem Bild auch ein kleines Haus-Icon angezeigt, das die gefundene(n) Sehenswürdigkeit(en) repräsentiert. Ein Beispiel dazu ist auf Abbildung 13 zu sehen.



Abbildung 13: Erfolgreiche Identifizierung einer Sehenswürdigkeit die auf dem Foto zu sehen ist.

Identify image using Wikipedia funktioniert grundsätzlich wie die zuvor beschriebene Identifizierung mittels Datenbank. Die Positionsdaten der Sehenswürdigkeiten stammen aber vom Datenbestand der Wikipedia. Um diese zu bekommen, muss ein Reverse Geocoding, siehe Kapitel 4.1, durchgeführt werden. In diesem Falle werden alle Sehenswürdigkeiten der Wikipedia gesucht, die innerhalb eines definierten Umkreises von der Aufnahmeposition des Bildes liegen. Wenn es für dieses Kriterium Resultate gibt, muss nur noch überprüft werden, ob sie innerhalb des Sichtfeldes liegen.

5.2.2 Grundlagen zur Nutzung des Sichtfeldes

Um rechnerisch feststellen zu können, ob sich eine bestimmte Sehenswürdigkeit innerhalb des Sichtfeldes befindet oder nicht, gibt es verschiedene Methoden²⁵. Das Problem kann auf die Frage reduziert werden, ob ein zu überprüfender Punkt innerhalb eines Dreiecks liegt oder nicht. Da es sich bei dem Prototypen um keine zeitkritische Anwendung handelt, wurde für die Berechnung folgende Methode verwendet.

Es wird für alle Seitenlinien des Dreiecks in sequentieller Reihenfolge überprüft, ob ein bestimmter Punkt links, rechts oder genau auf der Seitenlinie des Dreiecks liegt. Abbildung 14 verdeutlicht diesen Sachverhalt mit einer Grafik.

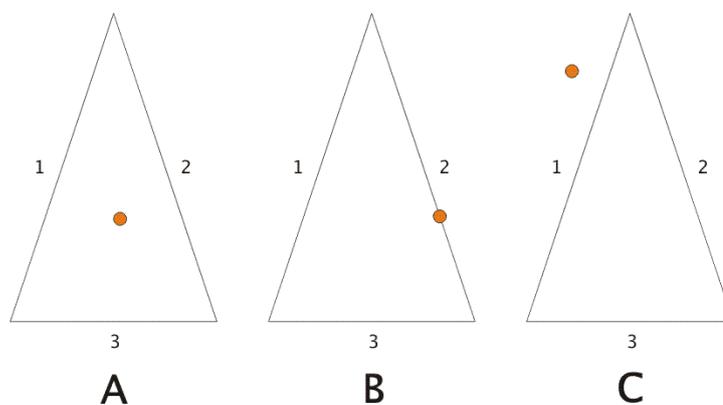


Abbildung 14: Drei möglichen Lagen eines Punktes zu einem Dreieck

Wenn der Punkt von allen Linien aus gesehen auf der gleichen Seite liegt, Abbildung 14 (A), dann ist er innerhalb des Dreiecks. Liegt er für eine Seitenlinie auf einer anderen Seite, Abbildung 14 (C), dann ist der Punkt außerhalb. Einen Sonderfall stellt die Situation dar, wenn der Punkt genau auf einer Seitenlinie des Dreiecks liegt, Abbildung 14 (B). In diesem Fall wird er als innenliegend gewertet.

Um die Lage eines Punktes zu einer Geraden zu beurteilen, genügt es die Steigungen zwischen der Geraden mit der Steigung, die zwischen dem Anfangspunkt der Geraden und dem Punkt vorliegt, zu vergleichen. Sind beide Steigungen gleich, dann muss der Punkt auf der Geraden liegen, ist die Steigung zum Punkt geringer dann liegt er rechts, ist die Steigung größer liegt er links. Abbildung 15 demonstriert diesen Sachverhalt an einer Grafik.

²⁵ Eine Übersicht bietet beispielsweise Ralph Wiebicke unter: <http://rw7.de/ralf/inffaq/polygon.htm>

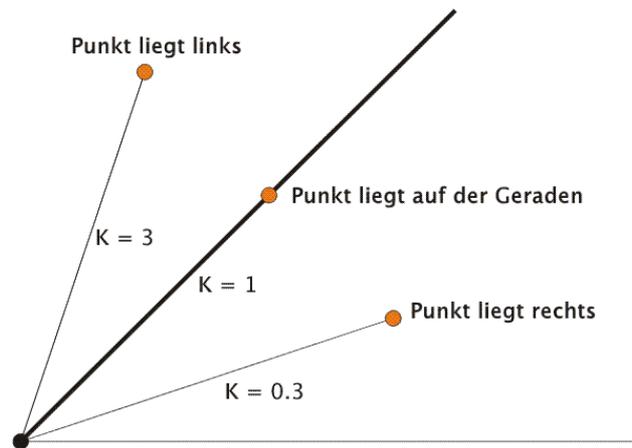


Abbildung 15: Drei möglichen Lagen eines Punktes zu einer Geraden im 2-dimensionalen Raum

Bei einer Steigung kleiner 1, liegt der Punkt in diesem Beispiel rechts von der Geraden, bei einer Steigung größer 1, liegt er links. Ist die Steigung gleich 1, liegt der Punkt genau auf der Geraden.

5.2.3 Gewonnene Erkenntnisse

Die Größe des Sichtfeldes sollte immer so gewählt werden, dass der Feldwinkel der verwendeten Brennweite und Bildsensorgröße entspricht, siehe Kapitel 4.2, und die eingestellte Entfernung den Großteil des abgebildeten Objekts bedeckt. Wenn die Entfernung zu kurz gewählt wird, kann es passieren, dass die Sehenswürdigkeit nicht im Sichtfeld liegt, obwohl sie eigentlich auf dem Foto erkennbar ist. Dieses Verhalten entsteht aus der Tatsache heraus, dass derzeit nur eine Position zu einer Sehenswürdigkeit gespeichert wird. Abbildung 16 demonstriert diesen Sachverhalt.

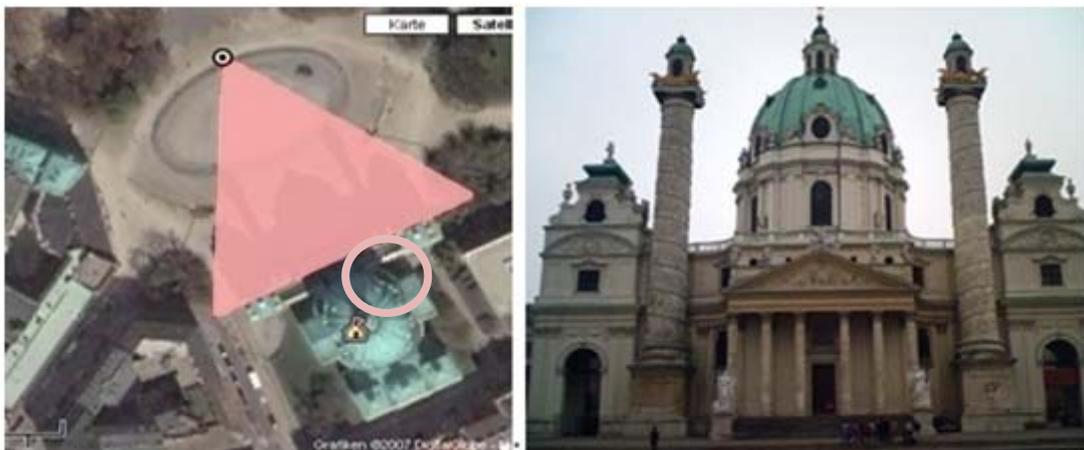


Abbildung 16: Probleme für FoVA die durch eine zu kurze Entfernungsangabe entstehen können

Das kleine Häuschen im Kreis das auf Abbildung 16 links zu sehen ist, symbolisiert die gespeicherte Position der Wiener Karlskirche. Da die Entfernung absichtlich zu kurz gewählt wurde (80 Meter) liegt die Kirche für FoVA nicht innerhalb des Sichtfeldes obwohl auf dem Foto selbst ausschließlich die Karlskirche zu sehen ist.

Hand in Hand mit diesem Problem geht die Tatsache, dass durch die Verwendung einer einzigen Distanz keine Sichtbarkeitsanalyse durchgeführt werden kann. Wenn auf dem Foto beispielsweise ein Baum ein sehenswertes Gebäude vollständig verdeckt, dann liefert die Suche nach dem Gebäude zwar ein Resultat, zu sehen wird aber nur der besagte Baum sein.

Trotzdem wird durch das Einbeziehen der Aufnahmerichtung und der Distanz des Sichtfeldes ein wichtiges Filterkriterium geschaffen, das die Resultate der Suche drastisch verbessern kann. Bei allen Suchanfragen wird zunächst nämlich nur eine reine Umkreissuche durchgeführt. Es werden also alle Resultate zurückgeliefert, die sich innerhalb eines Kreises um den Aufnahmepunkt befinden. Für all diese Resultate kann aber nun zusätzlich überprüft werden, ob sie innerhalb des Sichtfeldes der Kamera liegen oder nicht. Das Suchergebnis kann also weiter eingeschränkt werden. Abbildung 17 und Abbildung 18 verdeutlichen diese Tatsache.



Abbildung 17: Gebäude der österreichischen Nationalbank

Abbildung 17 zeigt das Gebäude der österreichischen Nationalbank, das von FoVA automatisch identifiziert werden soll. In Abbildung 18 ist zu sehen, dass ohne das Einbeziehen des Sichtfeldes, alle Sehenswürdigkeiten im Umkreis von 200 Metern zurückgeliefert werden. Der Anwender muss nun selbst entscheiden, welches Resultat für ihn das Richtige ist. Wenn allerdings das der Aufnahmerichtung

entsprechend ausgerichtetes Sichtfeld zur Verfügung steht, dann können die Resultate entsprechend gefiltert werden und der Benutzer muss sich zwischen weniger Alternativen entscheiden.



Abbildung 18: Vergleich zwischen einer Suche ohne und einer Suche mit Berücksichtigung des Sichtfeldes.

In diesem Beispiel werden die Wahlmöglichkeiten von 6 Sehenswürdigkeiten, die innerhalb von 200 Metern um den Aufnahmeort liegen, auf 2 reduziert. Diese Verbesserung des Suchergebnisses kann mit Methoden aus dem Information-Retrieval auch absolut berechnet werden. In diesem Fall wird durch FoVA die Präzision des Suchergebnisses, die so genannte Precision, erhöht. Sie ist definiert als der Anteil der gefundenen relevanten Objekte unter allen gefundenen Objekten²⁶. Eine Precision von 1 bedeutet, dass sich ausschließlich relevante Objekte im Suchergebnis wiederfinden. Ist die Precision gleich 0, dann enthält das Suchergebnis keinen relevanten Treffer.

Für die Suche ohne das Sichtfeld, ergibt sich für das genannte Beispiel folgende Precision:

$$Precision\ ohne\ Sichtfeld = \frac{Relevante\ Objekte}{Gefundene\ Objekte} = \frac{1}{6} = 0,1667$$

²⁶ siehe: http://information-retrieval.de/irb/ir.part_1.chapter_3.section_7.subdiv1_3.html

Die gleiche Suche unter Berücksichtigung des Blickfeldes ergibt einen deutlich besseren Wert:

$$\textit{Precision mit Sichtfeld} = \frac{\textit{Relevante Objekte}}{\textit{Gefundene Objekte}} = \frac{1}{2} = 0,5$$

Anhand dieser Zahlen kann festgehalten werden, dass die Berücksichtigung des Blickfeldes im Vergleich zu einer reinen Umkreissuche zu einer besseren Precision innerhalb der Suche führen kann. In diesem Beispiel wurde die Precision durch FoVA um den Faktor 3 verbessert.

Neben der Precision, kann auch die Vollständigkeit einer Suche überprüft werden. Der so genannte Recall ist definiert als der Anteil der relevanten gefundenen Objekte an allen existierenden relevanten Objekten. Ein Recall von 1 gibt an, dass sich alle relevanten Objekte auch im Suchergebnis wiederfinden. Ein Wert von 0 bedeutet, dass sich im Suchergebnis kein einziger relevanter Treffer befindet.

Für die Suche ohne das Sichtfeld, ergibt das genannte Beispiel folgenden Recall:

$$\textit{Recall ohne Sichtfeld} = \frac{\textit{Gefundene relevante Objekte}}{\textit{Relevante Objekte}} = \frac{1}{1} = 1$$

Die gleiche Suche unter Berücksichtigung des Blickfeldes ergibt:

$$\textit{Recall mit Sichtfeld} = \frac{\textit{Gefundene relevante Objekte}}{\textit{Relevante Objekte}} = \frac{1}{1} = 1$$

An diesem Beispiel ist zu sehen, dass der Recall in beiden Fällen optimal ist und sich das relevante Objekt immer innerhalb des Suchergebnisses befindet. Voraussetzung dafür ist allerdings, dass das Sichtfeld richtig berechnet und ausgerichtet ist. Ist das Sichtfeld so kurz gewählt, dass es nicht bis zum gesuchten Objekt reicht, wird der Recall in beiden Fällen gleich 0 sein, da das Suchergebnis kein relevantes Objekt enthalten wird.

Abschließend lässt sich an dieser Stelle festhalten, dass die Berücksichtigung des Sichtfeldes zu einer Verbesserung des Suchergebnisses führt. Durch das Sichtfeld wird ein zusätzliches Kriterium geschaffen, das nicht relevante Treffer aus dem Suchergebnis eliminieren und dieses dadurch präziser machen kann. Die Vollständigkeit der Suche wird, zumindest in diesem Beispiel, durch das Sichtfeld weder verbessert noch verschlechtert.

5.2.4 Verbesserungsmöglichkeiten

Beim Upload eines Bildes muss der Benutzer derzeit Feldwinkel und Distanz für das Sichtfeld selbst eingeben. Dieser Prozess könnte zumindest teilautomatisiert werden. Falls ein Bild die Bildsensorgröße der Kamera gemäß der exif-Spezifikation beinhaltet, könnte diese zusätzliche Information ausgelesen werden und zumindest der Feldwinkel automatisch berechnet werden.

Obwohl manche Kameras die gemessene Distanz des Autofokus im exif-Header speichern können, ist es nicht zu empfehlen, diesen Wert für das Sichtfeld zu verwenden, siehe Kapitel 4.2. Für eine genauere Distanzmessung würden sich Geräte mit lasergestützten Entfernungsmessern besser eignen. Inwieweit solche Systeme in eine Kamera integriert werden können, müssen weitere Untersuchungen zeigen. Bis dahin sollte die Entfernung daher noch mit Hilfe von FoVA geschätzt und manuell eingetragen werden.

5.3 GPSTool

Wie bereits in Kapitel 4.1 erwähnt, gibt es verschiedene Methoden um Bilder mit geografischen Metadaten anzureichern. Für diese Arbeit kamen aus Gründen der festgelegten Anforderungen nur zwei davon in Frage: das automatische Speichern der Daten beim Drücken des Auslösers bzw. das nachträgliche Verknüpfen per Zeitstempel. Die manuelle Zuordnung des Aufnahmeortes wäre prinzipiell auch möglich, schied aber, aufgrund der Anforderung an einen möglichst automatischen Prozess, aus.

Erstere Methode hat den großen Vorteil, dass die Daten sofort im exif-Header des Bildes gespeichert werden und der Benutzer keinerlei Mehraufwand betreiben muss. Der im Zuge dieser Arbeit verwendete HP iPAQ 6910 erfüllt diese Anforderung. Der Anwender kann mit der mitgelieferten Kamera-Software und dank des integrierten GPS-Empfängers auf einfache Weise geo-referenzierte Bilder aufnehmen. Leider fehlt dem iPAQ jedoch eine entscheidende Komponente, um alle Anforderungen dieser Arbeit zu erfüllen: Er besitzt keinen elektronischen Kompass, mit dem sich die Aufnahmerichtung feststellen lässt. Dies ist jedoch besonders wichtig, da nur damit die Richtung des Sichtfeldes bestimmt werden kann.

Um nun neben den Koordinaten auch die Aufnahmerichtung speichern zu können, kam ein externer GPS-Empfänger mit integriertem elektronischem Kompass zum Einsatz. Zur Anwendung kam der WBT-100 der Firma Wintec²⁷, der alle geforderten Eigenschaften erfüllt, siehe Anhang III. Da in dieser Arbeit nur mit einem zweidimensionalen Sichtfeld gearbeitet wird, ist es auch kein Problem, dass der Kompass nicht neigungskompensiert ist.

Aufgrund der fehlenden Möglichkeit zur Speicherung der Aufnahmerichtung, konnte die HP-eigene Kamerasoftware nicht mehr verwendet werden. Im Zuge dieser Arbeit wurde daher die Applikation GPSTool geschrieben. Diese ermöglicht auf dem PDA die Gewinnung und Speicherung der Testbilder inklusive ihrem Standort und der Aufnahmerichtung durchzuführen.

5.3.1 Programmbeschreibung

Das GPSTool ist ein kleines Programm für PDAs, die Windows Mobile 5 als Betriebssystem verwenden. Es ermöglicht dem Nutzer mit seinem Gerät Fotos inklusive geografischer Metadaten zu machen. Jedes mal, wenn er auf den Auslöseknopf drückt, werden die aktuelle Position und die Aufnahmerichtung vom GPS-Empfänger ausgelesen und neben dem Foto in einer separaten Datei gespeichert. Foto und Wegpunkte werden anschließend in einem eigenen Schritt am Computer geo-referenziert.

Als Entwicklungsplattform kam das .Net Framework²⁸ von Microsoft zum Einsatz da es dem Programmierer standardisierte Schnittstellen zur Hardware des PDAs anbietet. Damit ist es möglich, direkt aus dem Programm heraus die Bluetoothverbindung zum GPS-Empfänger zu steuern und die integrierte Kamera des PDAs zu nutzen. Als Programmiersprache innerhalb des Frameworks wurde aufgrund von Vorkenntnissen C# gewählt.

Das Programm selbst ist einfach zu benutzen. Der Benutzer wählt einen COM-Port und verbindet sich per Bluetooth mit dem GPS-Empfänger. Nach erfolgreicher Verbindung zeigt das Gerät Werte wie die Uhrzeit, die ermittelte Position und die Ausrichtung des Empfängers an. Nach dem Drücken von „Take Picture“, kann der

²⁷ http://www.wintec-gps.de/wintec_wbt-100.php

²⁸ <http://msdn2.microsoft.com/de-de/netframework/default.aspx>

Benutzer die integrierte Kamera des Gerätes nutzen, um ein Bild zu machen. Sobald dieses abgespeichert wird, werden auch die gerade aktuellen Positionsdaten und die Blickrichtung mitgespeichert.

Dieses Programmverhalten wird durch die Klassen *Gui* und *NMEAInterpreter* gesteuert, die im Folgenden kurz beschrieben werden.

Die Klasse *Gui* übernimmt die Darstellung und Steuerung des Benutzerinterfaces. Das Programm verfügt über drei Ansichten die in Abbildung 19 zu sehen sind.



Abbildung 19: Die drei Ansichten des GPSTools

Das Connection Fenster ermöglicht den Auf- bzw. Abbau der Verbindung zum GPS-Empfänger und zeigt die wichtigsten Positionsdaten auf dem Display an. Das Konsolenfenster zeigt zur Kontrolle die rohen NMEA-Sätze an. Das Take Picture Fenster wird auf jedem PDA anders dargestellt, da sich der Hersteller um die Kommunikation zwischen interner Kamera und dem .NET Framework kümmern muss. In allen Fällen wird aber eine Live-Ansicht der Kamera gezeigt und eine Möglichkeit zum Speichern des Bildes angeboten.

Um die Daten des GPS-Empfängers ständig updaten zu können, wird ein Timer verwendet, der sekundlich die Daten im Puffer des seriellen Ports ausliest und die Sätze zur Analyse an die Klasse *NmeaInterpreter* weiterreicht. Die *Gui*-Klasse wird

dann per Event über die Ankunft der neuen Daten benachrichtigt und aktualisiert die Ansicht.

Bei einem Klick auf den Button `Take Picture` wird ein `CameraCaptureDialogue` präsentiert, der es dem Anwender ermöglicht ein Bild aufzunehmen. Sobald er die Aufnahme mit OK bestätigt wurde, wird die Methode `writeGpx(String filename)` aufgerufen, die einen Wegpunkt nach GPX-Standard [29] erzeugt und in die Datei `[datum].gpx` schreibt.

Das GPS Exchange Format, kurz GPX, ist ein XML-basiertes Datenformat zur Speicherung von GPS-Daten. Der Standard wird durch ein Schema beschrieben, welches auf der Seite des Herstellers Topografix zu finden ist [29]. Innerhalb einer GPX Datei, können Wegpunkte und Routen, die ein GPS-Empfänger aufzeichnet, gespeichert und so auf einer standardisierten Basis mit anderen Programmen ausgetauscht werden. GPX wurde für diese Arbeit gewählt, weil es weit verbreitet ist und die Speicherung von Richtungsdaten unterstützt.

Durch die gleichzeitige Speicherung von Bild und Wegpunkt wird sichergestellt, dass die Zeitstempel im exif-Header des Bildes und in der GPX-Datei möglichst nahe bei einander liegen. Dies ist für die in späterer Folge angewandte Geo-Referenzierung sehr wichtig, denn die Zuordnung zwischen Wegpunkt und Bild erfolgt anhand der Uhrzeiten. Eine zusätzliche Schwierigkeit entsteht hier durch die Tatsache, dass es sich bei den zwei Uhrzeiten um Zeiten verschiedener Uhren handelt. Die Uhrzeit im exif-Header des Bildes stammt von der integrierten Uhr des PDAs, die Zeit des Wegpunktes ist die durch das GPS-System gemessene Uhrzeit. Wenn die beiden Uhren nicht synchronisiert wurden, muss die zeitliche Differenz bei der Georeferenzierung angegeben werden.

In der GPX-Datei selbst sind die Daten, die für die Referenzierung benötigt werden, abgespeichert. Abbildung 20 zeigt, welche Daten innerhalb der GPX-Datei abgespeichert werden.

```
<?xml version="1.0"?>
<gpx>
  <wpt lat="48.215017" lon="16.348055">
    <ele>157.5</ele>
    <time>2007-5-21T23:38:28Z</time>
    <magvar>143</magvar>
  </wpt>
  <wpt lat="48.2245245" lon="16.548632">
    <ele>160</ele>
    <time>2007-5-21T23:45:45Z</time>
    <magvar>220</magvar>
  </wpt>
</gpx>
```

Abbildung 20: GPX-Datei mit zwei Wegpunkten

Längen- und Breitengrad der Aufnahmeposition werden als Attribut zum Element Waypoint gespeichert. Dieses Element beinhaltet wiederum Kindelemente welche die Höhe, die GPS-Uhrzeit und die Ausrichtung des Empfängers in Grad angeben. Für jedes gespeicherte Bild wird ein eigenes Wegpunkt Element erzeugt, in der GPX-Datei können beliebig viele Wegpunkte abgelegt werden. Alle gespeicherten Dateien werden standardmäßig im Ordner „Eigene Dokumente/gpsPictures/[Datum]“ abgelegt, von wo aus sie zur weiteren Bearbeitung auf den PC übertragen werden können.

Die Klasse *NMEAInterpreter* dient zur Interpretation der NMEA-Sätze, die ein GPS-Empfänger standardmäßig über seine serielle Schnittstelle ausgeben kann, siehe Kapitel 3.1.3. Die Grundfunktionalität für einen solchen Interpreter stellt Jon Person unter [16] vor. Um den Anwendungsfällen dieser Arbeit gerecht zu werden, wurde seine Version um die Auswertung von Richtungs- und Gütedaten erweitert.

Herzstück der Klasse ist die Funktion `parse(String sentence)`, die einen NMEA-Satz anhand seiner Kennung identifiziert und anschließend in seine Einzelteile zerlegt. Die Funktion analysiert derzeit folgende Sätze die in Tabelle 6 beschrieben sind.

Satztyp	Werte des Satzes	Sendegerät
\$GPRMC (Recommended Minimum Sentence)	Uhrzeit, Datum, Position	GPS-Empfänger
\$GPGSV (Satellites in View Sentence)	sichtbare Satelliten	GPS-Empfänger
\$GPGSA (Fixed Satellites and DOP Sentence)	Art des Fixes (keiner, 2D, 3D), DOP (Güte des Signales)	GPS-Empfänger
\$GPGGA (Global Pos. Fixed Data Sentence)	Anzahl der sichtbaren Satelliten, Höhe	GPS-Empfänger
\$HCHDG (Heading Sentence)	Ausrichtung des Empfängers in Grad	Kompass

Tabelle 6: Übersicht welche Daten sich aus welchem NMEA Satz extrahieren lassen

Sobald ein Satz ausgewertet wurde, wird die Klasse *Gui* mittels eines Events über die neuen Daten informiert.

Information zur Ausrichtung des GPS-Empfängers werden ebenfalls über das NMEA-Protokoll übertragen und ausgelesen. Der Empfänger verwendet zur Übertragung den Satz \$HCHDG des integrierten digitalen Kompasses, der die magnetische Richtung als Dezimalzahl in Grad enthält.

Neben `parse(String sentence)` enthält die Klasse noch eine weitere wichtige Funktion: `isValid(String sentence)`. Diese Prozedur überprüft, ob die Prüfsumme die jeder NMEA-Satz am Ende trägt auch der Prüfsumme der übertragenen Daten entspricht. Bei einer Abweichung wird der Satz sofort verworfen und gar nicht weiter analysiert.

5.3.2 Gewonnene Erkenntnisse

Während der Arbeit an diesem Programm, wurden einige Erkenntnisse gewonnen, die auch für weitere Entwicklungen in diesem Bereich von Interessen sein könnten. Der folgende Abschnitt gibt daher einen Einblick in die Probleme die sich während der Programmierung ergeben haben und zeigt, wenn möglich, wie diese gelöst wurden.

Leider bietet das Microsoft Compact Framework 2.0 keine direkte Möglichkeit zum Schreiben von exif-Daten. Diese Funktion wird nur von dem umfangreicheren .NET Framework 2.0, das jedoch nur für den PC und nicht für PDAs zur Verfügung gestellt. Aus diesem Grund war es nicht möglich, die gewonnenen Positionsdaten direkt im Header des Bildes zu speichern und es musste eine Alternative gefunden werden. Als Lösung für das Problem bot sich der Umweg über die zusätzliche GPX-Datei an, in der zu jedem Bild ein Wegpunkt mit Zeitstempel gespeichert wird, über den sich die Daten später geo-referenzieren lassen. Die Entscheidung für GPX gab einerseits die hohe Verbreitung des Standards sowie die Tatsache, dass GPX auf XML basiert und dadurch mit Standardbibliotheken bearbeitet werden kann.

Das Compact Framework bietet zum Lesen der Daten, die über den seriellen Port geschickt werden, die einfache Methode `serialPort.readLine()` und das Event `serialPort.dataReceived()` an. Erstere Funktion besitzt den großen Vorteil, dass sie immer einen ganzen NMEA-Satz aus dem Puffer liest und diesen anschließend an den `NmeaInterpreter` weitergeben kann. Sie erzeugt dafür für jeden Lesevorgang einen eigenen Thread, liest und verarbeitet die Daten und beendet dann den Thread. Leider entsteht durch diese vielen Threads ein enormer Overhead an Verwaltungsdaten, der die Applikation auszubremsen beginnt. Die Anwendung kann die Daten nicht mehr schnell genug aus dem Puffer lesen und dieser beginnt sich langsam aber stetig zu füllen. Dies führt schlussendlich dazu, dass das Programm die aktuellen Positionswerte nicht mehr rechtzeitig anzeigen kann. In der Praxis, ging die auf dem Display angezeigte GPS-Uhrzeit bereits nach einer Minute um 15 Sekunden nach. Als Lösung bot sich die Verwendung eines Timers, an der nur jede Sekunde die Daten mittels `serialPort.readExiting()` aus dem Puffer liest und in einem einzigen Thread verarbeitet. Nach dieser Umstellung funktioniert das Programm nun in Echtzeit.

Beim Schließen des seriellen Portes über den Stop-Button kam es vereinzelt zu einem Deadlock innerhalb der Anwendung. Immer dann, wenn der Anwender versuchte den Port zu schließen, als dieser noch am Lesen war, reagierte das Programm nicht mehr. Um dieses Problem zu lösen, wurde die Funktion `closeSerialOnExit()` implementiert, die in einem eigenen Thread wartet, bis der Lesevorgang beendet ist und dann den Port schließt.

Wie bereits erwähnt, hängt die Implementierung der Kamera-Schnittstelle für das Compact Framework vom Hersteller des Gerätes ab. Da neben dem iPAQ auch ein MDA²⁹ von T-Mobile zur Verfügung stand, konnte das Program auf beiden PDAs getestet und folgendes Ergebnis festgestellt werden: Der iPAQ lässt innerhalb von .NET nur Fotos mit einer Auflösung von max. 320x240 Pixeln zu und schreibt beim Speichern des Fotos keinerlei exif-Daten in das Bild. Und das, obwohl die HP eigene Fotosoftware Bilder bis 1280x1024 aufnimmt und sogar GPS-Daten in den exif-Header schreiben kann. Für das HP-Camera SDK, das diese Funktionalität in C# nachrüsten könnte, verlangt der Hersteller 999.- US-Dollar. Beim MDA treten hingegen keine Probleme auf. Aus .NET heraus öffnet sich der gleiche Kamera-Dialog, der auch von T-Mobile vorinstalliert wurde. Damit sind Bilder bis 640x480 Pixel möglich, inklusive einem gültigen exif-Header.

Das größte Problem beim Geo-Referenzieren ist die Synchronisation der Zeitstempel im exif-Header des Bildes mit denen des Wegpunktes in der GPX-Datei. Bei Abweichungen von mehreren Sekunden, ist nicht mehr sichergestellt, dass einem Bild die richtigen Koordinaten zugewiesen werden. Eine Verbesserungsmöglichkeit im Programm selbst, bestünde jedoch in der automatischen Synchronisation zwischen der Uhrzeit des jeweiligen Gerätes mit der gemessenen GPS-Zeit. So kann gewährleistet werden, dass es bei der Geo-Referenzierung zu keiner Differenz kommt. Eine andere Möglichkeit wäre, die Abweichung zwischen Geräteuhrzeit und GPS-Zeit festzustellen und mit dem Wegpunkt selbst zu speichern. Das hat aber den Nachteil, dass die GPX-Datei nicht mehr dem GPX-Schema entspricht und die Weiterverarbeitung der Daten erschwert wird. Da zu Beginn der Messungen die Uhrzeit des PDAs mit der GPS-Zeit synchronisiert wurde, waren die gemessenen Zeiten innerhalb dieser Arbeit maximal um drei Sekunden verschoben. Die anschließende Geo-Referenzierung hat mit diesen Werten immer funktioniert. Eine Implementierung der hier genannten Verbesserungsvorschläge war daher nicht notwendig.

²⁹<http://www.t-mobile.com/shop/phones/Detail.aspx?device=8802ddeb-1ee4-477a-9608-d9cd1e2a903f>

5.4 Geo-Referenzierung

Da es trotz des GPSTools nicht gelang, die geografischen Metadaten direkt im exif-Header zu speichern, muss nun eine Geo-Referenzierung stattfinden, die die Fotos mit den Wegpunkten aus der GPX-Datei verknüpft.

Im Internet findet man viele verschiedene Tools, die das nachträgliche Geo-Referenzieren von Fotos erlauben [24]. Aus Gründen der einfachen Erweiterbarkeit, kam im Zuge dieser Arbeit das Perl-Skript `gpsPhoto.pl`³⁰ von `carto.net` zum Einsatz. Das Skript ermöglicht den einfachen Abgleich von GPX-Wegpunkten bzw. Tracks mit einem Verzeichnis von Bildern. Zum Speichern der Positionsdaten im exif-Header kommt die ebenfalls frei verfügbare Perl Bibliothek `exifTool` von Phil Harway³¹ zum Einsatz.

Das Skript selbst bietet zurzeit keine grafische Oberfläche und muss daher von der Kommandozeile gestartet werden. Dabei können dem Programm verschiedene Parameter übergeben werden, die wichtigsten sind in Abbildung 21 kurz beschrieben.

```
--dir          Das Verzeichnis in dem sich die unreferenzierten
               Fotos befinden
--gpsFile      Die GPX-Datei mit den Wegpunkten
--maxTimeDiff  Die maximale zeitliche Abweichung der beiden
               Zeitstempel in Sekunden
--timeOffset   Die Differenz zwischen Kamerauhr und GMT (GPS-Zeit)
```

Abbildung 21: Wichtige Parameter von `gpsPhoto.pl`

Bei einem gültigen Aufruf, beginnt das Skript die Wegpunkte anhand ihres Zeitstempels den Bildern zuzuordnen. Dabei wird der Punkt verwendet, der die geringste zeitliche Abweichung zur Aufnahmezeit des Bildes hat und dessen Abweichung kleiner als der Wert von `maxTimeDiff` ist. Sollte kein entsprechender Wegpunkt gefunden werden, teilt das Programm dies dem Benutzer mit. Das Skript ist auch in der Lage, die Aufnahmeposition zwischen zwei Wegpunkten zu interpolieren. Dies ist innerhalb dieser Arbeit jedoch nicht notwendig, da das

³⁰ <http://www.carto.net/projects/photoTools/gpsPhoto/>

³¹ <http://www.sno.phy.queensu.ca/~phil/exiftool/>

GPSTool den Zeitpunkt des entsprechenden Wegpunktes gleichzeitig mit der Aufnahme des Bildes erstellt.

5.4.1 Modifizierungen

Um das Skript für diese Arbeit nutzen zu können, musste es an zwei Stellen modifiziert werden. In der Originalversion erlaubt das Skript keinerlei Verarbeitung der Aufnahmerichtung. Aufgrund der guten Dokumentation konnte jedoch eine entsprechende Funktion implementiert werden. Die Aufnahmerichtung wird nun ebenfalls aus der GPX-Datei ausgelesen und im dafür vorgesehenen exif-Feld `GPSTimeDirection` in Grad abgespeichert.

Da die Bilder des HP iPAQs in Kombination mit dem GPSTool keinen exif-Header aufweisen, musste außerdem eine Möglichkeit geschaffen werden, einen anderen Zeitstempel als das `DateTimeOriginal` Feld des exif-Headers zur Geo-Referenzierung verwenden zu können. Als Lösung bot sich hier die Verwendung des `LastModified` Zeitstempels der Datei selbst an, der bei der ersten Speicherung des Bildes vom Dateisystem selbst erzeugt wird. Die Geo-Referenzierung funktioniert mit dieser neuen Methode genau so zuverlässig wie vor der Modifikation.

Die Modifizierte Version des Skriptes wird im weiteren Verlauf der Arbeit mit dem Namen `gpsPhotoWithBearing.pl` bezeichnet

5.5 Datenspeicherung

Der folgende Abschnitt behandelt die serverseitige Speicherung der Testdaten. Darunter fallen alle Bilder die mittels FoVA auf den Server geladen werden und die Informationen über einige Sehenswürdigkeiten nach denen mit FoVA gesucht werden kann.

5.5.1 MetadatenSpeicherung von Bildern

Obwohl die geografischen Informationen bereits im exif-Header der Fotos gespeichert wurden, ist es aus Performancegründen sinnvoll, sie zusätzlich in einer relationalen Datenbank abzulegen.

Das Schema der dafür verwendeten Tabelle *image_data* ist in Tabelle 7 zu sehen.

Feld	Datentyp	Beschreibung	Datenherkunft
<i>id</i>	Integer, primary key	Eindeutige Identifikationsnummer	Datenbank
<i>name</i>	String	Dateiname des Bildes	Dateiname
<i>lat</i>	Double	Breitengrad des Aufnahmeortes	Bildmetadatum
<i>long</i>	Double	Längengrad des Aufnahmeortes	Bildmetadatum
<i>alt</i>	Double	Höhe des Aufnahmeortes	Bildmetadatum
<i>bearing</i>	Integer	Ausrichtung der Kamera in Grad	Bildmetadatum
<i>datetime</i>	DateTime	Zeitpunkt der Aufnahme	Bildmetadatum
<i>angle</i>	Integer	Angegebener Feldwinkel	Benutzereingabe
<i>distance</i>	Double	Angegebene Distanz	Benutzereingabe
<i>destPointLeftLat</i>	Double	Breitengrad linker Zielpunkt	Berechnet
<i>destPointLeftLong</i>	Double	Längengrad linker Zielpunkt	Berechnet
<i>destPointRightLat</i>	Double	Breitengrad rechter Zielpunkt	Berechnet
<i>destPointRightLong</i>	Double	Längengrad rechter Zielpunkt	Berechnet

Tabelle 7: Schema der Tabelle *image_data*

Die Tabelle speichert neben dem Dateinamen des Bildes, auch dessen genaue Aufnahmeposition als Längengrad, Breitengrad und Höhe, dessen Aufnahmerichtung und dessen Aufnahmezeitpunkt. Zusätzlich werden der angegebene Feldwinkel und die Distanz des Sichtfeldes gespeichert. Damit die beiden Zielpunkte des Dreiecks nicht jedes Mal neu berechnet werden müssen, werden auch deren Koordinaten in der Tabelle abgelegt. Abbildung 22 zeigt einen Auszug mit Testdaten der Tabelle *image_data*.

id	name	lat	long	alt	bearing	datetime	angle	distance	destPointLeftLat	destPointLeftLong	destPointRightLat	destPointRightLong
1	HPIM0026.jpg	48.19...	16.37...	188	152	2007-04-24 09:06:30	60	300	48.19764167...	16.374281663494	48.1963744065	16.370705889074
2	HPIM0030.jpg	48.19...	16.37...	188	120	2007-04-24 09:08:11	60	300	48.19874437...	16.374536017255	48.1964066256...	16.372510971786
3	HPIM0033.jpg	48.21...	16.36...	190	300	2007-04-24 09:59:20	60	300	48.21428326...	16.35728219828	48.2166211163...	16.359307673317
4	IMAGE_00006.jpg	48.21...	16.35...	192	1	2007-06-03 20:45:12	60	300	48.21764898...	16.35235584164	48.21760187	16.356406624042
5	IMAGE_00008.jpg	48.21...	16.35...	190	90	2007-06-03 20:47:20	60	350	48.21715161...	16.359280346161	48.2140022423...	16.359280094384
6	IMAGE_00013.jpg	48.21...	16.36...	198	295	2007-06-03 20:56:38	60	300	48.21435665...	16.356752274953	48.21680324483	16.358464250026

Abbildung 22: Auszug aus der Tabelle *image_data*

5.5.2 MetadatenSpeicherung von Sehenswürdigkeiten

Damit in späterer Folge festgestellt werden kann, ob und wenn ja welche Sehenswürdigkeiten auf einem Bild dargestellt sind, ist es notwendig ihre Namen und Positionen in der realen Welt zu kennen. Im Zuge dieser Arbeit wurden zwei Methoden implementiert die einerseits eine eigene Datenbasis und andererseits den Datenbestand der Wikipedia nutzen um die Koordinaten von bestimmten Sehenswürdigkeiten zu ermitteln.

Wenn man bereits einen eigenen Datenbestand zur Verfügung hat, bzw. der Umfang der zu speichernden Sehenswürdigkeiten gering ist, ist es sinnvoll eine eigene Datenbasis zu erstellen, die den Namen der Sehenswürdigkeit, ihre Position und ev. weitere Felder wie Beschreibung und URL enthält. Die Verwendung einer eigenen Datenbank hat den Vorteil, den Datenbestand selbst verwalten und auch offline nutzen zu können. Im Vergleich zu einem reinen Webservice sind solche Lösungen meist performanter. FoVA nutzt dafür eine Tabelle, die nach dem in Tabelle 8 dargestellten Schema erzeugt wurde.

Feld	Datentyp	Beschreibung
<i>id</i>	Integer, primary key	Eindeutige Identifikationsnummer
<i>name</i>	String	Name der Sehenswürdigkeit
<i>lat</i>	Double	Breitengrad des Standortes
<i>long</i>	Double	Längengrad des Standortes
<i>url</i>	String	Link um dem Anwender mehr Informationen zur Sehenswürdigkeit bieten zu können
<i>thumbnaiUrl</i>	String	URL eines Vorschaubildes der Sehenswürdigkeit

Tabelle 8: Schema der Tabelle *sights_data*

Zusätzlich zu den Namen und Koordinaten der Sehenswürdigkeit, enthält die Tabelle Felder für einen Link und ein Vorschaubild um den Anwender mit zusätzlichen Informationen zur Sehenswürdigkeit versorgen zu können. Abbildung 23 zeigt einen Auszug mit Testdaten der Tabelle *sights_data*.

id	name	lat	long	url	thumbnailUrl
1	Karlskirche	48.1982	16.3714	http://de.wikipedia.org/wiki/Wiener_Karlskirche	http://www.geonames.org/img/wikipedia/1000/thumb-647-100.jpg
2	Votivkirche	48.2154	16.3586	http://de.wikipedia.org/wiki/Votivkirche_Wien	http://www.geonames.org/img/wikipedia/52000/thumb-51808-100.jpg

Abbildung 23: Auszug aus der Tabelle *sights_data*

Neben dem Sammeln und Bereitstellen unzähliger Beiträge, betreibt die Wikipedia ein eigenes Projekt zur Geo-Referenzierung³². Darin erlaubt sie, zu allen Artikeln geografische Koordinaten nach einem festgelegten Schema zu speichern. Durch die große Nutzergemeinde gibt es bereits über 580.000 Artikel in mehr als 17 Sprachen, die eine geografische Referenz aufweisen [30].

Aufgrund des großen Datenbestandes der Wikipedia und der Tatsache, dass ihr gesamter Inhalt frei verfügbar unter der GNU Free Documentation License³³ abrufbar ist, wurde nach einer Möglichkeit gesucht, die geo-referenzierten Artikel der Wikipedia mit FoVA zu Verknüpfen. Die Wikipedia bietet zu diesem Zweck das Wikipedia Geocoding Webservice³⁴ an, das mittels eines standardisierten HTTP-Requests einen komfortablen Zugriff auf die Daten der Wikipedia bietet. Angeboten werden Reverse Geocoding, Volltextsuche und eine Rechteck-Suche:

Die *Suche nach nahegelegenen Wikipedia Artikeln* (Reverse Geocoding) ermöglicht die Suche nach Artikeln ausgehend von einer geografischen Position. Diese kann entweder direkt als Lat/Long Koordinatenpaar oder aber auch indirekt über einen anderen geo-referenzierten Artikel, beispielsweise über den Standort eines berühmten Gebäudes, angegeben werden. Zusätzlich muss ein Suchradius definiert werden, der die geografischen Grenzen für die Suche festlegt. Zurückgeliefert werden alle Artikel, die innerhalb des gewünschten Umkreises liegen. Abbildung 24 zeigt eine Beispielsanfrage an das Webservice.

`http://ws.geonames.org/findNearbyWikipedia?lat=48.1982&lng=16.3714&radius=0.01`

Abbildung 24: Reverse Geocoding Anfrage an das Wikipedia Webservice

³² siehe: <http://de.wikipedia.org/wiki/WP:GEO>

³³ <http://www.gnu.org/licenses/fdl.txt>

³⁴ siehe: <http://www.geonames.org/export/wikipedia-webservice.html>

Diese Anfrage liefert alle Wikipedia Artikel zurück, die sich innerhalb eines 10 Meter Radius um den Punkt 48.2 geografischer Länge und 16.4 geografischer Breite.

Die *Wikipedia Volltextsuche* ermöglicht die Suche nach einem bestimmten Suchbegriff. Zurückgeliefert werden alle Artikel, die diesen Begriff beinhalten. Abbildung 25 zeigt eine Beispielanfrage.

`http://ws.geonames.org/wikipediaSearch?q=Stephansdom&maxRows=5`

Abbildung 25: Volltextanfrage an das Wikipedia Webservice

Diese Anfrage liefert alle georeferenzierten Artikel zurück, die das Wort „Stephansdom“ beinhalten. Durch die Angabe von `maxRows=5` wird das Suchergebnis auf die ersten fünf Treffer beschränkt.

Die *Wikipedia Rechteck-Suche* ermöglicht die Suche nach Artikeln, die sich innerhalb eines definierten rechteckigen Bereichs befinden. Als Suchparameter genügen die Koordinaten für alle vier Eckpunkte des Rechtecks. Abbildung 26 zeigt eine Beispielanfrage.

`http://ws.geonames.org/wikipediaBoundingBox?north=44.1&south=-9.9&east=-22.4&west=55.2`

Abbildung 26: Rechteckanfrage an das Wikipedia Webservice

Diese Anfrage liefert alle georeferenzierten Artikel zurück, die sich innerhalb des Koordinaten Rechtecks 44.1 nördlicher und -9.9 südlicher Breite und 22.4 östlicher und 55.2 Grad westlicher Länge liegen.

Alle drei Services bieten zusätzlich noch einige Suchparameter, wie Sprache oder maximale gefundene Artikel an, mit der die Suche weiter spezifiziert und eingeschränkt werden kann. Die Resultate selbst werden je nach Aufruf als XML³⁵ oder JSON³⁶ Dokument zurückgeliefert, zwei Beispieldokumente dazu finden sich im Anhang II. Da für beide Formate ein Schema definiert wurde, ist es möglich, Inhalte der gefundenen Artikel automatisch weiterzuverarbeiten. Innerhalb dieser Arbeit wurde ausschließlich das XML-Webservice genutzt, da zum Umgang mit XML bereits ein breites Vorwissen existierte.

³⁵ <http://www.w3.org/XML/>

³⁶ <http://json.org/json-de.html>

Kapitel 6

Zusammenfassung

Von Beginn an war es das Ziel dieser Arbeit, die Suche nach digitalen Erinnerungen zu vereinfachen bzw. zu verbessern. Durch die Berücksichtigung von zusätzlichen geografischen Informationen wie dem Aufnahmeort, der Aufnahmerichtung und des Sichtfeld eines Fotos, sollten dem Benutzer neue Möglichkeiten zum schnelleren und präziseren Auffinden bestimmter Bilder in die Hand gegeben werden.

Um dem Leser ein besseres Verständnis für die Begriffe Metadaten bzw. geografische Metadaten zu geben, wurden diese zu Beginn genau definiert. Anschließend wurden einige wichtige Standards zur Speicherung von Bildmetadaten angeführt und diskutiert. Hier zeigte sich, dass die viel genutzten Metadatenstandards der digitalen Fotografie, allen voran exif, bereits eine standardisierte Möglichkeit für die Speicherung von geografischen Informationen bereitstellen.

Die folgenden Kapitel beschäftigten sich mit der Gewinnung bzw. Erzeugung von geografischen Daten und den dafür notwendigen technischen Voraussetzungen. Zur Bestimmung von Aufnahmeort und Blickrichtung eines Fotos wurden an dieser Stelle insbesondere die Vorzüge und Einschränkungen des GPS-Systems und der Einsatz von elektromagnetischen Kompassen diskutiert. Im Anschluss wurde außerdem gezeigt, wie sich die gewonnenen Daten nutzen lassen, um das zweidimensionale Sichtfeld einer Kamera errechnen zu können.

Am Ende der Arbeit wurden die gewonnenen Kenntnisse in die Praxis umgesetzt und untersucht, inwieweit geografische Informationen das Auffinden von bestimmten Fotos verbessern können. Zu diesem Zweck wurden die Anwendungen FoVA und

GPSTool entwickelt, die den gesamten Prozess, also Gewinnung, Speicherung und Auswertung von geografischen Metadaten abdecken und zwei definierte Anwendungsfälle auf anschauliche Weise realisieren.

Besonders die Verwendung eines zweidimensionalen Sichtfelds zur Verbesserung des Suchprozesses stellte sich in der Praxis als empfehlenswert heraus. Ziel war es, die Berechnung des Sichtfeldes möglichst automatisch durchführen zu können, doch eine vollständige Automatisierung scheiterte leider trotz der eigentlich verfügbaren Technik und Standards. Wie in Kapitel 2.3.1 erwähnt, stellt die exif-Spezifikation zwar Felder für alle notwendigen Parameter des Sichtfeldes zur Verfügung, doch die Angaben zur Bildsensorgröße und die vom Autofokus gemessene Distanz zum Objekt fehlen mangels Unterstützung von manchen Kameraherstellern bei vielen Bildern. Aber auch wenn diese Werte vorhanden sein sollten, erweist sich besonders die Distanz als wenig praxistauglich. Bei einer hohen Tiefenschärfe wird sie, obwohl sich die Sehenswürdigkeit vielleicht nur wenige Metern vor einem befindet, oft viel zu hoch angegeben. Das dazugehörige Sichtfeld wäre dann viel zu lang und könnte fälschlicherweise auch Sehenswürdigkeiten miteinbeziehen, die hinter dem fotografierten Objekt liegen.

Trotzdem wird durch die Einbeziehung der Aufnahme­richtung und des Sichtfeldes ein wichtiges Filterkriterium geschaffen, das die Resultate der Suche drastisch verbessern kann. Dies wurde durch einen Vergleich zwischen einer reinen Umkreissuche und einer Umkreissuche mit Berücksichtigung des Sichtfeldes festgestellt. Durch das Sichtfeld können die Treffer des Suchergebnisses weiter eingeschränkt werden, was zu einer höheren Präzision der Suche führt. Weiters konnte festgestellt werden, dass die Vollständigkeit der Suche durch das Sichtfeld weder verbessert noch verschlechtert wird. Der Einsatz des Sichtfeldes stellt daher eine deutliche Verbesserung in der ortsbasierten Suche da. Es zeigte sich aber auch, dass das Verbesserungspotential maßgeblich von der Qualität bzw. Genauigkeit der geografischen Metadaten, des Feldwinkels und der Distanz zum fotografierten Objekt abhängt.

Ein weiterer Teil dieser Arbeit beschäftigte sich mit den Möglichkeiten der Geo-Referenzierung von Fotos. Dabei zeigte sich, dass die direkte Verbindung von Kamera und GPS-Empfänger die bequemste und genaueste Art der Referenzierung

ist. Leider verhindern derzeit der hohe Stromverbrauch von GPS-Empfängern und die teilweise langen Einschaltzeiten eine hohe Verbreitung dieser Methode. Die semiautomatische Referenzierung bringt, entsprechend niedrige Messintervalle vorausgesetzt, die gleichen Ergebnisse, erfordert dazu aber immer einen zusätzlichen Arbeitsschritt. Die im Zuge dieser Arbeit entwickelten bzw. modifizierten Programme *GPSTool* und *gpsPhotoWithBearing.pl* zeigen jedoch auch, dass dies für den Anwender nicht zwangsläufig mit einem höheren Aufwand verbunden sein muss, denn der Abgleich der Zeitstempel zwischen Wegpunkt und Bild funktioniert dank *gpsPhotoWithBearing.pl* vollkommen automatisch.

Kapitel 7

Ausblick

Obwohl Techniken zur Positions- und Richtungsbestimmung schon lange existieren, werden diese geografischen Zusatzinformationen erst seit kurzer Zeit zur Organisation von digitalen Erinnerungen genutzt. Die Nutzer schätzen die Vorteile, die für sie durch Verwendung von geografischen Informationen entstehen, allen voran die Gruppierung der Bilder nach ihrem Aufnahmeort. Doch die Methoden zur Auswertung der geografischen Metadaten sind bei weitem noch nicht ausgereizt. Gerade in der Analyse des Sichtfeldes eines Fotos steckt noch viel Potential für zukünftige Untersuchungen und Entwicklungen. Durch die Verwendung eines separaten Distanzmessers könnte beispielsweise die Größe des Sichtfeldes viel genauer und vor allem automatisch berechnet werden. Gleichzeitig wäre damit das Problem der Sichtbarkeit von verdeckten Objekten zumindest teilweise gelöst. Der nächste logische Schritt bestünde darin, die zweidimensionale Ebene zu verlassen und durch die Verwendung von Neigungssensoren die Lage und Ausrichtung der Kamera im Raum feststellen zu können.

Den ganz großen Durchbruch verhindern aber derzeit noch die Techniken zur Positionsbestimmung. Das GPS-System funktioniert prinzipbedingt nicht in Innenräumen und andere Methoden, die dafür das GSM- oder WiFi-Netz nutzen, arbeiten meist zu ungenau. An diesem grundlegenden Problem wird auch das neue GALILEO Positionierungssystem nichts ändern können. Durch neue, multistandardkompatible Empfänger wäre es jedoch denkbar, dass Satelliten verschiedener Anbieter gleichzeitig genutzt werden könnten um die Positionierung,

besonders in Gebieten mit schlechten Empfangseigenschaften, zu verbessern. Eine interessante Erweiterung bieten auch die derzeit entstehenden Assisted-GPS Systeme. Diese haben das Ziel, GPS-Empfänger mit den Lokalisierungsmethoden des GSM-Netzes zu koppeln um eine schnellere Positionsbestimmung zu ermöglichen.

Auch in den derzeitigen Suchmethoden, die geografische Informationen als Hilfsmittel zu einer treffsichereren Suche nutzen, steckt noch Verbesserungspotential. So wäre es denkbar, dass die Sortierung von Suchergebnissen abhängig vom Umfeld gemacht wird. Aufgrund der Aufnahmeposition könnte dabei festgestellt werden ob das Foto in städtischem oder ländlichem Gebiet gemacht wurde. Aufgrund kurzer Fotodistanzen könnte im ersten Fall davon ausgegangen werden, dass der Fokus auf den nähergelegenen Objekten liegt und diese somit an erster Stelle gereiht werden. Bei Fotos die während einer Gebirgswanderung aufgenommen wurden, sind jedoch eher die umliegenden Berge, die in weiter Entfernung liegen, von Interesse. In diesem Fall könnten diese an den Anfang des Suchergebnisses gestellt werden.

Abschließend kann jedoch festgehalten werden, dass der Weg, geografische Informationen als Hilfsmittel für den Suchprozess einzusetzen, ein Vielversprechender ist. Obwohl viele Techniken und Methoden noch in ihren Kinderschuhen stecken, sieht man anhand von Applikation wie FoVA wohin die Richtung gehen kann und wie geografische Metadaten das Suchergebnis verbessern können. Hier zeigt sich aber auch, dass das Verbesserungspotential maßgeblich von der Genauigkeit der verwendeten Positionierungsmethoden abhängt. Erst wenn die Bestimmung der notwendigen geografischen Informationen durch neue Techniken genau, zuverlässig und vor allem gänzlich unabhängig vom Standort funktioniert, werden die in dieser Arbeit behandelten Konzepte und Anwendungen von ihrem Nischendasein zur Massentauglichkeit aufsteigen können.

Abbildungsverzeichnis

Abbildung 1: Grafische Darstellung der Anforderungen an das Fotomotiv, die Positionsbestimmung und die Richtungsmessung	10
Abbildung 2: Auswirkung der Anzahl und der Verteilung der sichtbaren Satelliten auf die Güte der ermittelten Position	24
Abbildung 3: Reflexionsproblematik eines GPS-Signals	26
Abbildung 4: Typische NMEA-0183 Satzfolge.....	28
Abbildung 5: Interpretation eines \$GPRMC Satzes	28
Abbildung 6: Jelbert Geo-Tagger	36
Abbildung 7: Sichtfenster das durch horizontalen, vertikalen und diagonalen Sichtwinkel aufgespannt wird.....	39
Abbildung 8: Überführung eines 3-dimensionalen Sichtfeldes in den 2-dimensionalen Raum.....	41
Abbildung 9: Berechnung des Sichtfeldes	42
Abbildung 10: Auswirkung verschiedener Feldwinkel auf die Größe des Sichtfeldes	47
Abbildung 11: Kartenausschnitt mit vier Bildern	47
Abbildung 12: Suche nach Fotos von der Wiener Karlskirche	48
Abbildung 13: Erfolgreiche Identifizierung einer Sehenswürdigkeit die auf dem Foto zu sehen ist ..	49
Abbildung 14: Drei möglichen Lagen eines Punktes zu einem Dreieck.....	50
Abbildung 15: Drei möglichen Lagen eines Punktes zu einer Geraden im 2-dimensionalen Raum..	51
Abbildung 16: Probleme für FoVA die durch eine zu kurze Entfernungsangabe entstehen können ..	51
Abbildung 17: Gebäude der österreichischen Nationalbank	52
Abbildung 18: Vergleich zwischen einer Suche ohne und einer Suche mit Berücksichtigung des Sichtfeldes.....	53
Abbildung 19: Die drei Ansichten des GPSTools.....	57
Abbildung 20: GPX-Datei mit zwei Wegpunkten	59
Abbildung 21: Wichtige Parameter von gpsPhoto.pl.....	63
Abbildung 22: Auszug aus der Tabelle image_data.....	65
Abbildung 23: Auszug aus der Tabelle sights_data	67
Abbildung 24: Reverse Geocoding Anfrage an das Wikipedia Webservice.....	67
Abbildung 25: Volltextanfrage an das Wikipedia Webservice	68
Abbildung 26: Rechteckanfrage an das Wikipedia Webservice	68

Tabellenverzeichnis

Tabelle 1: Typische Metadaten für Bilder und Fotos	14
Tabelle 2: Vergleich von IIM/IPTC, Exif, XMP und MPEG-7.....	19
Tabelle 3: Auswirkung der behandelten Fehlern auf die Güte der GPS-Position.....	27
Tabelle 4: Gegenüberstellung der genannten Verfahren zur Gewinnung von geografischen Metadaten	37
Tabelle 5: Veränderung des Blickwinkels abhängig von der Brennweite (Kleinbildformat).....	40
Tabelle 6: Übersicht welche Daten sich aus welchem NMEA Satz extrahieren lassen.....	60
Tabelle 7: Schema der Tabelle <i>image_data</i>	65
Tabelle 8: Schema der Tabelle <i>sights_data</i>	66

Literaturverzeichnis

- [1] **Henry Lieberman and Hugo Liu** (2002). Adaptive Linking between Text and Photos Using Common Sense Reasoning. In De Bra, Brusilovsky, Conejo (Eds.): Adaptive Hypermedia and Adaptive Web-Based Systems, Second International Conference, AH 2002, Malaga, Spain, May 29-31, 2002, Proceedings. Lecture Notes in Computer Science 2347 Springer 2002, ISBN 3-540-43737-1, pp. 2-11. DOI: <http://web.media.mit.edu/~hugo/publications/papers/AH2002-aria.pdf>
- [2] **National Information Standards Organization**. Understanding Metadata. MD 20814 USA. NISO Press, 2004. DOI: <http://www.niso.org/standards/resources/UnderstandingMetadata.pdf>
- [3] **Burners-Lee, Tim** (1997). Web Architecture Metadata. Zugriff: 15. Juni 2007. <http://www.w3.org/DesignIssues/Metadata.html>
- [4] **IPT Council**. IPTC-Web. Zugriff: 17. April 2007 <http://www.iptc.org/IIM/>
- [5] **Wikipedia**. IPTC. Zugriff: 17. April 2007. <http://de.wikipedia.org/w/index.php?title=IPTC-NAA-Standard&oldid=29867771>
- [6] **Technical Standardization Committee on AV & IT Storage Systems and Equipment**. EXIF.org - EXIF and related resources. JEITA, 2002. <http://www.exif.org/Exif2-2.PDF>
- [7] **Wikipedia**. XMP. Zugriff: 17. April 2007 http://en.wikipedia.org/w/index.php?title=Extensible_Metadata_Platform&oldid=131214758
- [8] **Adobe Inc**. XMP Specification. September 2005. http://www.adobe.com/devnet/xmp/pdfs/xmp_specification.pdf
- [9] **Manjunath, B. S., Salembier, Philippe, Sikora Thomas**. Introduction to MPEG-7, Multimedia Content Description Interface. West Sussex: John Wiley & Sons Ltd., 2002
Yilmaz, Erkan, Dürschlag, Michael und Dabas, Ghasan. MPEG-7 Multimedia Content Description Interface. 2000. Zugriff: 26. April 2007. <http://www.aifb.uni-karlsruhe.de/WBS/sst/Teaching/SemanticWebWS00-01/MPEG7.pdf>
- [10] **Toyama, K., Logan, R., and Roseway, A**. 2003. Geographic location tags on digital images. In *Proceedings of the Eleventh ACM international Conference on Multimedia* (Berkeley, CA, USA, November 02 - 08, 2003). MULTIMEDIA '03. ACM Press, New York, NY, 156-166. DOI=<http://doi.acm.org/10.1145/957013.957046>
- [11] **Rötzer, Florian**. Telopolis - Konkurrenz für GPS und Galileo (14.11.2006). Hannover: Heise Zeitschriften Verlag, 2006. Zugriff: 08. August 2007. <http://www.heise.de/tp/r4/artikel/23/23971/1.html>
- [12] **Wikipedia**, COMPASS (Satellitennavigation). Zugriff: 08. August 2006.] http://de.wikipedia.org/w/index.php?title=Compass_%28Satellitennavigation%29&oldid=32102950
- [13] **Bauer, Manfred**. Vermessung und Ortung mit Satelliten. Heidelberg: Wichmann, 1997.
- [14] **Görrisch, Dieter**. GPS im Selbstbau. Pöing: Franzis Verlag, 2003.
- [15] **Kowoma.de**. Der Aufbau des GPS-Navigationssignales. Zugriff: 06. Juni 2007. <http://www.kowoma.de/gps/Signalaufbau.htm>
- [16] **Person, Jon**. Writing Your Own GPS Applications: Part 2. Zugriff: 3. Mai 2007 <http://www.codeproject.com/netcf/WritingGPSApplications2.asp>
- [17] **Kowoma.de**. Fehlerquellen bei der GPS-Positionsbestimmung. Zugriff: 06. Juni 2007. <http://www.kowoma.de/gps/Fehlerquellen.htm>
- [18] **National Marine Electronics Association**. NMEA-0183 Standard. <http://www.nmea.org/pub/0183/index.html>

-
- [19] **NMEA.de**. NMEA Datensätze. Zugriff: 07. Mai 2007.
<http://www.nmea.de/nmea0183datensaeetze.html>
- [20] **Schönfeld, Ralph**. Das GPS-Handbuch. Münster: Verlagshaus Monsenstein und Vannerdat, 2005.
- [21] **Voss, Tim und Biedermann, Niko**. Ein elektronischer Kompass auf Basis von Feldplattensensoren. 2001. Zugriff: 10. Mai 2007. <http://www.jufo-hermannsburg.de/pdfs/2001-kompass.pdf>. Arbeitsgemeinschaft Jugend Forscht des Christian-Gymnasiums Hermannsburg. 2001
- [22] **Baillie, L., Kunczler, H., and Anegg, H.** 2005. Rolling, rotating and imagining in a virtual mobile world. In *Proceedings of the 7th international Conference on Human Computer interaction with Mobile Devices & Services* (Salzburg, Austria, September 19 - 22, 2005). MobileHCI '05, vol. 111. ACM Press, New York, NY, 283-286. DOI=
<http://doi.acm.org/10.1145/1085777.1085833>
- [23] **Wikipedia**, Kompass. Zugriff: 11. Juni 2007.
<http://de.wikipedia.org/w/index.php?title=Kompass&oldid=32978620>
- [24] **Dr. Grobe, Thomas und Dr. Schult, Thomas J.** Geodaten für Fotos. c't special: Digitale Fotografie. 2007, 02/2007. S. 140. Hannover: Heise Zeitschriften Verlag, 2007.
- [25] **Wikipedia**. Angle of View. Zugriff: 05. Juni 2007.]
http://en.wikipedia.org/w/index.php?title=Angle_of_view&oldid=135985186
- [26] **Atkins, Bob**. Field of View - Rectilinear and Fishye Lenses. Zugriff: 5. Juni 2007.
http://www.bobatkins.com/photography/technical/field_of_view.html
- [27] **Schröder, Gottfried**. Technische Fotografie. Würzburg : Vogel Verlag, 1981.
- [28] **Simon, R., Fröhlich, P., and, Anegg, H.** (2006). Beyond Location Based - The Spatially Aware Mobile Phone. 6th International Symposium on Web and Wireless Geographical Information Systems (W2GIS 2006), Hong Kong, China, December 4-5, 2006. Springer Lecture Notes in Computer Science. pp. 12-21 DOI: 10.1007/11935148_2
- [29] **Foster, Dan**. GPX: the GPS Exchange Format. <http://www.topografix.com/gpx.asp>
- [30] **Alder, Tim**. Wikipedia mit Ortsbezug. c't. 2007, 13/2007. S. 12. Hannover: Heise Zeitschriften Verlag, 2007.

Anhang

I - Vergleich von exif-Informationen dreier Digitalkameras

Wie dieser Vergleich zeigt, bieten nicht alle Kameras Angaben zu Größe des verwendeten Bildsensors und der gemessenen Entfernung des Autofokus an. Von den drei getesteten Geräten, enthielt nur die Canon Digital Ixus 700 Felder für FocalPlaneXResolution, FocalPlaneYResolution, FocalPlaneResolutionUnit und Subject Distance. Die beiden anderen Geräte, Konica KD 400Z und Kodak DX7440, bieten diese Felder nicht.

Canon Digital Ixus 700

EXIF Tag	Value
DigitalZoomRatio	1.00 x
SceneCaptureType	Standard
ExposureTime	1/60 seconds
FNumber	2.80
ExifVersion	0220
DateTimeOriginal	2007:06:22 17:25:27
DateTimeDigitized	2007:06:22 17:25:27
ComponentsConfiguration	YCbCr
CompressedBitsPerPixel	5 [bits/pixel]
ShutterSpeedValue	1/60 seconds
ApertureValue	F 2.80
ExposureBiasValue	0.00
MaxApertureValue	F 2.80
MeteringMode	Multi-segment
Flash	Flash fired, auto mode, red-eye reduction mode
FocalLength	7.70 mm
UserComment	
FlashPixVersion	0100
ColorSpace	sRGB
ExifImageWidth	2304
ExifImageHeight	3072
FocalPlaneXResolution	10816.90
FocalPlaneYResolution	10816.90
FocalPlaneResolutionUnit	Inch
SensingMethod	One-chip color area sensor
FileSource	DSC - Digital still camera

Felder zur Bildsensorgroße

EXIF Tag	Value
Sharpness	Normal
ISO Value	Auto
Metering mode	Evaluative
Focus type	Auto
AF point selected	Auto AF point selection
Exposure mode	Easy shooting
Focal length	7700 - 23100 mm (1000 mm)
Flash activity	
Flash details	Internal
Focus mode 2	Single
White Balance	Auto
Sequence number	0
Flash bias	0.00 EV
Subject Distance	160
Image Type	IMG-DIGITAL Ixus 700-JPEG
Firmware Version	Firmware Version 1.01
Image Number	1292932
Owner Name	
Thumbnail:	
Compression	6 (JPG)
XResolution	180
YResolution	180
ResolutionUnit	Inch
JpegFOffset	2300
JpegFByteCount	5129

Entfernung des Autofokus

Konica KD 400Z

EXIF Tag	Value
DigitalZoomRatio	0.00 x
FocalLengthIn35mmFilm	39 mm
SceneCaptureType	Standard
GainControl	None
Contrast	Normal
Saturation	Normal
Sharpness	Normal
SubjectDistanceRange	Distant view
ExposureTime	1/59.5 seconds
FNumber	2.80
ISOSpeedRatings	100
ExifVersion	0220
DateTimeOriginal	2006:07:20 14:14:25
DateTimeDigitized	2006:07:20 14:14:25
ComponentsConfiguration	YCbCr
CompressedBitsPerPixel	1.70 (bits/pixel)
BrightnessValue	3.60
ExposureBiasValue	0.00
MaxApertureValue	F 2.83
MeteringMode	Center weighted average
LightSource	Auto
Flash	Flash fired, auto mode
FocalLength	8.00 mm
FlashPixVersion	0100
ColorSpace	sRGB
ExifImageWidth	1704
ExifImageHeight	2304

Copy to clipboard OK

Kodak DX7440

EXIF Tag	Value
ApertureValue	F 2.83
ExposureBiasValue	0.00
MaxApertureValue	F 2.83
MeteringMode	Multi-segment
LightSource	Auto
Flash	Flash fired, auto mode
FocalLength	5.50 mm
FlashPixVersion	0100
ColorSpace	sRGB
ExifImageWidth	2304
ExifImageHeight	1728
InteroperabilityOffset	2748
ExposureIndex	140
SensingMethod	One-chip color area sensor
FileSource	Other
SceneType	Other
CustomRendered	Normal process
ExposureMode	Auto
White Balance	Auto
DigitalZoomRatio	0.00 x
FocalLengthIn35mmFilm	33 mm
SceneCaptureType	Standard
GainControl	Low gain up
Contrast	Normal
Saturation	Normal
Sharpness	Normal
SubjectDistanceRange	Unknown

Copy to clipboard OK

II - XML und JSON Anfrage an das Wikipedia Webservice

Gesucht war der deutsche Eintrag über die „Wiener Karlskirche“. Die beiden Anfragen zeigen, welche Resultate das Wikipedia Webservice dazu liefern kann.

Anfrage XML:

```
http://ws.geonames.org/wikipediaSearch?q=karlskirche&lang=de&maxRows=1
```

Resultat als XML-Dokument:

```
<entry>
  <lang>de</lang>
  <title>Wiener Karlskirche</title>
  <summary>
    es in Wien, am Rande der Innenstadt und 200m außerhalb der
    Ringstraße. Sie ist einer der bedeutendsten barocken
    (barockklassizistischen) Kirchenbauten nördlich der Alpen und
    hat als spektakuläre Besonderheit eine Kuppel in Form eines
    verlängerten Ellipsoids. Seit der Karlsplatz Ende der 1980er
    wieder als Ensemble hergestellt wurde, wirkt die Karlskirche
    nicht nur durch ihre Kuppel und die (...)
  </summary>
  <feature/>
  <population>0</population>
  <elevation>0</elevation>
  <lat>48.1982</lat>
  <lng>16.3714</lng>
  <wikipediaUrl>
    http://de.wikipedia.org/wiki/Wiener_Karlskirche
  </wikipediaUrl>
  <thumbnailImg>
    http://www.geonames.org/img/wikipedia/1000/thumb-647-100.jpg
  </thumbnailImg>
</entry>
```

Anfrage JSON:

```
http://ws.geonames.org/wikipediaSearchJSON?q=karlskirche&lang=de&maxRows=1
```

Resultat als JSON-Dokument:

```
{"geonames":[{"summary":"es in Wien, am Rande der Innenstadt und
200m außerhalb der Ringstraße. Sie ist einer der bedeutendsten
barocken (barockklassizistischen) Kirchenbauten nördlich der Alpen
und hat als spektakuläre Besonderheit eine Kuppel in Form eines
verlängerten Ellipsoids. Seit der Karlsplatz Ende der 1980er wieder
als Ensemble hergestellt wurde, wirkt die Karlskirche nicht nur
durch ihre Kuppel und die (...)", "title":"Wiener
Karlskirche", "wikipediaUrl":"de.wikipedia.org/wiki/Wiener_Karlskirch
e", "elevation":0, "lng":16.3714222222222, "feature":"","thumbnailImg":
"http://www.geonames.org/img/wikipedia/1000/thumb-647-
100.jpg", "lang":"de", "lat":48.1982472222222, "population":0}]}
```

III - Wintec WBT-100 - Technisches Datenblatt

http://www.wintec-gps.de/wintec_wbt-100_techdata.php, 20.07.2007

Allgemeine Daten:

Kanäle	12
Chipsatz	u-Nav + iTrax 3. Generation
Prozessor Typ	ARM7TDMI
Empfangs Code	L1, C/A Code (SPS)
Empfangsfrequenz	1575.42 MHz
Empfangsverstärkung	eingebauter elektromagnetischer Kompass
Gewicht	ca. 48g
max. Update Rate	bis 4Hz (1Hz im Auslieferungszustand)
Größe	60.0 x 38.0 x 16.0 mm
Lagertemperatur	-40°C bis +85°C
Betriebstemperatur	-10°C bis +70°C
Leistungsaufnahme	< ca. 45mA

Aktualisierungszeiten:

Blitz Start	ca. 0,1 Sek.
Heiß Start	ca. 8 Sek.
Warm Start	ca. 38 Sek.
Kalt Start	ca. 45 Sek.
Automatische Update-Zeit	alle 1 Sek.
Wiederaufnahme bei Verlust	ca. 0.1 Sek.

Positionsgenauigkeit:

DGPS	-
WAAS / EGNOS	möglich

Dynamik:

Dyn. Höhe	ca. -1km bis 18km
Dyn. Geschwindigkeit	515m / Sek.
Beschleunigung	± Maximum 4G

Antenne:

eingebaute Antenne	Patch Antenne
Empfindlichkeit	ca. -152dBm

Protokoll:

NMEA Protokoll	NMEA 0183 V3.0 ASCII Ausgang
NMEA Format	GGA, GSA, GSV, RMC
NMEA Eingang	NMEA/SiRF Wahl; Höhe; Position, Datum & Zeit
optionales Format	GLL, VTG, ZDA
Baud Rate Bluetooth	38400 bps
Baud Rate Kabel	4800-9600 bps

Interface:

Bluetooth	Version 1.1, Klasse 2
-----------	-----------------------

Bluetooth Dienst	SPP Slave
Bluetooth Reichweite	bis 10m
Frequenz	2402 - 2480 MHz
Signalempfindlichkeit	-80dBm
Anschlüsse	Mini USB

Log Funktionen

Interner Speicher	für bis zu ca. 25360 Positionen
Daten Speicherung	Breiten- & Längengrad sowie Zeit & Datum
Daten Export	Über Bluetooth & Kabel möglich
Unterstützte Betriebssysteme	Windows® OS

LED- & Tasten Funktionen:

LED Anzeige	4 LED
Blau	Verbindungsbereitschaft über Bluetooth
Rot	GPS + Kompass
Grün	Akkuanzeige

Stromversorgung:

Ladungsspannung	+3.75 ~ +5.5V DC
externe Stromquelle	12-26V, 220V, 110V
Akku wechselbar	Mehr als 10 Stunden Dauerbetrieb mit Lithium-Ionen Akku bis zu 14 Stunden bei Logging